Sequential Detection of Parameter Changes in Dynamic Conditional Correlation Models

K. Pape, P. Galeano and D. Wied

TU Dortmund, Universidad Carlos III de Madrid, Universität zu Köln

Abstract

A multivariate monitoring procedure is presented to detect changes in the parameter vector of the Dynamic Conditional Correlation (DCC) model. The procedure can be used to detect changes in both the conditional and unconditional variances as well as in the correlation structure of the model. The detector is based on the contributions of individual observations to the gradient of the quasi-log-likelihood function. More precisely, standardized derivatives of quasi-log-likelihood contributions at time points in the monitoring period are evaluated at parameter estimates calculated from a historical period. The null hypothesis of a constant parameter vector is rejected if these standardized terms differ too much from zero. Critical values are obtained via a parametric bootstrap type procedure. Size and power properties of the procedure are examined in a simulation study. Finally, the behavior of the proposed monitoring scheme is illustrated with a group of asset returns.

Keywords: Dynamic Conditional Correlation model; Multivariate sequences; Online detection; Parameter changes; Threshold function.

JEL Classification: C12; C32; C58.

Acknowledgements: Financial support by Deutsche Forschungsgemeinschaft (SFB 823, project A1) and MINECO/FEDER/UE Grant ECO2015-66593-P are gratefully acknowledged.

1 Introduction

In recent years a lot of research has been focused on modeling volatilities and correlations as well as on testing for structural breaks. Research in the intersection between these fields is motivated by the importance of being informed about changes in the variances and covariances or in the parameters that determine these characteristics, as soon as possible after their occurrence. In particular, in financial applications, analysts may need the aforementioned information to construct optimal portfolios or to anticipate crises since volatilities and correlations tend to increase in turbulent market phases, see for instance Sandoval Jr. and De Paula Franca (2012) or Charles and Darné (2014).

While former monitoring methods for multiple asset returns often focus either on variances or correlations, see for instance, Wied and Galeano (2013) and Pape et al. (2016), among others, we aim at monitoring structural changes in both volatilities and correlations *jointly*. For this purpose, we consider the popular Dynamic Conditional Correlation (DCC) model by Engle (2002) and provide a method to monitor its parameters which steer the conditional and unconditional volatilities and correlations.[1] So, in contrast to Wied and Galeano (2013) and Pape et al. (2016), who propose methods which do not use a specific model assumption, our approach is model-based. This could in principle lead to efficiency gains as long as the model assumptions fit the data.

If the parameters are not constant in the observed period, then parameter estimates based on the constancy assumption are no longer reliable, as they yield biased volatility and correlation forecasts. Furthermore, the sum of the estimated autoregressive parameters of the conditional variance converges to one in univariate Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models if a parameter change is ignored, see Hillebrand (2005). This result should be kept in mind while dealing with multiple return series.

In contrast to the technique of Aue et al. (2009), who use a retrospective method to detect changes in the covariance structure of multivariate time series, we prefer a sequential monitoring procedure. That is, based on an historical period of observations, we obtain new data points bit by bit and

[1] This means that we do not consider models such as Audrino and Trojani (2011) in which time-varying parameters are included from the beginning on.

use them to refresh our detector in order to determine the presence of a changepoint in the model parameters as soon as possible once it has happened. Our approach is motivated by the work of Chu et al. (1996) and Berkes et al. (2004). First, Chu et al. (1996) suggest to use the information of a historical sample, that is entirely available when the monitoring starts and is assumed to be free from structural breaks. Then, a size controlled sequential test is developed to check the structural stability of linear models based on cumulated recursive residuals. Building on this procedure, Berkes et al. (2004) propose a sequential monitoring scheme to detect changes in the parameters of the GARCH model. For this purpose, a detector that depends on quasi-likelihood scores is used. The historical sample is used to estimate the model parameters, that are used to evaluate the contributions of the data from the monitoring period to the Gaussian quasi-log-likelihood (QLL) function. Hence, under the alternative of a parameter change in the monitoring period, it is expected that the absolute gradient contributions of post-break observations tend to infinity.

The procedure proposed in this paper is used to monitor changes in the parameters of the DCC model and can be seen as a multivariate extension of the monitoring scheme proposed by Berkes et al. (2004). Nonetheless, the extension is much more complex than it may seem. Models that allow for dynamic modelling of both the variances and correlations possess a far more complex structure than other multivariate extensions of the univariate GARCH model. The challenge of handling the model and its quasi-likelihood scores gets even more demanding if a multiplicative structure of the conditional covariance matrix is postulated as in the DCC model. Also, the DCC models and their properties are far less well investigated than univariate GARCH models and especially the classical GARCH model considered by Berkes et al. (2004). For the models with dynamic conditional correlations, important results like conditions for the existence and uniqueness of a stationary solution or for the existence of unconditional moments of higher order have just been proposed recently, see Fermanian and Malongo (2017), or remain to be established, which makes this type of model quite challenging in applications.

Even if we focus on the DCC model due to its enormous popularity for modeling multiple financial returns, the results of this paper may be extended to models with structure similar to the one of the DCC model of Engle (2002), e.g., the Constant Conditional Correlation (CCC) model of

Bollerslev (1990), the Varying Conditional Correlation (VCC) model of Tse and Tsui (2002) and the Asymmetric Generalized Dynamic Conditional Correlation (AG-DCC) model of Capiello et al. (2006), among others. On the contrary, the extension to other popular multivariate volatility models, e.g., the multivariate extensions of the GARCH models as proposed by Bollerslev et al. (1988) or the BEKK model proposed by Engle and Kroner (1995), that ensures the nonnegative definiteness of the conditional covariance matrix under milder conditions on the parameters, is more complex as the structure of these models is quite different than the structure of dynamic correlation models.

The rest of the paper is organized as follows. Section 2 briefly introduces the DCC model proposed by Engle (2002), presents necessary assumptions for the existence of a unique stationary solution of the DCC model, and explains quasi maximum likelihood estimation (QMLE) of the DCC model parameters. Then, Section 3 presents the monitoring problem and the proposed monitoring scheme. The performance of the procedure in finite time samples is investigated with the help of simulations and an application to a time series of returns in Sections 4 and 5, respectively. Finally, some concluding statements can be found in Section 6.

2 The Dynamic Conditional Correlation Model

2.1 The Model and Basic Assumptions

Let $\{y_t, t \in \mathbb{Z}\}$ be a sequence of p dimensional random vectors, $y_t = (y_{1t}, \ldots, y_{pt})'$, following a multivariate GARCH model given by

$$y_t = H_t^{1/2} \epsilon_t \tag{2.1}$$

where

$$H_t = \mathsf{Cov}\left(y_t | \mathcal{F}_{t-1}\right) \tag{2.2}$$

is the positive definite conditional covariance matrix of y_t given the information set $\mathcal{F}_{t-1} = \sigma\left\{y_{t-1}, y_{t-2}, \ldots\right\}$ and $\{\epsilon_t, t \in \mathbb{Z}\}$ a standard white noise sequence in $\mathbb{R}^p$, i.e. $E\left(\epsilon_t\right) = \mathbf{0}_p$, $\mathsf{Cov}\left(\epsilon_t\right) = \mathbb{I}_p$, $\forall t \in \mathbb{Z}$, and the vectors ϵ_t are mutually independent. In the following, $\mathbf{0}_p$, $\mathbf{0}_{p\times p}$ and $\mathbb{I}_p$ denote the

p dimensional vector of zeros, the $(p \times p)$ dimensional matrix of zeros, and the $(p \times p)$ dimensional identity matrix, respectively.

Of all the available specifications of the conditional covariance matrix H_t, we focus on the one presented by Engle (2002) for its high popularity. That is, we assume

$$H_t = D_t R_t D_t \tag{2.3}$$

with $D_t = \mathsf{diag}\left\{h_{1t}^{1/2}, \ldots, h_{pt}^{1/2}\right\}$, where h_{it}, $i = 1, \ldots, p$, are the individual variances, that can be specified for instance according to univariate GARCH(1,1) models:

$$h_{it} = \omega_i + \alpha_i y_{i,t-1}^2 + \beta_i h_{i,t-1}, \quad i = 1, \ldots, p, \tag{2.4}$$

for certain parameters ω_i, α_i and β_i. Furthermore, $R_t := \mathsf{Cor}\left(y_t | \mathcal{F}_{t-1}\right)$ is the conditional correlation matrix of y_t, which can be decomposed as

$$R_t = Q_t^* Q_t Q_t^*, \tag{2.5}$$

where Q_t is a $(p \times p)$ matrix that is recursively determined as

$$Q_t = (1 - \alpha - \beta)\,\bar{Q} + \alpha z_{t-1} z_{t-1}' + \beta Q_{t-1} \tag{2.6}$$

with $z_t = D_t^{-1} y_t$ the standardized vectors. The parameters α and β are nonnegative scalars, which satisfy $\alpha + \beta < 1$. $\bar{Q} = [\bar{q}_{ij}]_{i,j=1,\ldots,p}$ is both the unconditional covariance and correlation matrix of z_t in the special case of constant conditional correlations, see Aielli (2013). Motivated by this, we impose the restriction that the main diagonal elements are one, which is common in the literature. Consequently, the unknown parameters in the matrix $\bar{Q}$ are the entries of $\psi = \mathsf{vecl}\left(\bar{Q}\right) = (\bar{q}_{21}, \ldots, \bar{q}_{p,p-1})'$, where $\mathsf{vecl}(\cdot)$ is the operator that stacks the lower diagonal elements of a matrix into a vector. Finally, the normalizing matrix Q_t^* is given by

$$Q_t^* := \mathsf{diag}\left\{[Q_t]_{11}^{-1/2}, \ldots, [Q_t]_{pp}^{-1/2}\right\}$$

where $[Q_t]_{ii}$ denotes the i-th main diagonal entry of the matrix Q_t, for $i = 1, \ldots, p$.
In summary, the vector of parameters of the DCC model is given as

$$\boldsymbol{\theta} = (\omega_1, \alpha_1, \beta_1, \ldots, \omega_p, \alpha_p, \beta_p, \alpha, \beta, \bar{q}_{21}, \ldots, \bar{q}_{p,p-1})'$$

which leads to a total number of $d := \frac{1}{2}(p+1)(p+4)$ unknown parameters in the model. Note that $\boldsymbol{\theta}$ can be decomposed into $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')'$, where

$$\boldsymbol{\theta}_1 = (\omega_1, \alpha_1, \beta_1, \ldots, \omega_p, \alpha_p, \beta_p)' = \left(\phi_1', \ldots, \phi_p'\right)'$$

with $\phi_i := (\omega_i, \alpha_i, \beta_i)'$, for $i = 1, \ldots, p$, is the vector of variance parameters and

$$\boldsymbol{\theta}_2 = (\alpha, \beta, \bar{q}_{21}, \ldots, \bar{q}_{p,p-1})' = (\alpha, \beta, \psi')',$$

is the vector of correlation parameters.
An important issue in multivariate models with dynamic variances and correlations is that the positive definiteness of the conditional covariance matrix H_t has to be guaranteed for all $t \in \mathbb{Z}$ almost surely. Proposition 2 in Engle and Sheppard (2001) gives sufficient conditions for this property. Particularly, the matrix H_t, as specified in (2.3)-(2.6), is positive definite for all $t \in \mathbb{Z}$ almost surely, if Assumption 2.1 is satisfied:

Assumption 2.1. *1.* $\omega_i > 0, \ \forall i \in \{1, \ldots, p\}$.

2. $\alpha_i > 0$ *and* $\beta_i > 0$ *with* $\alpha_i + \beta_i < 1, \forall i \in \{1, \ldots, p\}$, *see also Nelson and Cao (1992).*

3. $h_{i0} > 0, \ \forall i \in \{1, \ldots, p\}$.

4. $\alpha > 0$ *and* $\beta > 0$ *with* $\alpha + \beta < 1$.

5. There exists $\delta_1 > 0$ *with* $\lambda_{\mathsf{min}}(\bar{Q}) > \delta_1$, *where* $\lambda_{\mathsf{min}}(\cdot)$ *is the smallest eigenvalue of a square matrix.*

Particularly, Assumption 2.1 ensures that Q_t is positive definite for all $t \in \mathbb{Z}$ because of the

decomposition

$$Q_t = \frac{1-\alpha-\beta}{1-\beta}\bar{Q} + \alpha\sum_{n=0}^{\infty}\beta^n z_{t-n-1}z'_{t-n-1}$$

and 6.70.(a) in Seber (2008):

$$\begin{aligned}\lambda_{\mathsf{min}}\left(Q_t\right) &\overset{a.s.}{\geq} \lambda_{\mathsf{min}}\left(\frac{1-\alpha-\beta}{1-\beta}\bar{Q}\right) + \lambda_{\mathsf{min}}\left(\alpha\sum_{n=0}^{\infty}\beta^n z_{t-n-1}z'_{t-n-1}\right)\\ &\overset{a.s.}{\geq} \frac{1-\alpha-\beta}{1-\beta}\lambda_{\mathsf{min}}\left(\bar{Q}\right) > \frac{1-\alpha-\beta}{1-\beta}\delta_1 > 0.\end{aligned}$$

Let $p^- := \frac{1}{2}p\left(p-1\right)$ be the number of unknown parameters in the constant matrix $\bar{Q}$. Analogously to Berkes et al. (2003) and Berkes et al. (2004), we consider the parameter space U,

$$\begin{aligned}U := \Big\{u :\ &\mathsf{max}\left\{t_1,\ldots,t_p,b,a+b,\left|q_1\right|,\ldots,\left|q_{p^-}\right|\right\} \leq \rho,\ \lambda_{\mathsf{min}}\left(\mathsf{F}_{\bar{Q}}(u)\right) > \delta_1,\\ &\text{and }\ \underline{u} < \mathsf{min}\left\{x_1,s_1,t_1,\ldots,x_p,s_p,t_p,a,b\right\} \leq \mathsf{max}\left\{x_1,s_1,t_1,\ldots x_p,s_p,t_p,a,b\right\} \leq \overline{u}\Big\}\end{aligned}$$

where $u = \left(x_1,s_1,t_1,\ldots,x_p,s_p,t_p,a,b,q_1,\ldots,q_{p^-}\right)'$ is a generic element of the constrained parameter space U and

$$\mathsf{F}_{\bar{Q}}(u) := \begin{pmatrix} 1 & q_1 & q_2 & \ldots & q_{p-1}\\ q_1 & 1 & q_p & \ldots & \vdots\\ \vdots & q_p & \ddots & & q_{p^--1}\\ \vdots & \vdots & & 1 & q_{p^-}\\ q_{p-1} & \ldots & q_{p^--1} & q_{p^-} & 1 \end{pmatrix}.$$

Moreover, $0 < \underline{u} < \overline{u}$ and $0 < \rho < 1$ are fixed constants. Throughout the paper, we assume that the true parameter $\boldsymbol{\theta}$ (both before and after potential breaks) is contained in this parameter space.

2.2 Existence of a Unique Stationary Solution

To verify the existence of a stationary and unique solution satisfying the DCC model, some additional assumptions have to be imposed. More precisely, as in Fermanian and Malongo (2017), the

model equations (2.1)-(2.6) can be written in autoregressive form as

$$X_t = T_t X_{t-1} + \zeta_t \tag{2.7}$$

with

$$X_t := \left(h_{1t}, \ldots, h_{pt}, y_{1t}^2, \ldots, y_{pt}^2, \mathsf{vecl}\left(Q_t\right)', \mathsf{vecl}\left(z_t z_t'\right)'\right)',$$
$$\zeta_t := \left(\omega_1, \ldots, \omega_p, \omega_1 z_{1t}^2, \ldots, \omega_p z_{pt}^2, (1-\alpha-\beta)\,\mathsf{vecl}\left(\bar{Q}\right)', \mathsf{vecl}\left(z_t z_t'\right)'\right)'$$

and

$$T_t = \left[\begin{array}{ccc|ccc|cc} \beta_1 & & 0 & \alpha_1 & & 0 & & \\ & \ddots & & & \ddots & & \mathbf{0}_{p\times p^-} & \mathbf{0}_{p\times p^-} \\ 0 & & \beta_p & 0 & & \alpha_p & & \\ \hline \beta_1 z_{1t}^2 & & 0 & \alpha_1 z_{1t}^2 & & 0 & & \\ & \ddots & & & \ddots & & \mathbf{0}_{p\times p^-} & \mathbf{0}_{p\times p^-} \\ 0 & & \beta_p z_{pt}^2 & 0 & & \alpha_p z_{pt}^2 & & \\ \hline \multicolumn{3}{c|}{\mathbf{0}_{p^-\times p}} & \multicolumn{3}{c|}{\mathbf{0}_{p^-\times p}} & \beta\,\mathbb{I}_{p^-} & \alpha\,\mathbb{I}_{p^-} \\ \hline \multicolumn{3}{c|}{\mathbf{0}_{p^-\times p}} & \multicolumn{3}{c|}{\mathbf{0}_{p^-\times p}} & \mathbf{0}_{p^-\times p^-} & \mathbf{0}_{p^-\times p^-} \end{array}\right].$$

Denote as $\boldsymbol{\mathcal{I}} := \{\mathcal{I}_t, t \in \mathbb{Z}\}$ the filtration of the process $\{X_t, t \in \mathbb{Z}\}$ in (2.7), i.e., $\mathcal{I}_t = \sigma\left(X_t, X_{t-1}, \ldots\right)$ is the information set at time t. In the following, consider the process $\{\eta_t, t \in \mathbb{Z}\}$ with $\eta_t := R_t^{-1/2} z_t$. Conditional upon the information up to time $t-1$, η_t behaves like an innovation vector and will be called conditional innovation. Analogously to Assumptions A0 and U0 and Assumptions U1-U3 in Fermanian and Malongo (2017), we assume:

Assumption 2.2. *The process $\{\eta_t, t \in \mathbb{Z}\}$ possesses the Markov property with respect to the filtration $\boldsymbol{\mathcal{I}}$. In particular, we have $\mathsf{E}\left(\eta_t | \mathcal{I}_{t-1}\right) = \mathsf{E}\left(\eta_t | X_{t-1}\right), \forall t \in \mathbb{Z}$. Furthermore, $\{\eta_t, t \in \mathbb{Z}\}$ is ergodic and stationary.*

Assumption 2.3. $\max\limits_{1\le i\le p} \alpha_i + \max\limits_{1\le i\le p} \beta_i < 1 \quad$ *and* $\quad |\beta| < 1.$

Then, under assumptions 2.2 and 2.3, Fermanian and Malongo (2017) have shown that the strict

stationarity of $\{X_t,\ t \in \mathbb{Z}\}$ and thus of the DCC process $\{y_t,\ t \in \mathbb{Z}\}$ can be obtained with Tweedie's Theorem, see Tweedie (1988), which implies the existence of a time invariant measure for the transition probabilities of the linear Markov chain given by (2.7). In addition, the uniqueness of the stationary solution can be obtained under the following assumption, which is equivalent to (14) in Fermanian and Malongo (2017):

Assumption 2.4.

$$\mathsf{E}\left[ln\left(\beta^2 + \alpha^2 \frac{4(2p+1)\sqrt{p}}{\sqrt{C_\lambda} C_q} \left\|\eta_t\right\|_2^2\right)\right] < 0,$$

where $C_\lambda = \frac{(1-\alpha-\beta)\lambda_{\min}(\bar{Q})}{1-\beta^2}$ *and* $C_q = \frac{(1-\alpha-\beta)\min\limits_{1\leq i\leq p} \bar{q}_{ii}}{1-\beta^2}$ *are constants.*

Assumption 2.4 brings up a restriction for the variance of η_t given the model parameters. It is more likely fulfilled the smaller α and β are. As Fermanian and Malongo (2017) state, the expectation can be simulated. So, it would be possible to check whether the assumption might be fulfilled by estimating the model parameters given the distribution of η_t, simulating the expectation and checking whether it is smaller than 0.

2.3 Estimation of the Model Parameters

Given an observed multivariate time series $y_1, \ldots, y_T$, the QMLE of $\boldsymbol{\theta}$ is a consistent parameter estimator that is obtained by maximizing the Gaussian QLL function

$$L_T(\boldsymbol{\theta}) := L_T\left(\boldsymbol{\theta}|y_1, \ldots, y_T\right) = \sum_{t=1}^{T} l_t\left(\boldsymbol{\theta}|y_1, \ldots, y_T\right), \tag{2.8}$$

with individual QLL contributions

$$l_t(\boldsymbol{\theta}) := l_t\left(\boldsymbol{\theta}|y_1, \ldots, y_T\right) = -\frac{1}{2}\left(p \cdot \log 2\pi + \log\det\left(H_t\right) + y_t' H_t^{-1} y_t\right). \tag{2.9}$$

Direct computation of the QMLE in one step is computationally expensive even for moderate dimensions of y_t. Alternatively, Engle and Sheppard (2001) proposed a two-step QMLE estimator

to reduce the calculation time[2]. The two-step one is based on maximizing the part of the likelihood that only depends on the volatility parameters, $\boldsymbol{\theta}_1$, and, after plug-in the estimates of these parameters and of the matrix $\bar{Q}$, maximize the rest of the likelihood that will only depends on the correlation parameters, $\boldsymbol{\theta}_2$. The step of estimating $\bar{Q}$ is called variance targeting, but this does not work in general, so that this procedure is in general not consistent (Aielli, 2013). Indeed, preliminary simulations showed that the one step QMLE yields distinctly better estimates than the two step one for the considered parameters values. Hence, we use the one step QMLE for our simulations and applications.[3] In the following, we denote this estimator calculated from a sample of T observations as $\hat{\theta}_T$.

3 The Monitoring Procedure

Next, we are interested in the hypothesis of model parameter stability of the sequence $\{y_t, t \in \mathbb{Z}\}$. For that, let $\theta_t \in \mathbb{R}^d$ be the parameter vector of the DCC model at time t. Assume that we have observed a sequence of y_t of length m, $y_1, \ldots, y_m$, that is not affected by any structural change over such historical period, i.e.:

Assumption 3.1. $\theta_1 = \ldots = \theta_m$ *with* m *a positive integer.*

Then, we are interested in testing the null hypothesis of a constant parameter vector

$$\mathsf{H}_0 : \ \theta_t = \boldsymbol{\theta}, \quad t = 1, \ldots, m, m+1, \ldots$$

against the alternative of a change in the vector of parameters at an unknown point in the monitoring period

$$\mathsf{H}_1 : \ \theta_t = \begin{cases} \boldsymbol{\theta}, & t = 1, \ldots, m, m+1, \ldots, m+k^*-1 \\ \boldsymbol{\theta^*}, & t = m+k^*, m+k^*+1, \ldots \end{cases}$$

[2]There is an alternative approach for estimating the model equation by equation (Francq and Zakoïan, 2016), which we have not used.

[3]Another possibility would have been to consider the cDCC-model proposed by Aielli (2013) and to use the estimator proposed there.

with $\boldsymbol{\theta} = \left(\phi_1', \ldots, \phi_p', \alpha, \beta, \psi'\right)'$ the parameter vector before and $\boldsymbol{\theta^*} = \left(\phi_1^{*\prime}, \ldots, \phi_p^{*\prime}, \alpha^*, \beta^*, \psi^{*\prime}\right)'$ the parameter vector after the change, where $\phi_i^* = (\omega_i^*, \alpha_i^*, \beta_i^*)'$, $i = 1, \ldots, p$, and $\psi^* = \mathsf{vecl}\left(\bar{Q}^*\right)$ with $\bar{Q}^* = \left[\bar{q}_{ij}^*\right]_{i,j=1,\ldots,p}$. Note that the change takes place at the k^*-th point of the monitoring period which is the $(m + k^*)$-th point in the entire time series.

To test H_0 against H_1, we will make use of a monitoring scheme. One option is to construct a procedure similar to the method proposed by Berkes et al. (2004) for univariate GARCH models. We take initially this option and then denote as $\nabla l_t(\theta)$ the gradient of the QLL contributions $l_t(\theta)$ with infinite past in (2.9) and $\nabla \hat{l}_t(\theta)$ the gradient with finite past, i.e. the gradient computed with initial value of the sequence y_0 and initial volatilities $h_{10}, \ldots, h_{p0}$.

For a more appropriate comparison of the simulation results for different lengths of the historical period, we denote the length of the monitoring period as mB. Therefore, B indicates how long the monitoring period is compared to the historical period. Furthermore, for any θ, define

$$\boldsymbol{D}(\theta) := \mathsf{E}\nabla l_t(\theta)\nabla l_t(\theta)'$$

and assume:

Assumption 3.2. $\boldsymbol{D} := \boldsymbol{D}(\boldsymbol{\theta})$ *is a finite and nonsingular matrix.*

For the observed sequence, $y_1, \ldots, y_m$, $\boldsymbol{D}$ can be estimated with:

$$\widehat{\boldsymbol{D}}_m = \frac{1}{m}\sum_{t=1}^{m} \nabla \hat{l}_t(\hat{\theta}_m)\nabla \hat{l}_t(\hat{\theta}_m)'$$

and the monitoring procedure can be based on the detector:

$$V_k = \sum_{t=m+1}^{m+k} \widehat{\boldsymbol{D}}_m^{-\frac{1}{2}} \nabla \hat{l}_t(\hat{\theta}_m)$$

with stopping rule

$$\tau_m = \mathsf{min}\left\{k \leq mB : \ |V_k| > m^{\frac{1}{2}}\left(1 + \frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)\right\}, \tag{3.1}$$

where $\mathsf{b}(\cdot)$ is a threshold function and $|\cdot|$ the norm that yields the maximum absolute entry of a vector and matrix. If $\tau_m < \infty$, a change in the parameters is indicated at some time point between $m+1$ and $m+\tau_m$. If the detector did not cross the threshold function in the monitoring period and no changepoint could be detected, the set on the right side of (3.1) would be the empty set and $\tau_m = \infty$ as $\mathsf{min}\ \emptyset = \infty$. As in Berkes et al. (2004), some moderate conditions are imposed on the form of the threshold function $\mathsf{b}(\cdot)$:

Assumption 3.3. $\mathsf{b}(\cdot)$ *is continuous on* $(0,\infty)$ *and* $\inf\limits_{0<t<\infty} \mathsf{b}(t) > 0$.

To avoid confusion with the model parameters, let $\tilde{\alpha} \in (0,1)$ be the significance level for testing the null hypothesis of no parameter change versus the alternative hypothesis of a change during the monitoring period. Therefore, the threshold function $\mathsf{b}(\cdot)$ or at least the variable parts of the function should be chosen such that

$$\lim_{m\to\infty} \mathsf{P}_{\mathsf{H}_0}\{\tau_m < \infty\} = \tilde{\alpha} \quad \text{and} \quad \lim_{m\to\infty} \mathsf{P}_{\mathsf{H}_1}\{\tau_m < \infty\} = 1.$$

Berkes et al. (2004) choose the threshold function $\mathsf{b}(\cdot)$ as a constant that is obtained via simulation. Preliminary simulations suggested that the empirical size of the proposed multivariate procedure depends strongly on the length of the monitoring period, that is on the parameter B, just as in the univariate case presented by Berkes et al. (2004). To reduce this effect, we include the length of the monitoring period into the stopping rule (3.1). Moreover, we prefer a curved threshold function to the linear one that results from choosing $\mathsf{b}(\cdot)$ as a constant function. In detail, we use the one proposed by Horváth et al. (2004) and also used by Wied and Galeano (2013) among others, i.e.

$$\mathsf{b}(x) = \mathsf{max}\left\{\left(\frac{x}{1+x}\right)^{\gamma}, \varepsilon\right\}$$

where $\gamma \in [0, 1/2)$ is a tuning parameter and $\varepsilon > 0$ a constant, that can be chosen arbitrarily small in applications. A larger value of γ results in a steeper threshold function that tends to detect early changes in the parameters with a higher probability. In contrast, a smaller value of the tuning parameter leads to a lower slope of the threshold function, which results in a higher probability to

detect changes that occur later in the monitoring period.

Additionally, we scale this threshold function by multiplying a constant $c = c(\tilde{\alpha})$ that is obtained via Monte Carlo simulations, such that the probability that the detector crosses the threshold function in the monitoring period equals the theoretical size $\tilde{\alpha}$.

Under several high-level assumptions, it is possible to derive a theorem regarding the asymptotic distribution of the detector. The proof is deferred to the appendix. Pape et al. (2017) shows that these high-level assumption also hold under lower-level assumptions, but the proofs need the assumption of bounded innovations.

Assumption 3.4. *1. The limit function $L(\cdot)$ of $L_m(\cdot)$ has a unique maximum in $\boldsymbol{\theta}$.*

2. For all $t = 1, \ldots, mB$, it holds

$$\mathsf{E}\left[\sup_{u \in U} l_t(u)\right] < \infty.$$

3. For all $t = 1, \ldots, mB$, the matrix of second derivatives of $l_t(u)$, $\nabla^2 l_t(u)$, exists and it holds

$$\mathsf{E}\left[\sup_{u \in U} \nabla^2 l_t(u)\right] < \infty.$$

4. Let $\mathbf{A}(\boldsymbol{\theta}) := \lim_{m \to \infty} \mathsf{E}\left[\frac{1}{m} \frac{\partial^2 L_m(\theta)}{\partial \theta \partial \theta'}\Big|_{\theta = \boldsymbol{\theta}}\right]$ be a finite and nonsingular matrix and $\{\theta_m^*, m \in \mathbb{N}\}$ a sequence with $\theta_m^* \xrightarrow{p} \boldsymbol{\theta}$ for $m \to \infty$. For $m \to \infty$, we have:

$$\frac{1}{m} \frac{\partial^2 L_m(\theta)}{\partial \theta \partial \theta'}\bigg|_{\theta = \theta_m^*} \xrightarrow{p} \mathbf{A}(\boldsymbol{\theta}).$$

5. Let $\mathbf{B}(\boldsymbol{\theta}) := \lim_{m \to \infty} \mathsf{E}\left[\frac{1}{m} \frac{\partial L_m(\theta)}{\partial \theta}\Big|_{\theta = \boldsymbol{\theta}} \frac{\partial L_m(\theta)}{\partial \theta'}\Big|_{\theta = \boldsymbol{\theta}}\right]$ be a finite and nonsingular matrix. For $T \to \infty$, we have:

$$\frac{1}{\sqrt{m}} \frac{\partial L_m(\theta)}{\partial \theta} \xrightarrow{d} \mathcal{N}(\mathbf{0}_d, \mathbf{B}(\boldsymbol{\theta})).$$

6. The estimator $\widehat{\boldsymbol{D}}$ is consistent for D.

7. It holds

$$\sup_{u \in U} \left| \sum_{i=1}^{n} \left(\nabla \hat{l}_i(u) - \nabla l_i(u) \right) \right| \overset{a.s.}{=} \mathcal{O}(1).$$

Theorem 3.1. *Under the null hypothesis and Assumptions 3.1, 3.2, 3.3 and 3.4 it holds*

$$\lim_{m\to\infty}\mathsf{P}_{\mathsf{H}_0}\left\{\tau_m<\infty\right\}=\lim_{m\to\infty}\mathsf{P}_{\mathsf{H}_0}\left(\sup_{1<k\leq mB}\frac{|V_k|}{m^{1/2}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)}\geq c\right)=\mathsf{P}_{\mathsf{H}_0}\left(\sup_{t\in(0,B]}\frac{|\mathcal{G}(t)|}{(1+t)\,\mathsf{b}\,(t)}\geq c\right)$$

where $\left\{\mathcal{G}(t)=(G_1(t),\ldots,G_d(t))',\ t\in[0,B]\right\}$ *is a d-variate stochastic process whose component processes are d independent mean zero Gaussian processes* $\{G_j(t),t\in[0,B]\}$ *with covariance function* $\mathsf{E}\left(G_j(k)G_j(l)\right)=\mathsf{min}\left\{k,l\right\}+kl$*, for* $j=1,\ldots,d$*, where d is the number of parameters in the DCC model.*

Along the lines of Berkes et al. (2004) or Galeano and Wied (2014) and denoting $\{W_i(t),\ t\in[0,\infty)\}$ for $i=1,\ldots,d$ as d independent one dimensional standard Brownian motions, we have that $|\mathcal{G}(t)|$ possesses the same distribution as $\underset{1\leq i\leq p}{\mathsf{max}}\left|(1+t)W_i\left(\frac{t}{1+t}\right)\right|$ for all $t\in\mathbb{Z}$, which yields

$$\begin{aligned}\sup_{t\in(0,B]}\frac{|\mathcal{G}(t)|}{(1+t)\mathsf{b}(t)}&\overset{d}{=}\sup_{t\in(0,B]}\underset{1\leq i\leq d}{\mathsf{max}}\frac{\left|W_i\left(\frac{t}{1+t}\right)\right|}{\mathsf{b}\,(t)}\overset{d}{=}\sup_{t\in(0,B]}\underset{1\leq i\leq d}{\mathsf{max}}\frac{\left|W_i\left(\frac{t}{1+t}\right)\right|}{\mathsf{max}\left\{\left(\frac{t}{1+t}\right)^{\gamma},\varepsilon\right\}}\\&\overset{d}{=}\sup_{\tilde{s}\in\left(0,\frac{B}{1+B}\right]}\underset{1\leq i\leq d}{\mathsf{max}}\frac{|W_i\,(\tilde{s})|}{\mathsf{max}\left\{\tilde{s}^{\gamma},\varepsilon\right\}}\end{aligned}$$

if $t=\frac{\tilde{s}}{1-\tilde{s}}$ is substituted. Furthermore, choosing $\tilde{s}=\frac{sB}{1+B}$ yields

$$\sup_{t\in(0,B]}\frac{|\mathcal{G}(t)|}{(1+t)\mathsf{b}(t)}\overset{d}{=}\sup_{s\in(\,0,1]}\underset{1\leq i\leq d}{\mathsf{max}}\frac{\left|W_i\left(\frac{sB}{1+B}\right)\right|}{\mathsf{max}\left\{\left(\frac{sB}{1+B}\right)^{\gamma},\varepsilon\right\}}\overset{d}{=}\left(\frac{B}{1+B}\right)^{1/2-\gamma}\sup_{s\in(\,0,1]}\underset{1\leq i\leq d}{\mathsf{max}}\frac{|W_i\,(s)|}{\mathsf{max}\left\{s^{\gamma},\varepsilon\left(\frac{1+B}{B}\right)^{\gamma}\right\}}.$$

Thus, we can use Monte Carlo simulations to obtain critical values $c=c\,(\tilde{\alpha})$ in dependence of the significance level $\tilde{\alpha}$ based on the equality

$$\begin{aligned}&\mathsf{P}_{\mathsf{H}_0}\left(\left(\frac{B}{1+B}\right)^{1/2-\gamma}\sup_{s\in(\,0,1]}\underset{1\leq i\leq d}{\mathsf{max}}\frac{|W_i\,(s)|}{\mathsf{max}\left\{s^{\gamma},\varepsilon\left(\frac{1+B}{B}\right)^{\gamma}\right\}}\geq c\,(\tilde{\alpha})\right)\\&=1-\left[\mathsf{P}_{\mathsf{H}_0}\left(\sup_{s\in(0,1]}\frac{|W_1(s)|}{\mathsf{max}\left\{s^{\gamma},\varepsilon\left(\frac{1+B}{B}\right)^{\gamma}\right\}}<\left(\frac{1+B}{B}\right)^{1/2-\gamma}c\,(\tilde{\alpha})\right)\right]^d=\tilde{\alpha}\end{aligned}$$

or alternatively

$$\mathsf{P}_{\mathsf{H}_0}\left(\sup_{s\in(0,1]}\frac{|W_1(s)|}{\max\left\{s^\gamma,\varepsilon\left(\frac{1+B}{B}\right)^\gamma\right\}}<\left(\frac{1+B}{B}\right)^{1/2-\gamma}c\left(\tilde{\alpha}\right)\right)=(1-\tilde{\alpha})^{1/d}.$$

Simulations showed that the critical values obtained by the use of the limit distribution of the detector yield infeasible high size distortions even for large-sized historical periods in finite samples. As a consequence the detector values tend to exceed the values of the scaled threshold function soon after the beginning of the monitoring period, whether a parameter change occurs or not. Nevertheless, to extenuate the resulting size distortions, the critical values can be obtained via a parametric Bootstrap type procedure. Recall that $\hat{\theta}_m$ is the estimate of the parameter vector calculated from the historical sample. We assume that the underlying DCC process features a similar behavior as the process determined by the parameters estimated from the historical period, if the latter one is sufficiently large. Hence, $b_{BT} = 199$ realizations of a DCC process whose structure is controlled by $\hat{\theta}_m$ (and the innovations follow a multivariate standard normal distribution) are simulated and denoted as $Y^{*(i)} := \left\{y_1^{*(i)},\ldots,y_{m(B+1)}^{*(i)}\right\}$, for $i \in \{1,\ldots,b_{BT}\}$.
An intuitive approach may be to calculate the detector values

$$\left|V_k^{*(i)}\right| = \left|\sum_{t=m+1}^{m+k}\left[\widehat{D}_m^{*(i)}\right]^{-1/2}\nabla\hat{l}_t^{*(i)}\left(\hat{\theta}_m\right)\right|$$

from each sample $Y^{*(i)}$ with $\nabla\hat{l}_t^{*(i)}\left(\hat{\theta}_m\right)$ the QLL contributions and

$$\widehat{D}_m^{*(i)} = \frac{1}{m}\sum_{t=1}^{m}\nabla\hat{l}_t^{*(i)}\left(\hat{\theta}_m\right)\left[\nabla\hat{l}_t^{*(i)}\left(\hat{\theta}_m\right)\right]^T$$

the estimate of the matrix $\boldsymbol{D}$ from Assumption 3.2 based on the first m observations of $Y^{*(i)}$. But since we are not interested in using the exact limit distribution of the detector, the matrix $\boldsymbol{D}$ is substituted by the identity matrix to avoid the additional uncertainty that goes along with the matrix estimation. Further simulations that are not part of this article showed that this approach yields a remarkable decrease of the size distortions compared to the use of an estimate of $\boldsymbol{D}$. Of

course, other choices can be used instead of the identity matrix such a matrix closer to $\boldsymbol{D}$, such as an estimate of the diagonal of $\boldsymbol{D}$. In any case, this would mean having to perform an estimation process that is precisely what we try to avoid.

Denote the resulting detector as $\left|\tilde{V}_k^{*(i)}\right|$ and the maximum of the scaled detector values gained from sample $Y^{*(i)}$ as

$$T^{*(i)} := \max_{1\leq k\leq [mB]} \frac{\left|V_k^{*(i)}\right|}{m^{1/2}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)}, \quad \text{for } i \in \{1,\ldots,b_{BT}\}.$$

The $(1-\tilde{\alpha})$ quantile of $\left\{T^{*(1)},\ldots,T^{*(b_{BT})}\right\}$ can be used as a critical value in finite sample applications. Although a detailed analysis of these critical values and their properties lies beyond the scope of this article, they show a satisfying behavior in simulations.

Lastly, we investigate the asymptotic distribution of the detector under a parameter change. Recall that under the alternative of a structural break at an unknown position in the monitoring period, the parameter vector changes from $\boldsymbol{\theta}$ to $\boldsymbol{\theta}^*$ at the k^*-th point of the monitoring period. Under similar assumptions as for Theorem 3.1, we can show that the detector is consistent, i.e.

Theorem 3.2. *Under the alternative of a structural break and Assumptions 3.1, 3.2, 3.3 and 3.4, we have*

$$\lim_{m\to\infty} \mathsf{P}_{\mathsf{H}_1}\{\tau_m < \infty\} = 1.$$

However, since it takes some time until the influence of the post break observations on the detector is strong enough to report the presence of a changepoint, it has to be assumed that in general the changepoint location is not consistent with the first hitting time τ_m. Once the presence of a change in the parameter vector is indicated, the position of the changepoint has to be estimated. For that, a plausible estimator, analogous to those in Wied et al. (2012) and in Wied and Galeano (2013), is defined as

$$\hat{k} := \operatorname*{arg\,max}_{1\leq k\leq \tau_m-1} \frac{k}{\sqrt{\tau_m}} \left| \frac{1}{\tau_m-1} \sum_{t=m+1}^{m+\tau_m-1} \nabla \hat{l}_t\left(\hat{\theta}_m\right) - \frac{1}{k}\sum_{t=m+1}^{m+k} \nabla \hat{l}_t\left(\hat{\theta}_m\right) \right|. \tag{3.2}$$

Although a detailed analysis of the properties of the estimator in (3.2) lies beyond the scope of this paper, estimators of this type showed satisfactory behavior in simulations and applications. That is why we use (3.2) to estimate the location of the changepoint throughout the next sections.

4 Simulations

This section is devoted to the investigation of the performance of the procedure under various simulation settings. Under the null as well as under alternative hypotheses, some parameters have to be specified. First, we choose the length of the historical period as $m \in \{500, 1000, 2000\}$. In terms of trading days, this roughly equals 2, 4 and 8 years, respectively. We assume that the length of the monitoring period is considerably smaller than the length of the historical period with $B \in \{0.1, 0.2, \ldots, 0.5\}$. The dimension of the random vectors is $p \in \{3, 5\}$ and the tuning parameter is chosen as $\gamma \in \{0, 0.2, 0.4\}$. These values support the ability of the monitoring procedure to detect early or later appearing structural breaks. In any case, we simulated 1000 time series and applied our procedure to them. Note that all of the simulations are carried out for a significance level of $\tilde{\alpha} = 0.05$ and critical values are obtained with the parametric Bootstrap type procedure. The innovations of the DCC models follow a multivariate standard normal distribution.

4.1 Simulations Under the Null Hypothesis

First, we investigate the performance of the monitoring scheme under the null hypothesis of no structural break in the parameter vector. For each vector component, the variance parameters are chosen either as $\phi_i = (0.01, 0.05, 0.9)'$ or as $\phi_i = (0.01, 0.2, 0.7)'$, for all $i \in \{1, \ldots, p\}$. Therefore, the second case indicates a stronger effect of single shocks on the volatility of future observations. The correlation structure is determined by the parameters $(\alpha, \beta) = (0.05, 0.9)$ and the constant unconditional correlation matrix $\bar{Q}_p$ which is defined as

$$\bar{Q}_3 = \begin{bmatrix} 1 & 0.5 & 0.1 \\ 0.5 & 1 & 0.5 \\ 0.1 & 0.5 & 1 \end{bmatrix} \quad \text{and} \quad \bar{Q}_5 = \begin{bmatrix} 1 & 0.5 & 0.3 & 0.2 & 0.1 \\ 0.5 & 1 & 0.5 & 0.3 & 0.2 \\ 0.3 & 0.5 & 1 & 0.5 & 0.3 \\ 0.2 & 0.3 & 0.5 & 1 & 0.5 \\ 0.1 & 0.2 & 0.3 & 0.5 & 1 \end{bmatrix}.$$

The results in Table 1, that shows the empirical sizes for all parameter combinations, suggest that the empirical size increases with B, which is plausible since larger values of this parameter imply a growing length of the monitoring period and thus more uncertainty. While larger values of m and γ reduce the size distortions, higher dimensions tend to increase the probability to commit a type I error. Importantly, the influence of variations in the length of the historical period and the dimension are as expected. Furthermore, the empirical size is distinctly higher when the variance parameters are chosen as $\phi_i = (0.01, 0.05, 0.9)'$. This result was expected since the sum $\alpha_i + \beta_i$ is closer to 1 in the first scenario than in the second.

4.2 Simulations Under Various Alternatives

In this section, we investigate the performance of the proposed procedure in view of different types of structural breaks. More precisely, we first simulate changes in the variance parameters followed by changes in the unconditional correlation matrix $\bar{Q}$.

Since the results under the null showed a strong dependency on the length of the monitoring period, the simulations under alternative scenarios will be limited to the case of monitoring periods with length $0.2m$. This choice of B yields small deviations between the empirical and the theoretical size as the results from Table 1 suggest. As the length of the monitoring period depends on m, we define the location of the changepoint k^* in terms of m as $k^* = [mB\lambda^*]$, where $[\cdot]$ is the largest integer smaller than a given real number and the fraction λ^* is chosen from $\{0.05, 0.3, 0.5\}$. This indicates changes located at the beginning or later in the monitoring period.

			$\phi_i = (0.01, 0.05, 0.9)'$			$\phi_i = (0.01, 0.2, 0.7)'$		
			$\boldsymbol{m = 500}$	$\boldsymbol{m = 1000}$	$\boldsymbol{m = 2000}$	$\boldsymbol{m = 500}$	$\boldsymbol{m = 1000}$	$\boldsymbol{m = 2000}$
$\boldsymbol{p = 3}$	$\boldsymbol{B = 0.1}$	$\boldsymbol{\gamma = 0}$	0.124	0.088	0.068	0.068	0.077	0.047
		$\boldsymbol{\gamma = 0.2}$	0.133	0.084	0.074	0.067	0.070	0.052
		$\boldsymbol{\gamma = 0.4}$	0.116	0.082	0.070	0.064	0.058	0.066
	$\boldsymbol{B = 0.2}$	$\boldsymbol{\gamma = 0}$	0.150	0.101	.0.094	0.069	0.057	0.088
		$\boldsymbol{\gamma = 0.2}$	0.151	0.086	0.091	0.069	0.077	0.087
		$\boldsymbol{\gamma = 0.4}$	0.118	0.080	0.083	0.065	0.060	0.060
	$\boldsymbol{B = 0.3}$	$\boldsymbol{\gamma = 0}$	0.177	0.111	0.089	0.120	0.087	0.081
		$\boldsymbol{\gamma = 0.2}$	0.151	0.094	0.073	0.095	0.075	0.083
		$\boldsymbol{\gamma = 0.4}$	0.143	0.084	0.086	0.071	0.054	0.069
	$\boldsymbol{B = 0.4}$	$\boldsymbol{\gamma = 0}$	0.213	0.113	0.118	0.106	0.109	0.104
		$\boldsymbol{\gamma = 0.2}$	0.193	0.120	0.105	0.110	0.098	0.106
		$\boldsymbol{\gamma = 0.4}$	0.147	0.094	0.072	0.077	0.066	0.063
	$\boldsymbol{B = 0.5}$	$\boldsymbol{\gamma = 0}$	0.197	0.118	0.141	0.129	0.110	0.122
		$\boldsymbol{\gamma = 0.2}$	0.179	0.135	0.097	0.100	0.109	0.115
		$\boldsymbol{\gamma = 0.4}$	0.166	0.111	0.090	0.086	0.073	0.073
$\boldsymbol{p = 5}$	$\boldsymbol{B = 0.1}$	$\boldsymbol{\gamma = 0}$	0.139	0.093	0.079	0.080	0.085	0.062
		$\boldsymbol{\gamma = 0.2}$	0.141	0.104	0.067	0.071	0.066	0.047
		$\boldsymbol{\gamma = 0.4}$	0.117	0.093	0.072	0.070	0.048	0.060
	$\boldsymbol{B = 0.2}$	$\boldsymbol{\gamma = 0}$	0.153	0.099	0.083	0.079	0.080	0.082
		$\boldsymbol{\gamma = 0.2}$	0.161	0.112	0.085	0.088	0.081	0.068
		$\boldsymbol{\gamma = 0.4}$	0.148	0.083	0.073	0.074	0.069	0.059
	$\boldsymbol{B = 0.3}$	$\boldsymbol{\gamma = 0}$	0.181	0.109	0.087	0.102	0.118	0.116
		$\boldsymbol{\gamma = 0.2}$	0.161	0.110	0.098	0.087	0.098	0.090
		$\boldsymbol{\gamma = 0.4}$	0.148	0.087	0.085	0.074	0.064	0.071
	$\boldsymbol{B = 0.4}$	$\boldsymbol{\gamma = 0}$	0.181	0.117	0.122	0.121	0.106	0.125
		$\boldsymbol{\gamma = 0.2}$	0.199	0.109	0.111	0.102	0.108	0.101
		$\boldsymbol{\gamma = 0.4}$	0.151	0.114	0.088	0.073	0.079	0.086
	$\boldsymbol{B = 0.5}$	$\boldsymbol{\gamma = 0}$	0.198	0.111	0.131	0.152	0.126	0.140
		$\boldsymbol{\gamma = 0.2}$	0.186	0.111	0.123	0.122	0.120	0.118
		$\boldsymbol{\gamma = 0.4}$	0.182	0.107	0.107	0.092	0.084	0.059

Table 1: Empirical size for all parameter combinations.

4.2.1 Changes in the Variance Parameters

We investigate two different settings of changes in the variance parameters. First, we assume that ϕ_i changes, first, from $\phi_i = (0.01, 0.05, 0.9)'$ to $\phi_i^* = (0.005, 0.2, 0.7)'$ and, second, from $\phi_i^* = (0.005, 0.2, 0.7)'$ to $\phi_i = (0.01, 0.05, 0.9)'$, for all $i \in \{1, \ldots, p\}$. These settings will be denoted as Scenario 1 and 2, respectively. Note that apart from the obvious change in parameters that causes a change in the conditional variance structure, Scenario 1 implies a decrease in the unconditional variances of all components while Scenario 2 involves a variance increase.

The results for Scenario 1 in Tables 2 and 3 and those for Scenario 2 in Tables 4 and 5 suggest that the power depends positively on the length of the historical period and negatively on the dimension

of the random vectors. While the first result is as expected, the negative influence of p on the power may be explained by the fact that the share of the $3p$ variance parameters in the group of all parameters decreases with growing dimension. Thus, changes in the variance parameters might be harder to detect if p gets large.

The ability to detect parameter changes is distinctly higher for changes located at the beginning of the monitoring period than for later ones which is a typical property of sequential monitoring schemes based on the information of a historical sample, see for instance Berkes et al. (2004), Wied and Galeano (2013) or Pape et al. (2016). Furthermore, parameter changes that lead to decreased unconditional variance can be detected much more reliably than changes that entail variance increases. This property underlines the results under the null, which suggest a stronger tendency of the detector to cross the threshold function if the initial variance parameters are chosen as $\phi_i = (0.01, 0.05, 0.9)'$ rather than $\phi_i = (0.01, 0.2, 0.7)'$.

In addition and consistently with the results under the null, the power decreases with the tuning parameter γ. This effect occurs even if the structural break is located at a later point in the monitoring period and should be detected more frequently if a higher value of γ is used.

The results concerning the estimated changepoint locations in Tables 2-5 suggest that the position of changepoints located at a fraction of $\lambda^* = 0.3$ of the monitoring period, can be estimated without large distortions while earlier and later changes, respectively, are systematically placed too early and late, respectively, in the dataset. Note that the results for the estimated changepoint locations depend strongly on the properties of the first hitting times since these define the length of the subsample, that is used to locate the changepoint.

4.2.2 Changes in the Correlation Parameters

Another possible alternative scenario is a change in the correlation structure. We assume that the variance parameters as well as α and β stay constant, whereas the matrix $\bar{Q}$ changes from $\mathbb{I}_p$ to $\bar{Q}^*$. The latter one is a matrix whose main diagonal entries are equal to one, while all of the diagonal entries are Δ with $\Delta \in \{0.1, \ldots, 0.9\}$. The variance parameters and α and β are chosen as in Section 4.1.

				Empirical first hitting times					Location estimator				
$\boldsymbol{\lambda^*}$	$\boldsymbol{m}$	$\boldsymbol{\gamma}$	Power	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
0.05	**500** ($k^* = 5$)	**0**	0.973	44.87	14.49	35	42	51	18.05	10.39	11	16	23
		0.2	0.971	40.28	15.77	30	38	48	16.08	9.76	9	14	21
		0.4	0.971	40.24	18.28	29	38	50	16.03	10.32	9	14	21
	1000 ($k^* = 10$)	**0**	1.000	60.40	15.46	51	57	68	24.40	11.71	17	22	29
		0.2	0.999	52.88	16.09	42	50	60	21.18	10.53	14	19	26
		0.4	1.000	47.09	19.83	36	45	56	19.20	11.89	12	17	24
0.3	**500** ($k^* = 30$)	**0**	0.916	66.47	15.56	58	66	77	31.62	9.50	27	31	36
		0.2	0.891	63.86	17.27	54	64	75	30.46	10.17	26	30	35
		0.4	0.849	64.25	22.60	54	68	79	31	12.26	26	31	38
	1000 ($k^* = 60$)	**0**	0.998	110.78	20.19	97	109	122	60.59	12.06	55	60	66
		0.2	0.998	107.62	22.10	94	106	119	59.84	13.31	55	60	66
		0.4	0.994	107.05	29.89	94	108	123	58.73	17.64	54	60	66
0.5	**500** ($k^* = 50$)	**0**	0.724	77.82	17.47	70.75	81	91	43.54	13.39	38	47	51
		0.2	0.703	76.83	19.48	70	81	90	42.58	14.30	36	47	52
		0.4	0.558	71.13	28.28	66	80	91	39.39	17.75	31	46	52
	1000 ($k^* = 100$)	**0**	0.965	148.56	22.61	135	149	163	92.07	20.31	87	97	103
		0.2	0.972	148.36	24.02	136	149	162	93.03	19.56	88	97	103
		0.4	0.931	150.02	36.68	140	156	172	92.43	24.40	90	98	104

Table 2: Power against changes in the parameters that imply a variance decrease (Scenario 1) and properties of the first hitting times τ_m and estimated changepoints $\hat{k}$ for $p = 3$.

				Empirical first hitting times					Location estimator				
$\boldsymbol{\lambda^*}$	$\boldsymbol{m}$	$\boldsymbol{\gamma}$	Power	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
0.05	**500** ($k^* = 5$)	**0**	0.927	44.46	13.82	36	42	50	17.23	9.07	11	15	22
		0.2	0.913	39.51	15.13	30	37	45	15.31	8.99	9	13	19
		0.4	0.892	38.88	16.83	29	37	48	15.46	10.01	9	13	20
	1000 ($k^* = 10$)	**0**	0.997	60.62	14.83	51	58	67	23.61	11.11	16	21	28
		0.2	0.997	51.60	14.42	42	49	59	20.66	10.31	14	18	25
		0.4	0.996	47.04	18.90	37	45	57	18.85	11.01	12	17	24
0.3	**500** ($k^* = 30$)	**0**	0.859	66.82	15.25	58	67	77	32.15	9.05	27	31	37
		0.2	0.848	64.89	16.99	56	65	76	31.01	10.60	26	30	37
		0.4	0.806	64.09	22.07	55	66	78	30.23	11.81	25.25	30	37
	1000 ($k^* = 60$)	**0**	0.993	108.95	17.86	97	107	119	59.93	12.89	54	60	65
		0.2	0.992	106.35	20.53	94	106	118	58.72	13.69	54	59	64
		0.4	0.992	107.68	28.43	97	109	123	58.45	16.38	54	60	66
0.5	**500** ($k^* = 50$)	**0**	0.680	79.04	16.92	73	82	91	43.53	12.98	37	46	51
		0.2	0.673	76.54	20.58	70	82	91	42.48	14.36	37	46	51
		0.4	0.553	70.67	30.36	65	82	93	38.90	18.00	31	46	51
	1000 ($k^* = 100$)	**0**	0.973	147.58	22.32	136	148	161	91.81	19.18	86	96	102
		0.2	0.972	145.77	25.43	135	148	161	90.26	21.93	84	97	102
		0.4	0.947	148.30	38.44	139	155	169	90.00	25.49	87	97	103

Table 3: Power against changes in the parameters that imply a variance decrease (Scenario 1) and properties of the first hitting times τ_m and estimated changepoints $\hat{k}$ for $p = 5$.

				Empirical first hitting times					Location estimator				
$\boldsymbol{\lambda^*}$	$\boldsymbol{m}$	$\boldsymbol{\gamma}$	Power	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
0.05	**500** ($k^* = 5$)	**0**	0.768	64.97	19.92	51	66	80	24.62	16.68	12	20	32
		0.2	0.702	56.90	23.55	39.25	57	75	21.57	15.39	10	17	29
		0.4	0.558	49.34	29.99	21.25	48	75.75	20.44	17.00	7	15	29
	1000 ($k^* = 10$)	**0**	0.974	116.05	32.61	95	114	138	39.92	30.82	16	29	55
		0.2	0.955	107.65	38.44	80	107	133	36.77	30.67	14	26	49
		0.4	0.810	107.81	49.94	72	109.5	146	39.06	35.14	14	25	53
0.3	**500** ($k^* = 30$)	**0**	0.527	76.44	15.92	65	79	90	32.30	11.46	27	32	37
		0.2	0.482	73.28	18.45	61	76	88	31.38	12.84	25	31	36
		0.4	0.271	62.76	27.40	46	68	85	29.00	14.92	20	30	36
	1000 ($k^* = 60$)	**0**	0.809	151.87	28.70	130	155	176	58.77	15.35	52	60	65
		0.2	0.727	145.37	33.39	121	149	172	55.86	16.11	48	59	65
		0.4	0.457	142.46	47.80	122	153	179	54.78	22.12	46	59	65
0.5	**500** ($k^* = 50$)	**0**	0.346	81.76	14.91	72	84	95	44.04	15.42	36	48.5	54
		0.2	0.292	76.91	19.49	70	81	92	40.72	16.24	29	46	52
		0.4	0.172	58.26	35.67	18	72	87.25	32.97	23.10	6.75	39.5	52
	1000 ($k^* = 100$)	**0**	0.505	167.32	24.59	153	172	187	86.40	22.86	74	94	102
		0.2	0.445	161.34	32.89	146	168	184	83.36	25.98	71	91	101
		0.4	0.221	138.81	61.53	122	164	182	74.26	37.55	52	90	101

Table 4: Power against changes in the parameters that imply a variance increase (Scenario 2) and properties of the first hitting times τ_m and estimated changepoints $\hat{k}$ for $p = 3$.

$\boldsymbol{\lambda^*}$	$\boldsymbol{m}$	$\boldsymbol{\gamma}$	Power	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
0.05	**500** ($k^* = 5$)	**0**	0.738	67.80	19.61	54	70.50	83	25.32	17.10	12.25	20	34
		0.2	0.644	60.38	23.72	43	62.5	80	23.15	17.21	11	18	31
		0.4	0.409	46.89	30.46	15	49	72	18.67	16.67	6	13	26
	1000 ($k^* = 10$)	**0**	0.964	127.63	32.68	105	127	151	43.43	31.97	18	33	62
		0.2	0.935	117.04	39.51	90	116	145	38.39	31.36	15	27	54
		0.4	0.704	112.02	55.44	73	120	157.25	41.77	37.69	14	27	59
0.3	**500** ($k^* = 30$)	**0**	0.491	76.75	16.17	66	80	89	30.91	11.62	25.5	31	36
		0.2	0.394	72.05	20.39	59.25	73	89	30.59	12.72	24	31.5	36
		0.4	0.193	56.33	34.87	30	64	88	25.09	17.89	8	29	34
	1000 ($k^* = 60$)	**0**	0.742	155.38	29.41	137	159	178	58.46	15.89	53	60	65
		0.2	0.671	149.38	33.91	128	154	176	57.44	15.32	51	60	65
		0.4	0.363	135.31	58.76	114.5	151	181	53.02	26.20	43.5	59	65
0.5	**500** ($k^* = 50$)	**0**	0.292	79.91	14.59	70	82	92	41.55	15.36	31	46.5	53
		0.2	0.265	76.12	19.48	65	81	91	39.81	15.82	29	46	52
		0.4	0.127	46.23	36.34	4.5	56	78	25.44	21.42	2	24	48
	1000 ($k^* = 100$)	**0**	0.461	166.82	23.85	150	171	186	83.79	21.66	71	91	100
		0.2	0.384	160.36	33.85	144.75	169	185	82.73	24.10	68.75	90.5	101
		0.4	0.175	125.46	71.73	64.5	156	182	65.53	39.96	31.5	80	99

Table 5: Power against changes in the parameters that imply a variance increase (Scenario 2) and properties of the first hitting times τ_m and estimated changepoints $\hat{k}$ for $p = 5$.

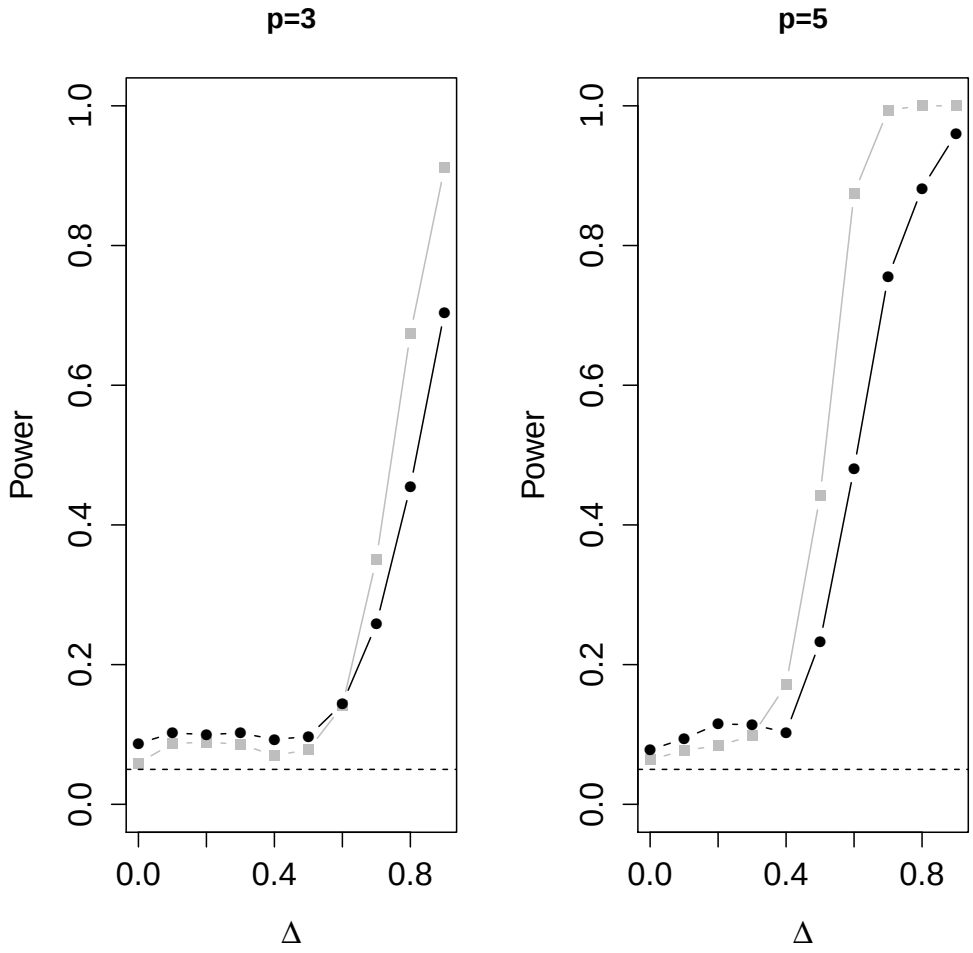

Figure 1: Power against correlation changes. Black: $\omega_i = 0.01,\ \alpha_i = 0.05,\ \beta_i = 0.9$, for $i = 1, \ldots, p$, Gray: $\omega_i = 0.01,\ \alpha_i = 0.2,\ \beta_i = 0.7$, for $i = 1, \ldots, p$.

The power results for changes at fraction $\lambda^* = 0.05$ of the monitoring period are illustrated in Figure 1 for simulated time series of dimension 3 or 5, a historical period consisting of 1000 data points and tuning parameter $\gamma = 0.2$. Please note that we focus on $\gamma = 0.2$ here because the simulation results for changes in the variance parameters suggest that this value yields a good balance between high power (obtained for larger γ) and early breakpoint detection (smaller γ). The results reveal problems to detect correlation changes of moderate magnitude for both choices of the vector of variance parameters. However, the power curve has a large slope for higher values and converges to one. While smaller changes in the correlation parameters can be detected more frequently if

the variance parameters are chosen as $\phi_i = (0.01, 0.05, 0.9)'$ rather than $\phi_i = (0.01, 0.2, 0.7)'$, the opposite is true for larger values of Δ.

The fact that some of the power results are quite low, suggests that the QLL function seems to be very flat in some regions, so parameter changes of small magnitude are hard to detect.

5 Empirical Results

To investigate the performance of the proposed monitoring scheme under real conditions, the procedure is applied to a group of asset returns. Due to the fact that a conjoint modeling seems to be appropriate for the returns of assets from the same industrial sector and monetary area, we choose the assets of several insurance companies, which are listed at different stock exchanges throughout Europe. More precisely, we monitor the log returns of the assets of *Allianz* (abbreviated by *All*), *AXA*, *Generali* (*Gen*), *ING* and *Munich Re* (*MRe*) in the time from 2003-01-03 to 2016-10-25. The data set is available upon request. Engle (2002) argued that the DCC model is in principle well-suited to model the typical features of multivariate return time series. Furthermore, Bollerslev (1986) pointed out, that even GARCH models of order $(1, 1)$ are capable of explaining the behavior of log returns very well. Thus, we use GARCH$(1, 1)$ models for the univariate conditional variance equations (2.4), which is in line with our approach in Section 2. The starting date (which lies after the peak of the dotcom bubble) is chosen such that the assumption of constant parameters in the initial period is reasonable.

As the results in Table 1 suggest that the size increases considerably with the length of the monitoring period and hence with B, we monitor the data by the use of a stepwise approach, which works as follows:

1. Use the first m observations as historical period and monitor the following Bm observations for a parameter change.

2. There are two possibilities:

   (a) If a changepoint is detected in the subsample, estimate the location of the changepoint

and cut-off all of the pre-change observations. Then, restock the subsample to m observations and return to step 1. If there are not enough observations left to restock the historical sample to m observations or the monitoring dataset to mB observations, terminate the procedure.

(b) If no changepoint is detected in the subsample, cut-off the first mB data points of the historical sample and add the data of the previous monitoring period to the historical dataset. Then, return to step 1. If there are not enough observations left to restock the historical sample to m observations or the monitoring dataset to mB observations, terminate the procedure.

The procedure is similar to that used in Pape et al. (2016). Clearly, during this procedure, it is not ruled out that there might be false changepoint detections (type I errors) which influence subsequent results of the tests. The development of similar procedure which keep a predetermined size seems a challenging problem. One possibility might be to use a critical level that decreases with each new changepoint detected, so that the size of the procedure remains fixed. However, this process would require estimating a new critical value by simulation with each new change point detected, which could make the computational cost of the procedure prohibitive. Therefore, we prefer to leave this problem for future research.

The results for $m \in \{500, 1000\}$, $B = 0.2$ and $\gamma \in \{0, 0.2, 0, 4\}$ can be seen in Table 6. While in general, $\gamma = 0.2$ yields a good balance between high power and early breakpoint detection, the break dates do not depend on γ in this example. On the other hand, there is one breakpoint more for $m = 500$ and we focus on this case in the following, as it seems to be more informative about the true data generating process. Interestingly, the first stopping time is the same for both $m = 500$ and $m = 1000$. The estimated changepoint locations are shown in Figure 2 with two of the monitored time series. The time series of asset returns are split up into calm and more turbulent phases. The estimated break point locations in Table 6 can be linked to important economic events of the last two decades. The calmer period from 2002 on was interrupted in 2008 by the financial crisis followed by the debt crisis. Some years later, we observe the beginning of a recovery phase

of the stock markets.

Estimates of the model parameters calculated from the data between two successive changepoints can be seen in Table 7. To measure the magnitude of the changes in the estimated parameters, the table also contains the squared Euclidean norm of the persistence parameters estimated from the subsamples. One observes considerable differences, whereas the parameters are in general larger in the period after the climax of the financial crisis in 2008 compared to the periods directly before and in particular after. This can be linked to the often observed phenomenon that correlations and variances of asset returns tend to increase in times of crisis, see Sandoval Jr. and De Paula Franca (2012) among others.

6 Conclusion

We present a method to detect changes in the parameter vector of the DCC model proposed by Engle (2002) which is based on quasi-log-likelihood scores and allows to detect changes in the conditional and unconditional variance and covariance structure. We analyze the size and power properties of the presented procedure and apply it to a group of log returns that belong to the assets of several insurance companies. In applications it turns out as a difficult problem that the assumption of a historical period which is free from structural breaks cannot be checked with a known retrospective method. The search for a solution for this problem is left as a task for future research. Also, the statistic used is designed for situations in which many model parameters change. Thus, this statistic may not useful for detecting changes in one or few model parameters. Even if such changes are detected it would be complicated to distinguish which model parameters have changed already. The development of statistics adapted to these situations is also left for future research. Finally, note that we have focus on detecting the presence of changepoints. A complete development and understanding of the implications of these changepoints in issues such as forecasting or the estimation of risk measures deserve their own space.

$\gamma \in \{0, 0.2, 0.4\}, B = 0.2,\ m = 500$	
$\boldsymbol{\tau_m}$	$\hat{\boldsymbol{k}}$
2008/01/17	2007/10/22
2010/02/03	2009/10/23
2015/07/13	2015/06/19
$\gamma \in \{0, 0.2, 0.4\}, B = 0.2,\ m = 1000$	
$\boldsymbol{\tau_m}$	$\hat{\boldsymbol{k}}$
2008/01/17	2008/01/15
2014/03/19	2013/09/09

Table 6: First hitting times and estimated changepoint locations.

	2003/01/02 -2007/10/22	2007/10/23 -2009/10/23	2009/10/26 -2015/06/19	2015/06/22 2016/10/25
$\hat{\omega}_{All}$	< 0.00001	0.00003	0.00001	0.00006
$\hat{\omega}_{AXA}$	0.00001	0.00002	0.00001	0.00028
$\hat{\omega}_{Gen}$	< 0.00001	< 0.00001	< 0.00003	0.00002
$\hat{\omega}_{ING}$	0.00001	0.00004	0.00001	0.00007
$\hat{\omega}_{MRe}$	< 0.00001	0.00002	0.00001	0.00004
$\hat{\boldsymbol{\alpha}}_{All}$	0.078	0.142	0.077	0.111
$\hat{\boldsymbol{\alpha}}_{AXA}$	0.073	0.153	0.089	0.551
$\hat{\boldsymbol{\alpha}}_{Gen}$	0.065	0.099	0.052	0.057
$\hat{\boldsymbol{\alpha}}_{ING}$	0.101	0.182	0.069	0.223
$\hat{\boldsymbol{\alpha}}_{MRe}$	0.069	0.180	0.094	0.114
$\hat{\boldsymbol{\beta}}_{All}$	0.904	0.836	0.901	0.683
$\hat{\boldsymbol{\beta}}_{AXA}$	0.901	0.846	0.895	0.042
$\hat{\boldsymbol{\beta}}_{Gen}$	0.908	0.898	0.936	0.917
$\hat{\boldsymbol{\beta}}_{ING}$	0.877	0.817	0.918	0.676
$\hat{\boldsymbol{\beta}}_{MRe}$	0.915	0.801	0.868	0.700
$\hat{\boldsymbol{\alpha}}$	0.010	0.027	0.034	0.035
$\hat{\boldsymbol{\beta}}$	0.965	0.896	0.832	0.805
$\left\|\hat{\boldsymbol{\theta}}\right\|_2$	5.021	4.454	4.810	3.288

Table 7: Persistence parameters estimated from the data between successive detected changepoints for $m = 500$ from Table 6 and the euclidical norm of the estimated parameter vectors.

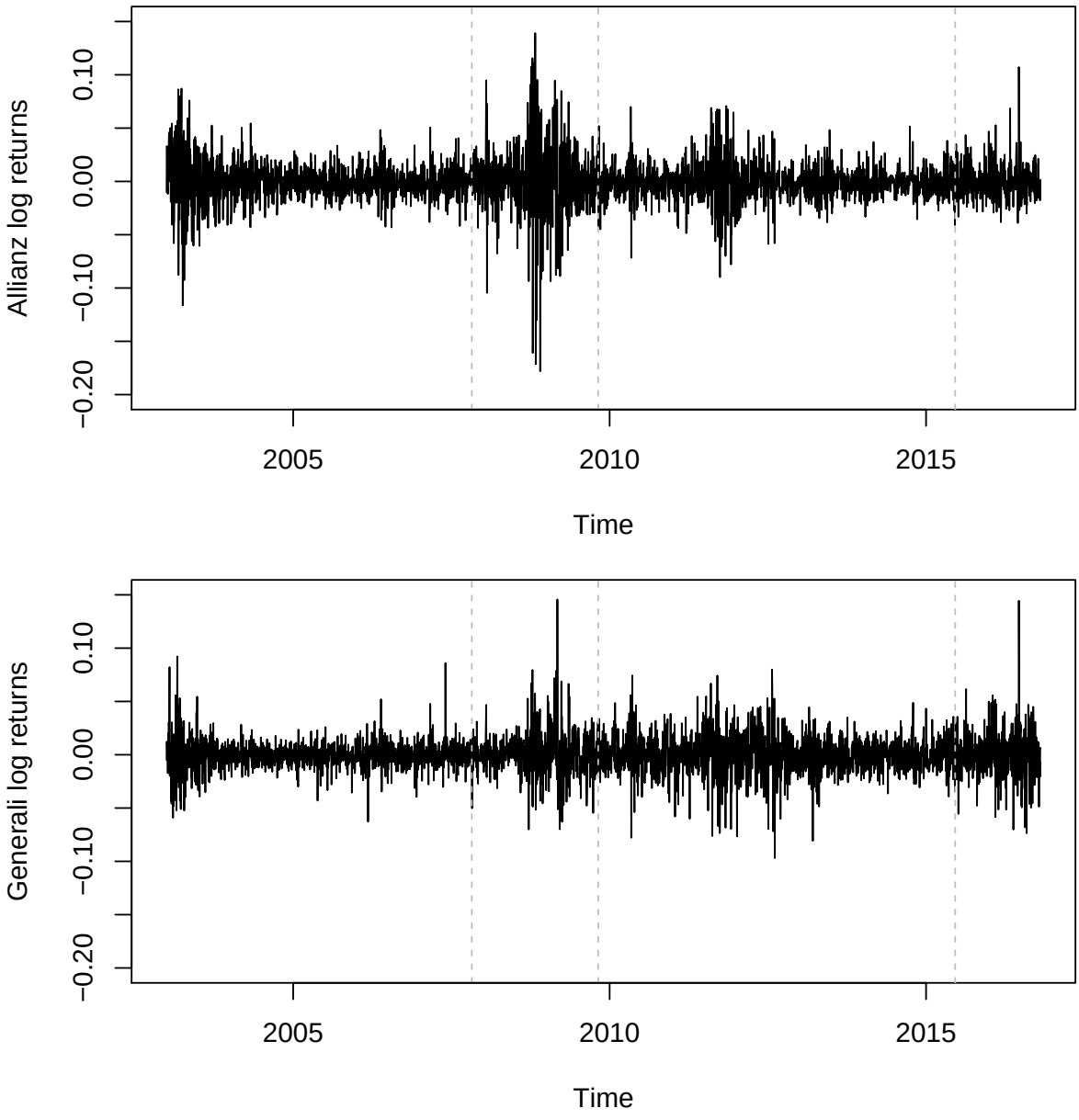

Figure 2: Log returns of the *Allianz* and *Generali* assets with the detected changepoints for $m = 1000$ from Table 6 (dashed gray lines).

A Proofs

Proof of Theorem 3.1

The proof follows the arguments of the proof of Theorem 3.1 in Berkes et al. (2004). It makes use of the fact that the process $\left\{\nabla \hat{l}_t\left(\hat{\theta}_m\right),\ t \in \mathbb{Z}\right\}$ behaves similarly as $\{\nabla l_t(\boldsymbol{\theta}),\ t \in \mathbb{Z}\}$. This is a stationary and ergodic martingale difference sequence, which yields the convergence to a Gaussian limit process.

First, we have for $m \to \infty$ the following statement that equals Lemma 6.3 in Berkes et al. (2004):

$$\sup_{1 \le k < \infty} \frac{\sum_{i=m+1}^{m+k} \nabla l_i\left(\hat{\theta}_m\right) - \left[\sum_{i=m+1}^{m+k} \nabla l_i(\boldsymbol{\theta}) + \left(\hat{\theta}_m - \boldsymbol{\theta}\right) k \mathsf{E}\left(\nabla^2 l_0(\boldsymbol{\theta})\right)\right]}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} = o_P(1). \tag{A.1}$$

The validity of (A.1) is implied by a Taylor approximation and the uniform convergence $\nabla^2 l_t(\cdot)$ to $\mathsf{E}\left[\nabla^2 l_0(\cdot)\right]$. The latter one is implied by Theorem A.2.2 in White (1994). Since U is a compact set and $\nabla^2 l_t(u)$ is continuous in u for all y_t as well as measurable in y_t for all $u \in U$, the dominance condition remains to be verified. We choose the dominating function as $\sup_{u \in U}\left|\nabla^2 l_t(u)\right|$ whose expectation if finite by Assumption 3.4.3.

Therefore, Theorem A.2.2 in White (1994) implies

$$\sup_{u \in U}\left|\frac{1}{m}\sum_{i=1}^{m}\nabla^2 l_i(u) - \mathsf{E}\left[\nabla^2 l_0(u)\right]\right| = \frac{1}{m}\sup_{u \in U}\left|\sum_{i=1}^{m}\Big(\nabla^2 l_i(u) - \mathsf{E}\left[\nabla^2 l_0(u)\right]\Big)\right| \overset{a.s.}{\to} 0.$$

Analogously to Lemma 6.4 in Berkes et al. (2004) and by the use of Assumptions 3.4.(1,2,4,5), we have for $m \to \infty$:

$$\left(\hat{\theta}_m - \boldsymbol{\theta}\right) = \frac{1}{m}\sum_{i=1}^{m}\nabla l_i(\boldsymbol{\theta})\left[\mathsf{E}\Big(\nabla^2 l_0(\boldsymbol{\theta})\Big)\right]^{-1}\left[1 + o(1)\right].$$

Furthermore, Assumptions 3.4.(1,2,4,5) and 3.3 and Lemma 6.5 in Berkes et al. (2004) yield:

$$\sup_{1 \le k < \infty} \frac{\left|\sum_{i=m+1}^{m+k} \nabla l_i\left(\hat{\theta}_m\right) - \left[\sum_{i=m+1}^{m+k} \nabla l_i(\boldsymbol{\theta}) - \frac{k}{m}\sum_{i=1}^{m}\nabla l_i(\boldsymbol{\theta})\right]\right|}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} = o_P(1).$$

Then, by the use of Lemma 6.6 in Berkes et al. (2004) we have:

$$\sup_{1\leq k<\infty} \frac{\left|\sum_{i=m+1}^{m+k} \nabla l_i\left(\boldsymbol{\theta}\right) - \frac{k}{m}\sum_{i=1}^{m} \nabla l_i\left(\boldsymbol{\theta}\right)\right|}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} \xrightarrow{d} \sup_{0<t<\infty} \frac{|\boldsymbol{W_D}(1+t) - (1+t)\boldsymbol{W_D}(1)|}{(1+t)\mathsf{b}(t)} \tag{A.2}$$

where $\{\boldsymbol{W_D}(t), t \in [0,B]\}$ is a d-variate Gaussian process with

$$\mathsf{E}\left[\boldsymbol{W_D}(s)\right] = \mathbf{0}_d, \quad \forall\ s \in [0,B), \quad \text{and} \quad \mathsf{E}\left[\boldsymbol{W'_D}(k)\boldsymbol{W_D}(l)\right] = \mathsf{min}\left\{k,l\right\}\boldsymbol{D}, \quad \forall\ k,l \in [0,B].$$

As a consequence of (A.2) and Assumption 3.4.(6,7) we have:

$$\sup_{1\leq k<mB} \frac{\left|\sum_{i=m+1}^{m+k} \widehat{D}_m^{-\frac{1}{2}} \nabla \hat{l}_i\left(\hat{\theta}_m\right)\right|}{m^{\frac{1}{2}}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} \xrightarrow{d} \sup_{0<t\leq B} \frac{\left|\boldsymbol{D}^{-\frac{1}{2}}\left[\boldsymbol{W_D}(1+t) - (1+t)\boldsymbol{W_D}(1)\right]\right|}{(1+t)\mathsf{b}\left(t\right)}.$$

A simple recalculation of the properties of the resulting process indicates that

$$\left\{\boldsymbol{D}^{-\frac{1}{2}}\left[\boldsymbol{W_D}(1+t) - (1+t)\boldsymbol{W_D}(1)\right], t \in [0,\infty)\right\}$$

and $\{\mathcal{G}(t), t \in [0,\infty)\}$ possess the same distribution. ■

Proof of Theorem 3.2

Under the alternative of a change in the vector of parameters, it is appropriate to decompose the detector as in Berkes et al. (2004):

$$\frac{\sum_{i=m+1}^{m+k} \nabla l_i\left(\hat{\theta}_m\right)}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} = \frac{\sum_{i=m+1}^{m+k^*-1} \nabla l_i\left(\hat{\theta}_m\right)}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} + \frac{\sum_{i=m+k^*}^{m+k} \nabla l_i\left(\hat{\theta}_m\right)}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)}. \tag{A.3}$$

The first summand on the righthand side of (A.3) is based on the gradient contributions of observations before the parameter change. It can be treated analogously to the proof of Theorem 3.1 or to the proof of Theorem 3.1. in Berkes et al. (2004). Note that $m \to \infty$ implies $k^* \to \infty$ and we

have:

$$\sup_{k^* \leq k < \infty} \frac{\left| \sum_{i=m+1}^{m+k^*-1} \nabla l_i\left(\hat{\theta}_m\right) \right|}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} \xrightarrow{d} \sup_{t \in [\lambda^* B, \infty)} \frac{|\boldsymbol{W_D}\left(1+\lambda^* B\right) - \left(1+\lambda^* B\right)\boldsymbol{W_D}(1)|}{(1+t)\mathsf{b}\left(t\right)}. \tag{A.4}$$

The second summand on the righthand side of (A.3) is based on the gradient contributions of the observations after the change which are driven by the parameter vector $\boldsymbol{\theta}^*$. Thus for $m \to \infty$, a taylor series expansion of $l_i'\left(\hat{\theta}_m\right)$ centered in $\boldsymbol{\theta}^*$ yields:

$$\sup_{k^* \leq k < \infty} \frac{\left| \sum_{i=m+k^*}^{m+k} \nabla l_i\left(\hat{\theta}_m\right) - \left[\sum_{i=m+k^*}^{m+k} \nabla l_i\left(\boldsymbol{\theta}^*\right) + \left(\hat{\theta}_m - \boldsymbol{\theta}^*\right)' \sum_{i=m+k^*}^{m+k} \nabla^2 l_i\left(\boldsymbol{\theta}^*\right) \right] \right|}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} = o_P(1).$$

Analogously to the proof of Theorem 4.4 in Berkes et al. (2004) Assumptions 3.4.1,2,4,5 imply:

$$\left(\hat{\theta}_m - \boldsymbol{\theta}^*\right) = \left(\hat{\theta}_m - \boldsymbol{\theta}\right) + \left(\boldsymbol{\theta} - \boldsymbol{\theta}^*\right) = -\frac{1}{m}\sum_{i=1}^{m} \nabla l_i(\boldsymbol{\theta}) \left[\mathsf{E}\left(\nabla^2 l_0(\boldsymbol{\theta})\right)\right]^{-1} \left(1 + o_P(1)\right) + \left(\boldsymbol{\theta} - \boldsymbol{\theta}^*\right).$$

Furthermore, as $\boldsymbol{\theta}^*$ is contained in the compact parameter space, there exists a neighborhood U_2 of $\boldsymbol{\theta}^*$ where the function $\frac{1}{m}\sum_{i=1}^{m} \nabla^2 l_i(u)$ converges uniformly to its theoretical counterpart with Theorem A.2.2 in White (1994). Additionally, the uniform convergence implies the convergence in probability to zero of:

$$\sup_{k^* \leq k < \infty} \frac{\left| \left(\hat{\theta}_m - \boldsymbol{\theta}^*\right)' \sum_{i=m+k^*}^{m+k} \nabla^2 l_i\left(\boldsymbol{\theta}^*\right) \right.}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)}$$
$$\left. - \frac{\left[-\frac{k-k^*+1}{m} \left[\mathsf{E}\left(\nabla^2 l_0(\boldsymbol{\theta})\right)\right]^{-1} \mathsf{E}\left(\nabla^2 l_0\left(\boldsymbol{\theta}^*\right)\right) \sum_{i=1}^{m} \nabla l_i(\boldsymbol{\theta}) + (k-k^*+1)\left(\boldsymbol{\theta} - \boldsymbol{\theta}^*\right)' \mathsf{E}\left(\nabla^2 l_0\left(\boldsymbol{\theta}^*\right)\right) \right]}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} \right|.$$

Moreover, with the triangle inequality yields, we have:

$$\sup_{k^* \leq k < \infty} \frac{\left| \left[\sum_{i=m+k^*}^{m+k} \nabla l_i\left(\boldsymbol{\theta}^*\right) - \frac{k-k^*+1}{m} \left[\mathsf{E}\left(\nabla^2 l_0(\boldsymbol{\theta})\right)\right]^{-1} \mathsf{E}\left(\nabla^2 l_0\left(\boldsymbol{\theta}^*\right)\right) \sum_{i=1}^{m} \nabla l_i(\boldsymbol{\theta}) \right] \right|}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)}$$

$$- \frac{(k-k^*+1)\,(\boldsymbol{\theta}-\boldsymbol{\theta}^*)'\,\mathsf{E}\Big(\nabla^2 l_0\,(\boldsymbol{\theta}^*)\Big)\Big|}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)}$$

$$\geq \left| \sup_{k^*\leq k<\infty} \frac{\left|\sum_{i=m+k^*}^{m+k} \nabla l_i\,(\boldsymbol{\theta}^*) - \frac{k-k^*+1}{m}\left[\mathsf{E}\Big(\nabla^2 l_0(\boldsymbol{\theta})\Big)\right]^{-1}\mathsf{E}\Big(\nabla^2 l_0\,(\boldsymbol{\theta}^*)\Big)\sum_{i=1}^{m}\nabla l_i(\boldsymbol{\theta})\right|}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} \right. \tag{A.5}$$

$$\left. - \sup_{k^*\leq k<\infty} \frac{\Big|(k-k^*+1)\,(\boldsymbol{\theta}-\boldsymbol{\theta}^*)'\,\mathsf{E}\Big(\nabla^2 l_0\,(\boldsymbol{\theta}^*)\Big)\Big|}{\sqrt{m}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} \right|. \tag{A.6}$$

Thereby, for $m \to \infty$ and because of $k^* = \lambda^* m B$ also for $k^* \to \infty$, the minuend in (A.5) converges in distribution to:

$$\sup_{t\in(\lambda^*,\infty)} \frac{\Big|\boldsymbol{W}_{\boldsymbol{D}^*}(1+t) - \boldsymbol{W}_{\boldsymbol{D}^*}(1+\lambda^*) - (t-\lambda^*)\left[\mathsf{E}\Big(\nabla^2 l_0(\boldsymbol{\theta})\Big)\right]^{-1}\mathsf{E}\Big(\nabla^2 l_0\,(\boldsymbol{\theta}^*)\Big)\boldsymbol{W}_{\boldsymbol{D}}(1)\Big|}{(1+t)\,\mathsf{b}\,(t)}, \tag{A.7}$$

where $\boldsymbol{D}^* := \mathsf{Cov}\left[\nabla l_0\,(\boldsymbol{\theta}^*)\right]$ and $\{\boldsymbol{W}_{\boldsymbol{D}^*}(t),\ t\in[0,\infty)\}$ is a d-variate Gaussian process with:

$$\mathsf{E}\left[\boldsymbol{W}_{\boldsymbol{D}^*}(s)\right] = \mathbf{0}_d, \quad \forall\ s\in[0,B), \qquad \text{and} \qquad \mathsf{E}\left[\boldsymbol{W}'_{\boldsymbol{D}^*}(k)\boldsymbol{W}_{\boldsymbol{D}^*}(l)\right] = \mathsf{min}\left\{k,l\right\}\boldsymbol{D}^*, \quad \forall\ k,l\in[0,B).$$

Furthermore, for $m\to\infty$ we have for the subtrahend in (A.6):

$$\sqrt{m}\sup_{k^*\leq k<\infty}\frac{\left|\frac{k-k^*+1}{m}\,(\boldsymbol{\theta}-\boldsymbol{\theta}^*)'\,\mathsf{E}\Big(\nabla^2 l_0\,(\boldsymbol{\theta}^*)\Big)\right|}{\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} \overset{a.s.}{\to} \infty$$

Since the limits in (A.4) and (A.7) are stochastically bounded and the variable parts of the function $\mathsf{b}(\cdot)$ are chosen such that the procedure keeps its size under the null hypothesis, the detector values diverge for $m\to\infty$. This implies that any change in the vector of parameters can be detected if the historical period is long enough. ■

References

Aielli, G. (2013): "Dynamic Conditional Correlation: On Properties and Estimation," *Journal of Business and Economic Statistics*, 31(3), 282–299.

Audrino, F. and F. Trojani (2011): "A General Multivariate Threshold GARCH Model for Dynamic Correlations," *Journal of Business and Economic Statistics*, 29(1), 183–149.

Aue, A., S. Hörmann, L. Horváth, and M. Reimherr (2009): "Break Detection in the Covariance Structure of Multivariate Time Series Models," *The Annals of Statistics*, 37(6B), 4046–4087.

Berkes, I., E. Gombay, L. Horváth, and P. Kokoszka (2004): "Sequential Change-point Detection in GARCH(p,q) Models," *Econometric Theory*, 20(6), 1140–1167.

Berkes, I., L. Horváth, and P. Kokoszka (2003): "GARCH Processes: Structure and Estimation," *Bernoulli*, 9(2), 201–227.

Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31(3), 307–327.

——— (1990): "Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model," *The Review of Economics and Statistics*, 72(3), 498–505.

Bollerslev, T., R.F. Engle, and J.M. Wooldridge (1988): "A Capital Asset Pricing Model with Time-varying Covariances," *Journal of Political Economy*, 96(1), 116–131.

Capiello, L., R. Engle, and K. Sheppard (2006): "Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns," *Journal of Financial Econometrics*, 4(4), 537–572.

Charles, A. and O. Darné (2014): "Large Shocks in the Volatility of the Dow Jones Industrial Average Index: 1928-2013," *Journal of Banking and Finance*, 43, 188–199.

Chu, C.-S. J., M. Stinchcombe, and H. White (1996): "Monitoring Structural Change," *Econometrica*, 64(5), 1045–1065.

Engle, R. (2002): "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroscedasticity Models," *Journal of Business & Economic Statistics*, 20(3), 339–350.

Engle, R. and K. Sheppard (2001): "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH," *NBER Working Paper No. 8554*.

Engle, R.F. and K.F. Kroner (1995): "Multivariate Simultaneous Generalized ARCH," *Econometric Theory*, 11(1), 122–150.

Fermanian, J.-D. and H. Malongo (2017): "On the Stationarity of Dynamic Conditional Correlation Models," *Econometric Theory*, 33(3), 636–663.

Francq, C. and J.-M. Zakoïan (2016): "Estimating multivariate volatility models equation by equation," *Journal of the Royal Statistical Society Series B*, 78(3), 613–635.

Galeano, P. and D. Wied (2014): "Multiple Change Point Detection in the Correlation Structure of Financial Assets," *Computational Statistics and Data Analysis*, 76, 262–282.

Hillebrand, E. (2005): "Neglecting Parameter Changes in GARCH Models," *Journal of Econometrics*, 129, 121–138.

Horváth, L., M. Hušková, P. Kokoszka, and J. Steinebach (2004): "Monitoring Changes in Linear Models," *Journal of Statistical Planning and Inference*, 126(1), 225–251.

Nelson, D.B. and C.Q. Cao (1992): "Inequality Constraints in the Univariate GARCH model," *Journal of Business & Economic Statistics*, 10(2), 229–235.

Pape, K., P. Galeano, and D. Wied (2017): "Sequential Detection of Parameter Changes in Dynamic Conditional Correlation Models," *SFB 823 Discussion Paper*, 7.

Pape, K., D. Wied, and P. Galeano (2016): "Monitoring Multivariate Variance Changes," *Journal of Empirical Finance*, 39(A), 54–68.

Sandoval Jr., L. and I. De Paula Franca (2012): "Correlation of Financial Markets in Times of Crisis," *Physica A: Statistical Mechanics and its Applications*, 391(1-2), 187–208.

Seber, G.A.F. (2008): *A Matrix Handbook for Statisticians*, John Wiley & Sons, Hoboken, NJ, USA.

Tse, Y.K. and A.K.C. Tsui (2002): "A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations," *Journal of Business and Economic Statistics*, 20(3), 351–362.

Tweedie, R.L. (1988): "Invariant Measures for Markov Chains with no Irreducibility Assumptions," *Journal of Applied Probability*, 25, A Celebration of Applied Probability, 275–285.

White, H. (1994): *Estimation, Inference and Specification Analysis*, Cambridge University Press, New York.

Wied, D., M. Arnold, N. Bissantz, and D. Ziggel (2012): "A New Fluctuation Test for Constant Variances with Applications to Finance," *Metrika*, 75(8), 1111–1127.

Wied, D. and P. Galeano (2013): "Monitoring Correlation Change in a Sequence of Random Variables," *Journal of Statistical Planning and Inference*, 143(1), 186–196.