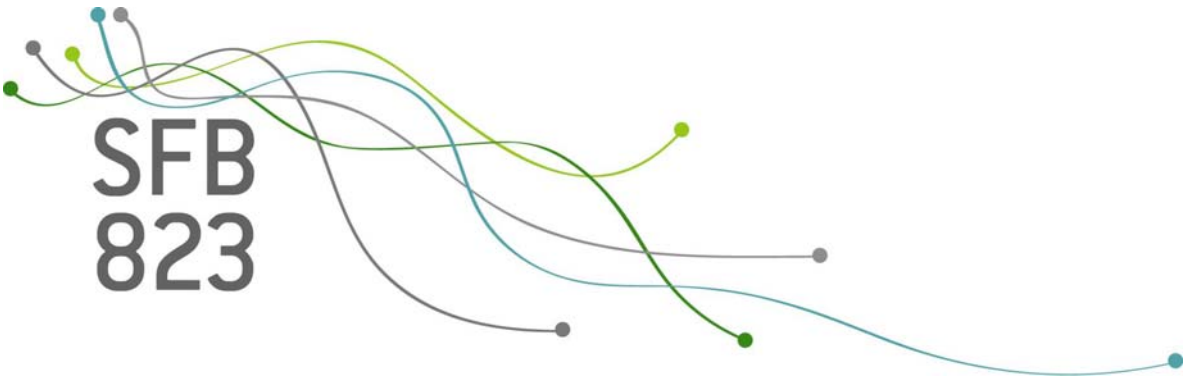
SFB 823

Monitoring multivariate variance changes

Katharina Pape, Dominik Wied,
Pedro Galeano

Nr. 29/2015

Discussion Paper



Monitoring multivariate variance changes

Katharina Pape, Dominik Wied, Pedro Galeano*

TU Dortmund and Universidad Carlos III de Madrid

This Version: August 4, 2015

Abstract

We propose a model-independent multivariate sequential procedure to monitor changes in the vector of componentwise unconditional variances in a sequence of p-variate random vectors. The asymptotic behavior of the detector is derived and consistency of the procedure stated. A detailed simulation study illustrates the performance of the procedure confronted with different types of data generating processes. We conclude with an application to the log returns of a group of *DAX* listed assets.

Keywords: Multivariate sequences; Online detection; Threshold function; Variance changes.

JEL Classification: C12, C14

*K. Pape and D. Wied: TU Dortmund, Fakultät Statistik, 44221 Dortmund, Germany. E-Mail: pape@statistik.tu-dortmund.de, Phone: +49 231 755 3869 (K. Pape). E-Mail: wied@statistik.tu-dortmund.de, Phone: +49 231 755 5419 (D. Wied). P. Galeano: Universidad Carlos III de Madrid, Departamento de Estadística, E-28903 Getafe, Madrid, Spain. E-mail: pedro.galeano@uc3m.es, Phone: +34 916 818 901. Financial support by Deutsche Forschungsgemeinschaft (SFB 823, project A1) and Ministerio de Ciencia e Innovación grant ECO2012-38442 is gratefully acknowledged.

1. Introduction

In financial contexts, variances and volatilities are of major interest since they can be used to evaluate the risk of financial instruments. Motivated by the fact that structural stability is important for forecasting and inference, this paper is concerned with possible structural changes in variances. From an empirical point of view, it is clear that, in general, variances of stock returns cannot be taken as constant over (a longer period of) time and tend to increase heavily in times of instability, see, e.g., Schwert (2011) or Charles and Darné (2014). There are many papers which deal with models for time-varying *conditional* variances; a prominent one is Bollerslev (1986) which proposes the well-known $GARCH(p, q)$ model. On the other hand, it is far from clear whether also the *unconditional* variances should be modeled in a time-varying way. Also, once such a possible change in the unconditional variance is detected, the question of dating the breakpoint arises.

This paper proposes methods for answering these two questions. In particular, our aim is to monitor the vector of variances of a series of random vectors of moderate dimension p. In practice, it is often important to get informed about changes in the model structure as soon as possible after their appearance to be able to react to the change. Hence, a monitoring procedure could be of more practical relevance than a retrospective test. The focus lies on changes in the individual variances since Bissantz et al. (2011) show that the impact of fluctuations is distinctively larger for volatilities than for correlations. Thus, we refrain from monitoring the whole covariance matrix as proposed by Aue et al. (2009b) in the retrospective case. If the covariances are monitored as well, the vector of moments that are supervised tends to be of unpropitious high dimension even if the time series itself is of moderate dimension.

The procedure is based on the monitoring technique proposed by Chu et al. (1996) who used a similar but univariate sequential method based on fluctuations to detect structural breaks in the parameter vector of a linear regression model. Their approach was refined and further investigated by Horváth et al. (2004), Aue et al. (2006) and Aue et al. (2009a), among others. Groen

et al. (2013) expanded the approach to the multivariate case. Nevertheless, even if the main goal in Groen et al. (2013) is to monitor structural changes in multivariate sequences, as it is in our case, the focus is put on the parameters of the linear regression model and not in the individual variances of the components of the sequence, thus making their approach different to our own. While authors as Berkes et al. (2004) and Aue et al. (2011) extended the field of applications to the monitoring of parameters in univariate $GARCH(p, q)$ models and high frequency portfolio betas, respectively, Wied and Galeano (2013) presented a model-independent monitoring procedure to detect changes in the correlation of bivariate time series. Retrospective methods to detect changes in the covariance or correlation structure of random vectors were proposed by Aue et al. (2009b) respectively by Wied (2014). We combine the sequential approach with the attempt to survey moments of multivariate processes.

Since whole random vectors often provide more information than single random variables, the additional information should be used to develop a procedure which enables monitoring the vector of variances of the individual components. Although we are only interested in the vector of variances, such a procedure could be able to detect changes in one or several variances more efficiently than using several univariate procedures similar to the correlation monitoring procedure proposed by Wied and Galeano (2013) that could be adapted to the situation by using the Bonferroni-Holm method.

The rest of the paper is organized as follows. Section 2 introduces the proposed monitoring procedure for detecting a changepoint in the vector of variances of a multivariate random variable as soon as possible and derives the asymptotic properties of the chosen detector. Sections 3 and 4 present a detailed simulation study and an application to real data that illustrate the behavior of the procedure in finite settings. Finally, section 5 provides some conclusions. All proofs are presented in Appendix A, while Appendix B contains tables with the simulation results from section 3.

2. The monitoring procedure

Let $(X_t,\ t \in \mathbb{Z})$ be a sequence of p dimensional random vectors whose elements possess finite fourth moments and cross moments. W.l.o.g. we assume that $\mathsf{E}(X_t) = 0$, $t \in \mathbb{Z}$. This assumption is natural in financial contexts when one considers daily log returns of financial assets. The value of interest is the vector of variances associated with the single components of the random vector $X_t = \left(X_{t1}, \ldots, X_{tp}\right)'$ denoted by

$$\sigma_t^2 = \left(\sigma_{t1}^2, \ldots, \sigma_{tp}^2\right)' \quad \text{with} \quad \sigma_{tj}^2 = \mathsf{Var}\left(X_{tj}\right) = \mathsf{E}\left(X_{tj}^2\right), \quad \text{for } j = 1, \ldots, p.$$

Since often a non time-varying variance structure cannot be assumed, we are interested in a monitoring procedure that supervises the vector of variances and reports a potential structural break as soon as possible after it has occurred. We estimate the variances from growing subsamples and compare them with estimators obtained from a reference data set that is assumed as being not affected by a variance change. In the context of sequential testing, this implies using a historical sample to obtain a first estimator of the vector of variances. In the monitoring period the historical sample is successively extended by p dimensional data points that are used to update the chosen detector. This reflects the fact that data like daily asset or index prices is observed step by step. The formal constancy assumption for the historical period of length m is

Assumption 1. $\sigma_1^2 = \ldots = \sigma_m^2$, *where m is a positive integer.*

The validity of this assumption can be checked by performing retrospective changepoint detection procedures on the historical data set, for instance a procedure similar to the one proposed by Wied et al. (2012b). In practice, it is usually possible to find a sufficient amount of historical data points with a stable variance structure.

In the following, we want to test the null hypothesis of equal vectors of variances

$$\mathsf{H}_0 : \sigma_1^2 = \ldots = \sigma_m^2 = \sigma_{m+1}^2 = \ldots$$

versus the alternative H_1 that σ_t^2 changes at one or several unknown points in the monitoring

period. Let $\mathcal{F}_B$ be the set of functions $f : [0, \infty) \to \mathbb{R}$ that are bounded and can be approximated by step functions on the interval $[0, B + 1]$. Throughout the paper the variable B indicates how much longer the monitoring period is compared to the historical data set. We consider the alternative H_1 that the individual variances can be decomposed as

$$\mathsf{Var}\left(X_{tj}\right) = \sigma_{tj}^2 = \bar{\sigma}_j^2 + g_j\left(\frac{t}{m}\right), \quad j = 1, \ldots, p, \quad t \in \mathbb{Z}, \tag{1}$$

with $\bar{\sigma}_j^2$, $j = 1, \ldots, p$, time-invariant constants and structural stability determining functions $g_j(\cdot) \in \mathcal{F}_B$, $j = 1, \ldots, p$. While $g_j(z) = 0$, for $z \in [0, 1]$ and for all $j = 1, \ldots, p$, for at least one $j \in \{1, \ldots, p\}$ let $\int_1^{B+1} \left|g_j(z)\right| dz > 0$. The property $\int_1^{B+1} \left|g_j(z)\right| dz > 0$ indicates that the variance of the j-th vector component is affected by a structural change.

In order to derive asymptotic results concerning size and power of the procedure that will be presented below, some assumptions have to be imposed first. They are counterparts of the assumptions (A1)-(A3) in Wied et al. (2012a) and to (A2)-(A4) in Wied and Galeano (2013), respectively.

Assumption 2. *For* $U_t := \left(X_{t1}^2 - \mathsf{E}\left(X_{t1}^2\right), \ldots, X_{tp}^2 - \mathsf{E}\left(X_{tp}^2\right)\right)'$ *and* $S_j := \sum_{t=1}^{j} U_t$, $j \in \mathbb{N}$, *we have*

$$\lim_{m\to\infty} \mathsf{E}\left(\frac{1}{m} S_m S_m'\right) =: D_p$$

where D_p *is a finite and positive definite matrix.*

Assumption 3. *The r-th absolute moments of the components of* U_t *are uniformly bounded for some* $r > 2$.

Assumption 4. *The process* $(X_t, t \in \mathbb{Z})$ *is* L_2*-near epoch dependent, see e.g.Davidson (1994), with size* $-\frac{r-1}{r-2}$, *where r is from Assumption 3, and constants* (c_t), $t \in \mathbb{Z}$, *on a sequence* (Y_t), $t \in \mathbb{Z}$, *which is* α*-mixing of size* $\phi^* := -\frac{r}{r-2}$, *i.e.*

$$\|X_t - \mathsf{E}\left(X_t | \sigma(Y_{t-l}, \ldots, Y_{t+l})\right)\|_2 \leq c_t v_l$$

with $\lim_{l\to\infty} v_l = 0$*, such that* $c_t \leq 2\|U_t\|_2$ *with* U_t *from Assumption 3 and* $\|\cdot\|_2$ *the* L_2*-norm .*

The proposed procedure is inspired by the model-independent fluctuation test proposed by Wied and Galeano (2013) for the detection of changes in the correlation of two sequences of random variables. The fluctuations arise from the comparison of variance estimates calculated from several subsamples of the available data. Denote by $\left[\hat{\sigma}^2\right]_k^l$ the estimate of the vector of variances calculated from X_k to X_l, $k < l$:

$$\left[\hat{\sigma}^2\right]_k^l = \left(\left[\overline{X_1^2}\right]_k^l, \cdot\cdot, \left[\overline{X_p^2}\right]_k^l\right) \quad \text{with} \quad \left[\overline{X_j^2}\right]_k^l = \frac{1}{l-k+1}\sum_{t=k}^{l} X_{tj}^2, \text{ for } j = 1, \ldots, p.$$

Then, estimates of the vector of variances from growing samples are compared to estimates resulting from the historical data. Under the hypothesis of equal vectors of variances the estimate vectors should not differ too much. In Wied and Galeano (2013) the fluctuations could easily be defined as the absolute differences of the two correlation estimates. In the multivariate setting two vectors have to be compared. Let

$$V_k = \frac{k}{\sqrt{m}}\hat{D}^{-\frac{1}{2}}\left(\left[\hat{\sigma}^2\right]_{m+1}^{m+k} - \left[\hat{\sigma}^2\right]_1^m\right) = \frac{k}{\sqrt{m}}\hat{D}^{-\frac{1}{2}}\begin{pmatrix}\left[\overline{X_1^2}\right]_{m+1}^{m+k} - \left[\overline{X_1^2}\right]_1^m \\ \vdots \\ \left[\overline{X_p^2}\right]_{m+1}^{m+k} - \left[\overline{X_p^2}\right]_1^m\end{pmatrix}, \; k \in \mathbb{N},$$

with $\hat{D}_p$ an estimator of the matrix D_p defined in Assumption 2 that is calculated from the first m observations. Possible estimation methods are for instance kernel estimation as in Wied et al. (2012a) and Wied and Galeano (2013) or bootstrapping as in Wied (2014). Simulations reveal that both estimation methods give approximately equivalent results. We prefer the kernel estimator since its calculation requires less time. Define

$$\tilde{V}_t = \frac{1}{\sqrt{m}}\tilde{U}_t \quad \text{with} \quad \tilde{U}_t = \left(X_{t1}^2 - \left[\overline{X_1^2}\right]_1^m, \ldots, X_{tp}^2 - \left[\overline{X_p^2}\right]_1^m\right)'.$$

Then, a consistent estimator of D_p is given by

$$\hat{D}_p = \sum_{t=1}^{r} \sum_{u=1}^{r} k\left(\frac{t-u}{\delta_r}\right) \tilde{V}_t \tilde{V}_u{}' \quad \text{with} \quad k(x) = \begin{cases} 1 - |x|, & |x| \leq 1 \\ 0, & otherwise. \end{cases}$$

Here, $k(x)$ is the Bartlett kernel and δ_r the bandwidth that determines up to which lag outer products of the vectors $\tilde{V}_t$ are used to calculate the estimator. The choice of the kernel is motivated by the approach in Wied et al. (2012a). However, a different bandwidth was chosen since simulations show that $\delta_r = \left[r^{\frac{1}{4}}\right]$ is the most suitable one compared to alternative bandwidths. Consistency of the estimator $\hat{D}_p$ is necessary for deriving the asymptotic distribution of the detector that is presented later in Theorem 1.

As it is desirable to construct a one dimensional detector that can be compared to the values of a univariate threshold function, possible solutions are to use either the Euclidian norm or a quadratic form of the vector of differences. The latter was considered by Aue et al. (2009b) in the retrospective setting. The detector used by our monitoring procedure is the Euclidean norm of V_k. The value of $\|V_k\|_2$ is calculated online for every k in the monitoring period. The procedure stops when the detector exceeds the value of a scaled threshold function $w(\cdot)$. As soon as this happens, the null hypothesis of no variance change cannot be taken as valid anymore and is rejected. Accordingly, the stopping rule can be defined as

$$\tau_m = \min\left\{k \leq [mB] : \|V_k\|_2 > c \cdot w\left(\frac{k}{m}\right)\right\}, \tag{2}$$

with $w(\cdot)$ a positive and continuous function and c a constant chosen such that under a valid null hypothesis $\lim_{m\to\infty} \mathsf{P}(\tau_m < \infty) = \alpha \in (0,1)$ is the test significance level. Along the lines of Aue et al. (2011) we write $\tau_m < \infty$ to indicate that the detector has exceeded the threshold function $cw(\cdot)$ in the monitoring period which implies a rejection of the hypothesis of equal vectors of variances. If $\|V_k\|_2$ does not exceed the corresponding value of the threshold function in the whole monitoring period, we mark this by $\tau_m = \infty$, see Aue et al. (2011). This leads to our main result:

Theorem 1. *Under* H_0*, Assumptions 1-4 and for any* $B > 0$*,*

$$\lim_{m\to\infty} \mathsf{P}\left(\tau_m < \infty\right) = \lim_{m\to\infty} \mathsf{P}\left(\sup_{b\in[0,B]} \frac{\|V_{\lfloor m\cdot b\rfloor+2}\|_2}{w\left(b\right)} > c\right) = \mathsf{P}\left(\sup_{b\in[0,B]} \frac{\|\mathcal{G}(b)\|_2}{w\left(b\right)} > c\right), \tag{3}$$

where $\left\{\mathcal{G}(b) = \left(G_1(b), \ldots, G_p(b)\right)', \; b \in [0, B]\right\}$ *is a p-variate stochastic process whose component processes are p independent mean zero Gaussian processes* $\left\{G_j(b), b \in [0, B]\right\}$ *with covariance function* $\mathsf{E}\left(G_j(k)G_j(l)\right) = \min(k, l) + kl$*, for* $j = 1, \ldots, p$*.*

Theorem 1 establishes the asymptotic behavior of the monitoring procedure based on the stopping rule τ_m in (2). As argued in detail in Aue et al. (2011) and Wied and Galeano (2013) the limiting probability in (3) can be led back to the behavior of p independent standard Brownian motions $\left\{W_j\left(b\right) : b \in [0, 1]\right\}$, $j = 1, \ldots, p$. Since $\left\{G_j\left(b\right) : b \in [0, B]\right\}$ has the same distribution as $\left\{(1 + b)\, W_j\left(b/\left(1 + b\right)\right) : b \in [0, B]\right\}$, for $j = 1, \ldots, p$, we have

$$\sup_{b\in[0,B]} \frac{\|\mathcal{G}(b)\|_2}{w\left(b\right)} = \sup_{b\in[0,B]} \frac{\sqrt{\sum_{j=1}^{p}\left[G_j(b)\right]^2}}{w\left(b\right)} \overset{L}{=} \sup_{b\in[0,B]} \frac{(1 + b)\,\sqrt{\sum_{j=1}^{p}\left[W_j\left(\frac{b}{1+b}\right)\right]^2}}{w\left(b\right)} \tag{4}$$

with $A_1 \overset{L}{=} A_2$ indicating that A_1 and A_2 possess the same distribution. As in Wied and Galeano (2013) the threshold function $w(\cdot)$ can be chosen as

$$w\left(b\right) = (1 + b) \cdot \max\left\{\left(\frac{b}{1 + b}\right)^{\gamma}, \epsilon\right\} \tag{5}$$

with $\gamma \in \left[0, \frac{1}{2}\right)$ and $\epsilon > 0$ a fixed constant that solely serves to guarantee the divisibility by $w(\cdot)$ and can be chosen arbitrarily small in applications. The parameter γ can be used to adjust the procedure such that it performs at its best in a certain expected situation. As discussed in Wied and Galeano (2013) in detail, there is a trade off between the aim to detect arisen structural breaks as soon as possible and the purpose to reduce the probability of type I errors to the significance level. A value of γ chosen closely to $\frac{1}{2}$ tends to cause a soon rejection of the null hypothesis. This is desirable if a structural change is expected to take place shortly after the beginning of the monitoring period, but also tends to produce type I errors, while the null is still

valid. In contrast, using a smaller value for γ rather results in a reduction of type I errors but also leads to a testing routine that is less capable of indicating structural breaks arising early in the monitoring period. Further simulations show that large values of γ lead to unacceptable high percentages of falsely rejected null hypotheses especially for higher dimensions of the random vectors under supervision. Hence, in the following no γ values larger than 0.25 are considered. Substituting $w(\cdot)$ from (5) and defining $u = \frac{b}{1+b}$ as well as $s = \frac{u(1+B)}{B}$ allows to write expression (4) along the lines of Wied and Galeano (2013) as

$$\sup_{b\in[0,B]} \frac{\|\mathcal{G}(b)\|_2}{w(b)} \stackrel{L}{=} \sup_{u\in\left[0,\frac{B}{1+B}\right]} \frac{\sqrt{\sum_{i=1}^{p}\left[W_i(u)\right]^2}}{\max\left\{u^\gamma, \epsilon\right\}} \stackrel{L}{=} \sup_{s\in[0,1]} \left(\frac{B}{1+B}\right)^{\frac{1}{2}-\gamma} \frac{\sqrt{\sum_{i=1}^{p}\left[W_i(s)\right]^2}}{\max\left\{s^\gamma, \epsilon\left(\frac{1+B}{B}\right)^\gamma\right\}}.$$

Since under the conditions of Theorem 1

$$\lim_{m\to\infty} \mathrm{P}(\tau_m < \infty) = \mathrm{P}\left(\sup_{s\in[0,1]} \left(\frac{B}{1+B}\right)^{\frac{1}{2}-\gamma} \frac{\sqrt{\sum_{i=1}^{p}\left[W_i(s)\right]^2}}{\max\left\{s^\gamma, \epsilon\left(\frac{1+B}{B}\right)^\gamma\right\}} > c \right),$$

Monte Carlo simulations can be used to obtain the constant $c = c(\alpha)$ such that

$$\mathrm{P}\left(\sup_{s\in[0,1]} \left(\frac{B}{1+B}\right)^{\frac{1}{2}-\gamma} \frac{\sqrt{\sum_{i=1}^{p}\left[W_i(s)\right]^2}}{\max\left\{s^\gamma, \epsilon\left(\frac{1+B}{B}\right)^\gamma\right\}} > c(\alpha) \right) = \alpha,$$

for any $\alpha \in (0,1)$. Thus, the probability of a false alarm is approximately α if m is large enough.

		$\alpha = 0.01$			$\alpha = 0.05$			$\alpha = 0.1$		
γ	B	$p = 2$	$p = 5$	$p = 10$	$p = 2$	$p = 5$	$p = 10$	$p = 2$	$p = 5$	$p = 10$
0	0.5	1.9062	2.3268	2.8462	1.5514	2.0265	2.5802	1.3991	1.8817	2.4146
	1	2.2924	2.8653	3.5217	1.9039	2.4659	3.1544	1.7003	2.3122	2.9439
	2	2.6246	3.3371	4.0214	2.1915	2.8704	3.6375	1.9737	2.6447	3.4005
0.25	0.5	2.5231	3.1579	3.8898	2.1439	2.7760	3.4385	1.9431	2.5872	3.2596
	1	2.8124	3.4880	4.2737	2.3881	3.0361	3.8051	2.1627	2.8457	3.6051
	2	2.9854	3.7461	4.5824	2.5351	3.2927	4.1315	2.3001	3.0523	3.8723

Table 1: Simulated critical values $c(\alpha)$.

Simulated critical values for all combinations of $p \in \{2, 5, 10\}$, $B \in \{0.5, 1, 2\}$, $\gamma \in \{0, 0.25\}$ and for significance levels of $\alpha \in \{0.01, 0.05, 0.1\}$ can be taken from Table 1. To obtain the values

of c, 10.000 Brownian motions are simulated on a grid of 10.000 equidistant points.

Up to now, we have focused on the behavior of the detector under the null hypothesis. In the considered case the alternative is rather broad including scenarios with a single or multiple structural breaks in one or several vector components as well as variance changes of minor or major magnitude. This suggests investigating the testing power against local alternatives.

Assumption 5. *For the process* $(X_t,\ t \in \mathbb{Z})$ *with* $X_t = \left(X_{t1}, \ldots, X_{tp}\right)'$ *the variances of the individual vector components can be decomposed as*

$$\mathsf{Var}\left(X_{j,t}\right) = \bar{\sigma}_j^2 + \frac{1}{\sqrt{m}} g_j\left(\frac{t}{m}\right), \quad j = 1, \ldots, p,$$

with $\bar{\sigma}_j^2$ *and* $g_j(\cdot),\ j = 1, \ldots, p$, *as in (1).*

Theorem 2. *Under a sequence of local alternatives, Assumptions 1-5 and for any* $B > 0$

$$\lim_{m\to\infty} \mathsf{P}\left(\tau_m < \infty\right) = \lim_{m\to\infty} \mathsf{P}\left(\sup_{b\in[0,B]} \frac{\|V_{\lfloor m\cdot b\rfloor+2}\|_2}{w\left(b\right)} > c\right) = \mathsf{P}\left(\sup_{b\in[0,B]} \frac{\|\mathcal{G}(b) + h(b)\|_2}{w\left(b\right)} > c\right),$$

with $\left\{\mathcal{G}(b) = \left(G_1(b) \ldots G_p(b)\right)',\ b \in [0, B]\right\}$ *as in Theorem 1 and* $h(\cdot) = H \cdot \left(h_1(\cdot), \ldots, h_p(\cdot)\right)'$. *Up to a constant,* H *is the limit of* $\hat{D}_p$ *under* H_0, *while the function* $h_j(b) := \int_1^{b+1} g_i(u)du = 0$ *for all* $b \in [0, B]$ *if and only if the* j*-th component is not affected by a variance change.*

Theorem 2 yields that even a small variance change in just one single component can be detected with high probability if the historical period is large enough. To obtain general statements about the testing power, the magnitude of a variance change is assumed to tend to ∞. This can be modeled by defining one of the structural stability determining functions $g_j(\cdot)$, $j = 1, \ldots, p$, as a scaled function $g_*(\cdot)$ and assume the scaling factor to tend to ∞ implying an increasing magnitude of a shift in the respective component of the vector of variances.

Assumption 6. *At least one of the structural stability determining functions* $g_j(\cdot) \in \mathcal{F}_B$ *with* $\int_1^{b+1} \left|g_j(z)\right| dz > 0$ *can be decomposed as* $g_j(\cdot) = M \cdot g_*(\cdot)$ *with* $g_*(\cdot) \in \mathcal{F}_B$.

Under the alternative of at least one structural break in the vector of variances in the monitoring

period and Assumption 6, let $\mathsf{P}_{\mathsf{H}_1}(M)$ be the probability that the detector exceeds the threshold function during the monitoring period for given M.

Theorem 3. *Let $g_*(\cdot) \in \mathcal{F}_B$ be arbitrary but fixed. Under Assumptions 1-4 and 6, for every $\varepsilon > 0$ there exists an M_ε such that for all $M > M_\varepsilon$*

$$\lim_{m\to\infty} \mathsf{P}_{\mathsf{H}_1}(M) > 1 - \varepsilon.$$

Theorem 3 yields that a variance change of sufficiently high magnitude will be detected with given probability if the length of the historical period tends to ∞ even if just one single component is affected by the change or if multiple components experience contrary variance changes.

If the detector actually exceeds the threshold function, the presence of a structural change is indicated. This leads to the challenge to determine the location of the changepoint. This does not necessarily have to coincide with the first hitting time τ_m. In fact, an abrupt change of the variances will often take time to affect the detector strongly enough to get identified by the procedure. A possible estimator of the changepoint location is a multivariate equivalent to the one used by Wied et al. (2012a) and Wied and Galeano (2013):

$$\hat{k} = \underset{1\leq j\leq \tau_m - 1}{\arg\max} D_{j,\tau_m} \quad \text{with} \quad D_{j,\tau_m} := \frac{j}{\sqrt{\tau_m}} \hat{D}^{-\frac{1}{2}} \left\| \begin{pmatrix} \left[\overline{X_1^2}\right]_{m+1}^{m+j} - \left[\overline{X_1^2}\right]_{m+1}^{m+\tau_m-1} \\ \vdots \\ \left[\overline{X_p^2}\right]_{m+1}^{m+j} - \left[\overline{X_p^2}\right]_{m+1}^{m+\tau_m-1} \end{pmatrix} \right\|_2 .$$

This type of estimator led to satisfying results in the univariate case, hence we use it to estimate the location of an indicated changepoint. However, a detailed analysis of the estimator's properties lies beyond the scope of this paper. The performance of the proposed procedure as well as the properties of the first hitting times τ_m, the estimated changepoint locations $\hat{k}$ and the estimated location fractions $\hat{\lambda} = \frac{\hat{k}}{mB}$ will be investigated in the following section.

3. Simulations

This section is devoted to a performance analysis of the proposed monitoring procedure in finite samples. Each of the regarded scenarios is constructed using different tuning parameters. First, the dimension of the random vectors is chosen as $p \in \{2, 5, 10\}$. Since all of the asymptotics are based on the length of the historical period tending to ∞, large values of m are considered. We choose $m \in \{500, 1.000, 2.000\}$. In the context of financial data like asset returns, these values correspond to time periods of approximately 2, 4 and 8 years. It is important to note that smaller values of m may lead to noninvertible estimates of D_p in practice, especially for higher dimensions p. On the other hand, since the historical period must be assumed to be free from variance changes, larger values for m can hardly be found in practice. Furthermore, we choose $B \in \{0.5, 1, 2\}$ implying that the monitoring period is shorter, of the same length or longer than the historical period. Finally, the parameters in the threshold function $w(\cdot)$ have to be specified: ϵ is chosen as 10^{-6} in all of the following simulation settings and $\gamma \in \{0, 0.25\}$. These values represent the aim to detect changes that are expected to occur earlier or later in the monitoring period. The theoretical size used for all of the simulations is $\alpha = 0.05$. In each case 10.000 time series are simulated. To simulate $c(\alpha)$ 10.000 Brownian motions are simulated on a grid of 10.000 equidistant points.

3.1. Monitoring time series of i.i.d. random vectors

To begin with, we investigate the size of the proposed procedure under the null hypothesis of no structural break. First, we simulate time series that capture neither serial nor cross-sectional dependence to gain reference values to which the performance in more complex scenarios can be compared. As simplest possible case, realizations of processes of i.i.d. random vectors are simulated. These simulation results can work as a benchmark for more complex simulation scenarios. The random vectors under consideration are i.i.d. multivariate normal and multivariate t distributed with $\nu = 8$ degrees of freedom. As covariance matrix the matrices

$$\Sigma_2 = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}, \; \Sigma_5 = \begin{pmatrix} 1 & 0.3 & 0.4 & 0.5 & 0.6 \\ 0.3 & 1 & 0.3 & 0.4 & 0.5 \\ 0.4 & 0.3 & 1 & 0.3 & 0.4 \\ 0.5 & 0.4 & 0.3 & 1 & 0.3 \\ 0.6 & 0.5 & 0.4 & 0.3 & 1 \end{pmatrix} \text{ and } \Sigma_{10} = \begin{pmatrix} 1 & 0.1 & 0.11 & 0.12 & 0.15 & 0.2 & 0.25 & 0.3 & 0.35 & 0.4 \\ 0.1 & 1 & 0.1 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0.35 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0.35 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0.1 & 1 & 0.1 \\ 0.4 & 0.35 & 0.3 & 0.25 & 0.2 & 0.15 & 0.12 & 0.11 & 0.1 & 1 \end{pmatrix}$$

are used as well as identity matrices I_p of corresponding dimension p. To enable a reasonable comparison of the results, the covariance matrices in the case of the t distribution are standardized by multiplying $\frac{\nu-2}{\nu}$. The results are given in Tables 4 and 5 in Appendix B and illustrated in Figure 1. For the sake of clarity and since the results differ only slightly for the different values of the tuning parameter γ and the different types of covariance matrix, the figure only shows the empirical sizes for $\gamma = 0$ and an identity covariance matrix.

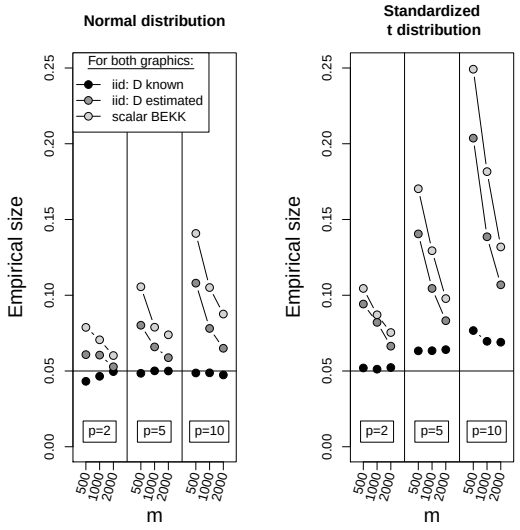


Figure 1: Size comparison: i.i.d. random vectors with matrix D_p known and estimated, respectively, and scalar $BEKK$ time series.

The fact that there is hardly a difference in the empirical sizes depending on whether the covariance matrix is diagonal or not was expected as our procedure is only based on estimates of the main diagonal elements of the covariance matrix and not of the remaining entries. In general, the empirical size increases with the dimension. In order to determine the source of this, we use the actual matrix D_p that can be easily calculated for an identity covariance matrix

I_p and its standardized analogue $\frac{\nu-2}{\nu} I_p$, respectively. The matrix is given as $D_p = 2I_p$ for normal distributed random vectors and as $D_p = \frac{2(\nu-1)}{\nu-4} I_p$ for t distributed ones. The results are also given in Tables 4 and 5 in Appendix B and illustrated in Figure 1. They state that the main fraction of the increased size is caused by an insufficient estimation of the matrix D_p. Furthermore, heavy tails in the distribution of the random vectors entail an additional size increase. Unfortunately, the estimation of the matrix D_p could not be improved by using an alternative bandwidth or estimation procedure. The empirical size is distinctly larger in the case of the t distributed random vectors and decreases with growing length of the historical period m. This convergence to the theoretical size goes back to the fact that all of the asymptotic statements are established for $m \to \infty$. While for the different values of B - indicating different lengths of the monitoring period - no tendency in the empirical sizes can be recognized, larger values of γ result in a slight increase of the sizes. This is a plausible result as larger values of this parameter tend to sensitize the procedure for changes that are expected early in the monitoring period at the expense of increased probabilities of false alarms.

In the following, the power of the monitoring procedure is investigated considering two different types of scenario. In both cases the covariance matrix in the pre-break period equates the matrix Σ_p whose diagonal elements are affected by a structural break later in the series. In the first setting, the variances of all components increase from 1 to 1.3. In the second one, the variance of only one of the components jumps to 1.5. In both scenarios the power to detect an early and a later occurring change are compared. Since the length of the monitoring period depends on the parameters m and B, we assume that, independent of the length of the time series, the change happens at the same fraction of the monitoring period indicated by $\lambda^* \in (0, 1)$. We choose $\lambda^* \in \{0.05, 0.5\}$ to mark changepoints located at the beginning ($k^* = 0.05mB$) or in the middle ($k^* = 0.5mB$) of the monitoring period. The results for the first scenario are given in Tables 6 and 7 in Appendix B and illustrated in Figure 2, while those for the second scenario are presented in Tables 8 and 9 and illustrated in Figure 3. They state that the power increases considerably with growing length of the historical and the monitoring period. If all

of the variances are affected by a change, the power increases with growing dimension of the random vectors. If only one of the variances experiences a change, the frequency of detecting the change decreases for growing dimension p since the portion of variance components that are not struck by the change increases. Early changes can be detected reliably in both scenarios. However, the power gets quite low if the changepoint is located in the advanced series, especially for t distributed random vectors as the rejection fractions in Tables 7 and 9 state. The direct comparison of the two scenarios shows that a major change in just one of the variances can be detected more frequently than a minor change that affects all of the variances only when

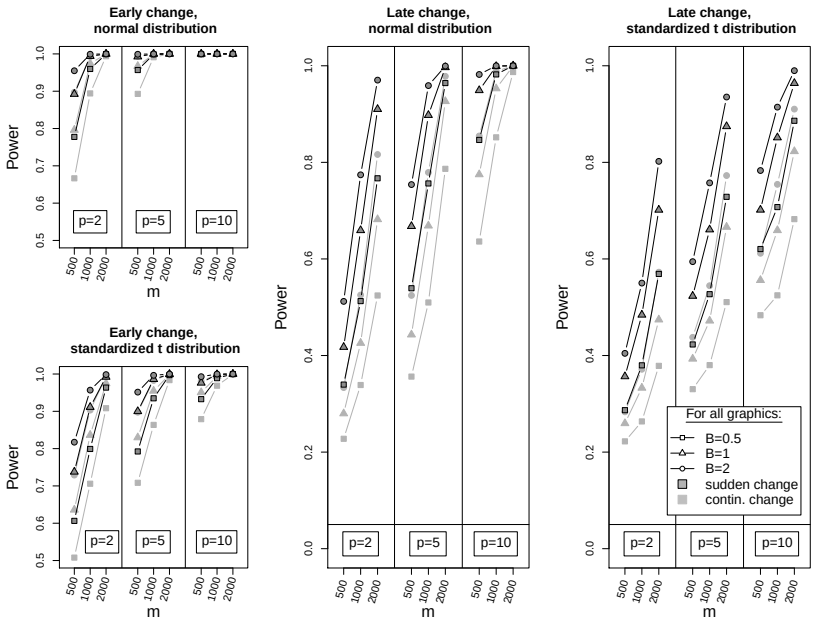


Figure 2: Power: i.i.d. random vectors when all of the variances are affected by a change.

the dimension is rather small and the change occurs not too late in the monitoring period. In all of the settings the procedure performs worse in the case of t distributed random vectors, but the differences to the normal distribution results are declining with m. Also, in most cases the power is lower for the higher value of γ. While for a later change this is a plausible result, it contradicts the expectation that early changes can be detected more frequently using a higher value of γ. An explanation for this result is that in both cases the values of the detector are compared to the values of the scaled threshold function that has a higher slope in the case of the larger tuning parameter. Since both functions intersect the down scaling of the differences

by multiplying $\frac{k}{\sqrt{m}}$ can cause an earlier crossing of the threshold function for $\gamma = 0$ than for $\gamma = 0.25$. Overall, changes that occur right after the beginning of the monitoring period can be detected much more frequently than those located in the advanced monitoring period no matter how the tuning parameter was chosen.

Now, the results can be compared to scenarios of continuously appearing changes, i.e., a slow linear increase of the affected variances that starts at $\lambda_1^* = 0.05$ and 0.5, respectively, and is completed at $\lambda_2^* = 0.3$ and 0.75, respectively. The results are also illustrated in Figures 2 and 3 and presented in Tables 6-9 in Appendix B as values in parentheses. The impact of variations

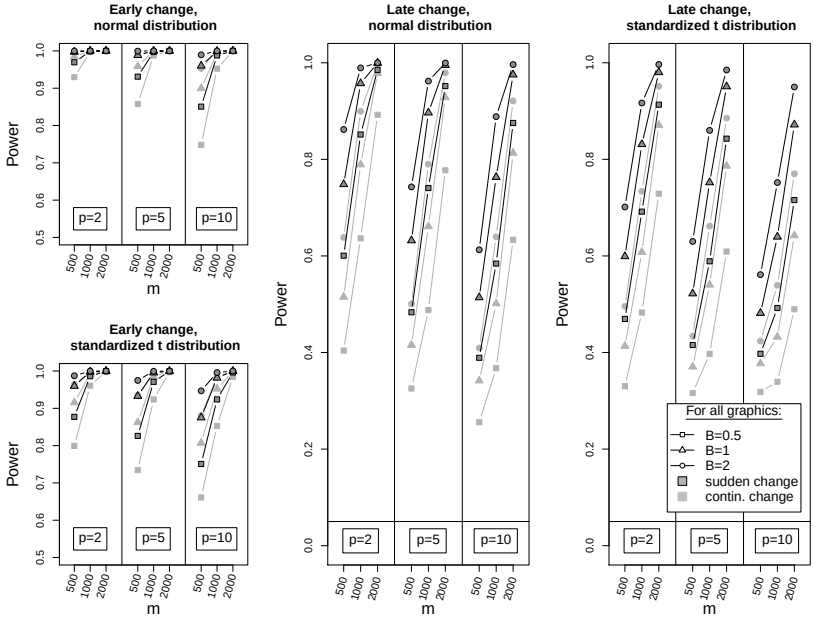


Figure 3: Power: i.i.d. random vectors when only one of the variances is affected by a change.

in the parameters remains the same as in the situation of a sudden variance change. However, the power is considerably lower in the case of a slow increase. Since the power simulations for sudden changes suggest that later changes can be detected less frequently it is clear that changes that are completed later in the monitoring period are more difficult to be detected. Although in our simulations the detectability of changes that start in the advanced monitoring period is kind of low especially for short historical periods, the power increases quickly with growing length of the historical period.

3.2. Monitoring scalar $BEKK$ time series

Up to now, the random vectors under consideration only possessed dependence between the individual vector components. In practice, time series that additionally exhibit serial dependence, which is permitted in moderate magnitude by Assumption 4, are of larger interest. In financial contexts, it may be desirable to detect changes in the vector of unconditional variances of random vectors whose conditional covariance matrices are expected to be time-varying. A common way to model this behavior and to explain volatility clusters that usually can be observed in financial time series is to use a multivariate $GARCH$ model. Assume

$$X_t = H_t^{\frac{1}{2}} \varepsilon_t, \tag{6}$$

where $(\varepsilon_t,\ t \in \mathbb{Z})$ is a sequence of i.i.d. $\mathbb{R}^p$-valued random vectors, $H_t^{\frac{1}{2}}$ is the square root of the conditional covariance matrix $H_t = Cov(X_t|\mathcal{I}_t)$ and $\mathcal{I}_t = \sigma(X_{t-1}, X_{t-2}, \ldots)$ is the information set at time t. Since Bollerslev (1986) states that even $GARCH$ models of low order are able to explain the behavior of many financial time series well, we will focus on models that are solely based on lagged conditional covariance matrices and observations of first order. To specify the conditional covariance matrix we use the two parameter model, see Ding and Engle (2001), that arises from the scalar diagonal model when performing variance targeting as in Engle and Mezrich (1996) and that is a special case of the $BEKK(1,1,1)$ model proposed by Engle and Kroner (1995). We will refer to the scalar BEKK model in the following. Here, the conditional covariance matrix is recursively defined by

$$H_t = (1 - \alpha - \beta) H + \alpha X_{t-1} X_{t-1}' + \beta H_{t-1}, \tag{7}$$

where α and β are positive scalars with $|\alpha + \beta| < 1$ to guarantee stationarity and H is the unconditional covariance matrix of X_t, $t \in \mathbb{Z}$. The following lemma, whose proof can be found in Appendix A, provides a useful help to check the validity of Assumption 3. Let $\Gamma := \mathsf{E}\left(X_t^2 X_t^{2\prime}\right)$ with $X_t^2 = \left(X_{t1}^2, \ldots, X_{tp}^2\right)'$ be the matrix of fourth moments and cross moments of X_t. Denote by $vec(\cdot)$ the vec operator that stacks the columns of a matrix in a vector of dimension p^2

and $vech(\cdot)$ the *vech* operator that stacks only the lower triangular part including the main diagonal of a symmetric matrix in a vector of dimension $d = \frac{1}{2}p(p+1)$. Let D_p and L_p with $vec(A) = D_p vech(A)$ and $vech(A) = L_p vec(A)$ be the duplication and elimination matrix, respectively. These matrices are, as well as the transformation matrix G_p, defined in Hafner (2003).

Lemma 1. *The matrix* Γ *exists if and only if all the eigenvalues of*

$$Z := (A_1 \otimes A_1)\, G_p + A_1 \otimes B_1 + B_1 \otimes A_1 + B_1 \otimes B_1$$

with $A_1 = \alpha L_p D_p$ *and* $B_1 = \beta L_p D_p$ *are smaller than one in modulus.*

This postulation leads to a strong restriction of the parameter space. Only small values of α define processes that satisfy the assumptions even for random vectors of dimension 5 or 10. Thus, when applying the procedure to real data there might be a large chance to falsely detect changes that in fact did not appear. Additional simuations show that there are indeed considerable size distortions and that we get false signals right at the beginning of the observations period. In our application to real asset returns this has to be kept in mind. An approach to tackle this problem is presented in section 4.

Since the parameter space is restricted by the assumptions, the results do not differ very much for different choices of the parameters. Thus, we consider just one parameter combination chosen as $(\alpha, \beta) = (0.03, 0.45)$. The innovation vectors ε_t are i.i.d. multivariate standard normal distributed and standardized t distributed with 8 degrees of freedom. All variable parameters are chosen as in the i.i.d. case. According to the dimension $p \in \{2, 5, 10\}$, the unconditional covariance matrix H is chosen as Σ_p from Section 3.1. Results concerning the empirical size are presented in Tables 10 and 11 in the Appendix B. To simplify the comparison to the i.i.d. case, the results are also illustrated in Figure 1 for $\gamma = 0$. The size is slightly higher in the case of serial dependence. The influence of parameter variations is similar to the i.i.d. case disregarding the fact that large values of B cause a size decrease when serial dependence is present.

Since each of the considered time series consists of a bundle of p univariate possibly correlated processes one could think about monitoring the single component series with the univariate equivalent of the procedure with detector

$$\|V_k\|_2 = \frac{k}{\sqrt{m}} \hat{D}_1^{-\frac{1}{2}} \left| \left[\hat{\sigma}_j^2 \right]_{m+1}^{m+k} - \left[\hat{\sigma}_j^2 \right]_1^m \right|, \quad j = 1, \ldots, p, \tag{8}$$

where $\hat{D}_1$ is a scalar. To fortify why the multivariate approach should be preferred to the univariate one we compare size and power when monitoring scalar $BEKK$ time series. To guarantee that asymptotically the probability of type I error, i.e. that one of the p detectors (8) exceeds the threshold function during the monitoring period, does not exceed $\alpha = 0.05$, the significance levels are adjusted by using the Bonferroni-Holm method. The simulated sizes

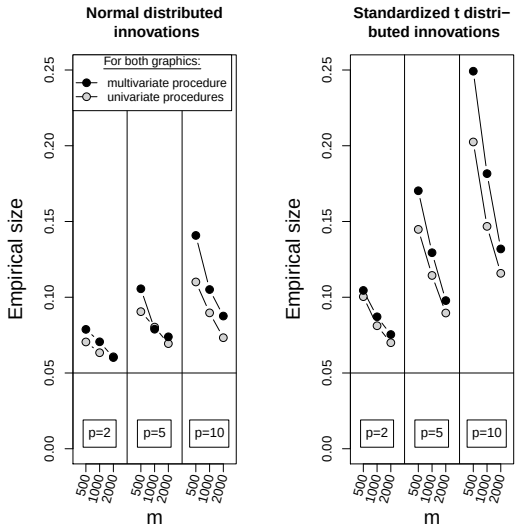


Figure 4: Size comparison: scalar BEKK time series under the use of a multivariate or several univariate procedures.

are presented in Tables 10 and 11 in Appendix B. Also, the results for $\gamma = 0$ and $B = 1$ are illustrated in Figure 4 for the multivariate and the univariate procedure. The size is slightly lower for the univariate procedures, but the differences decline with m. Moreover, the problem of an increased error I probability when monitoring realizations of random vectors with heavy tailed distribution cannot be avoided by using univariate procedures.

Next, the multivariate and univariate procedure are confronted with alternative scenarios cor-

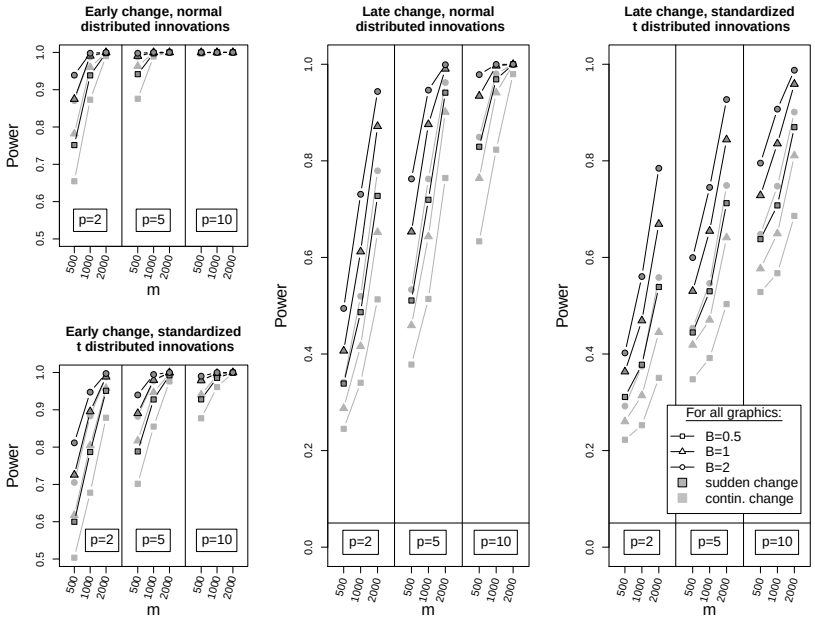


Figure 5: Power: scalar *BEKK* time series under the use of the multivariate procedure when all of the variances are affected by the change.

responding to those presented in 3.1. The results for the multivariate procedure are given in Tables 12-15 and those for univariate procedures in Tables 16-17 in Appendix B. Figures 5 and 6 illustrate the performance of the multivariate procedure. Although the time series exhibit

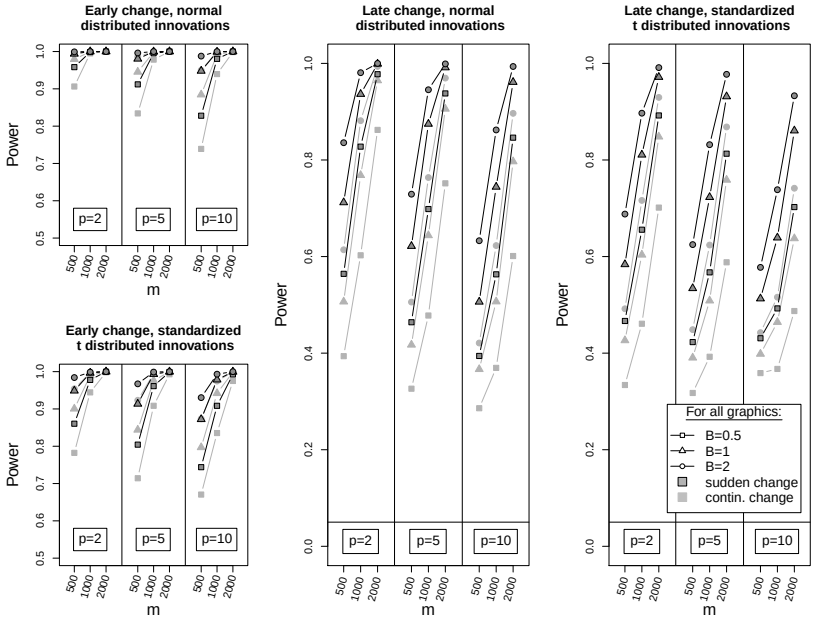


Figure 6: Power: scalar *BEKK* time series under the use of the multivariate procedure when only one of the variances is affected by the change.

serial dependence, the results resemble those of the i.i.d. case very strongly. The power is slightly lower, but the impact of changes in the variable parameters remains the same.

The power results for the univariate procedure are illustrated in Figure 7. Since the results resemble strongly those of the multivariate procedure, the figure only shows the values for sudden changes and normal distributed innovations. To simplify the comparison, the graphic also contains the rejection frequencies for the multivariate procedure. As in the i.i.d. case, early changes can be detected reliably by both procedures while later changes are detected by the multivariate procedure more frequently. The latter one especially shows its strength when all of the variances experience a minor change or if just one of the variances is affected by a larger change but the historical period is rather short. Unfortunately, the higher power of the multivariate procedure goes along with a slightly increased size compared to the univariate procedure.

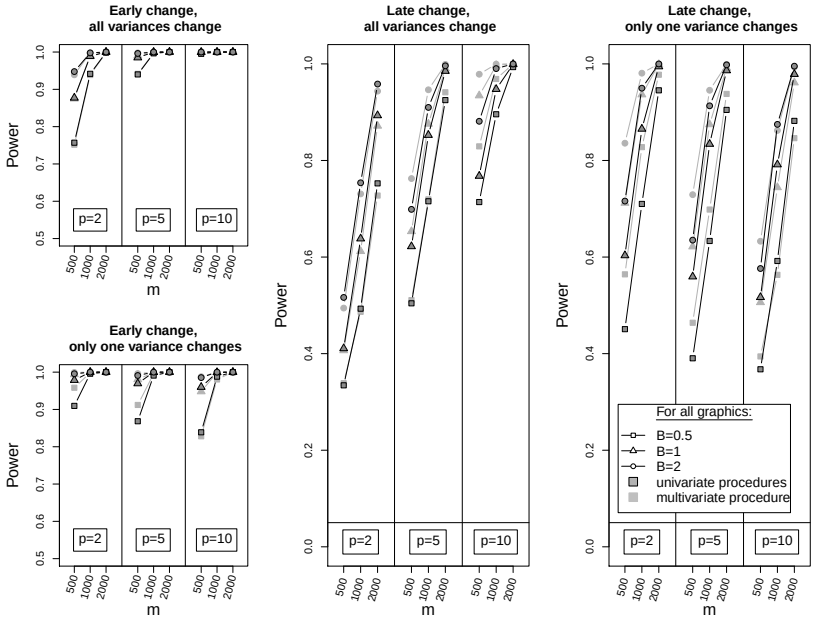


Figure 7: Power: scalar $BEKK$ time series with $N\left(0, \Sigma_p\right)$ distributed innovations under the use of several univariate procedures.

Since it is not only of interest to detect changes in the vector of variances but also to signalize their presence as soon as possible after they have occurred, we look closer at the properties of the first hitting times τ_m and the estimated changepoint locations $\hat{k}$ or location fractions $\hat{\lambda} = \frac{\hat{k}}{mB}$.

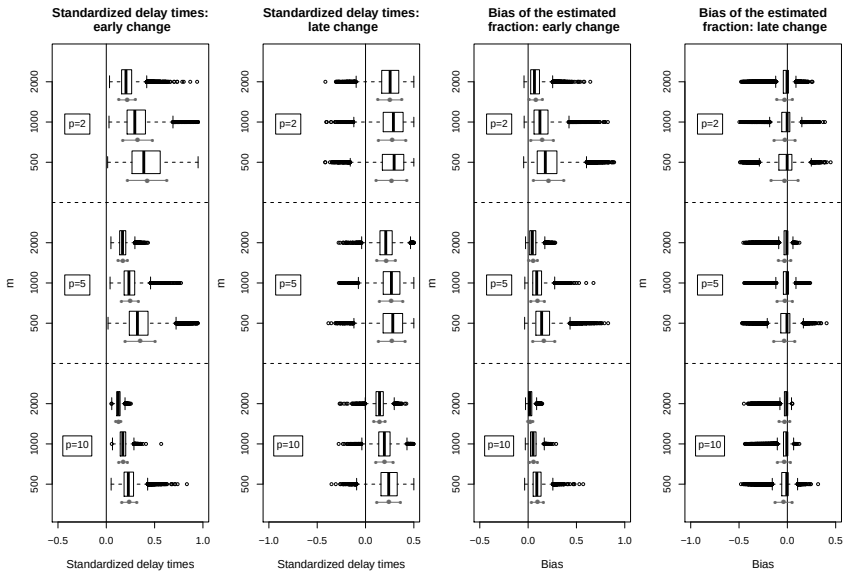


Figure 8: Properties of the standardized delay times and the estimated location fractions $\hat{\lambda}$ when monitoring scalar $BEKK$ time series with $N\left(0, \Sigma_p\right)$ distributed innovations and all of the variances are affected by the change.

While only the results for $\gamma = 0$ are visualized, the remaining parameters take the same values as before. To simplify the comparison for different sample lengths, Figures 8 and 9 illustrate the standardized delay times $d_m := \frac{\tau_m - k^*}{mB}$ and the bias of the location fraction estimator. Right under the boxplots, the graphics also show the means ± the standard deviations of the respective

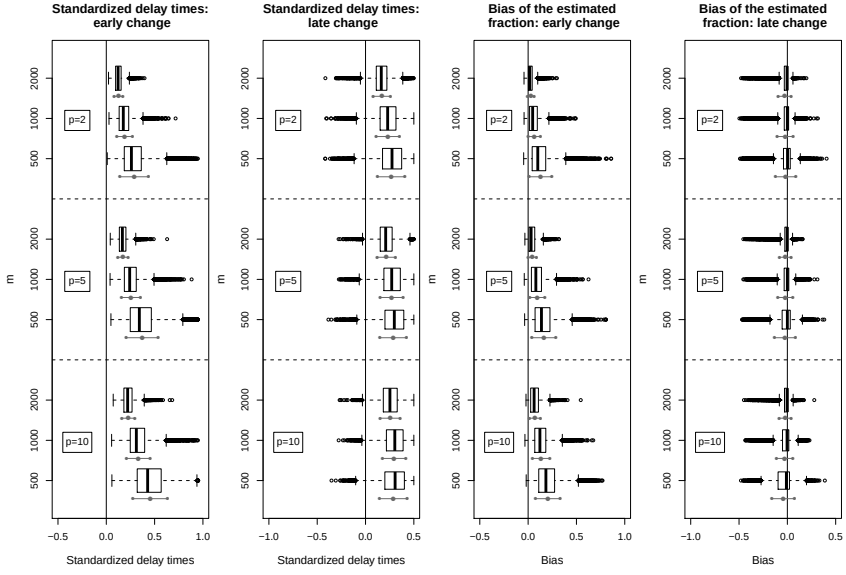


Figure 9: Properties of the standardized delay times and the estimated location fractions $\hat{\lambda}$ when monitoring scalar $BEKK$ time series with $N\left(0, \Sigma_p\right)$ distributed innovations and only one of the variances is affected by the change.

group. In general, the delay times decrease with growing length of the historical period and dimension. For small dimensions the procedure stops earlier if only a part of the variances is affected by a mayor change while for higher dimensions the delay time is shorter for smaller changes that affect more or all of the variances. This is in line with the power results discussed before.

Since the first hitting times determine which fraction of the data set is used to estimate the changepoint location, it is expected that the properties of the location fraction estimator resemble those of the first hitting times.

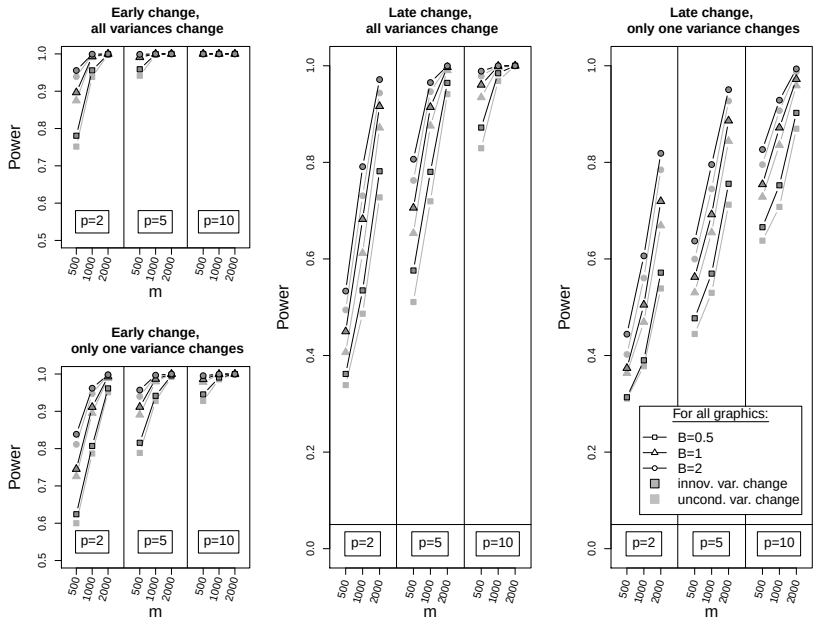


Figure 10: Power: scalar $BEKK$ time series with $N\left(0,\Sigma_p\right)$ distributed innovations and the variances of the innovations change.

So far, we considered changes that affect the diagonal elements of the unconditional covariance matrix H directly. Now, assume that the variances of the innovation vectors ε_t jump from 1 to 1.5. The power results are presented in Table 18 in Appendix B and illustrated in Figure 10 along with the results for comparable changes that affect the elements of the unconditional covariance matrix H directly. Changes in the innovations' variances can be detected almost as reliably as changes that affect the main diagonal entries of H. This is a plausible result considering the model structure in (6).

To complete the simulation study, we illustrate the behavior under increasing magnitudes of the changes. Figure 11 shows the rejection frequencies of the procedure given different magnitudes of an earlier or later shift that affects all of the variances. Assume that all of the variances equal 1 before the change and experience a change of magnitude $\Delta \in \{-0.7, -0.6, \ldots, 0.6, 0.7\}$. The investigation is limited to the case of a historical period consisting of 1.000 observations, a monitoring period that is as long as the historical data set and scalar $BEKK$ time series with multivariate normal or standardized t distributed innovation vectors. To ensure invertibility of the covariance matrix, H is chosen as the identity and standardized identity matrix, respectively.

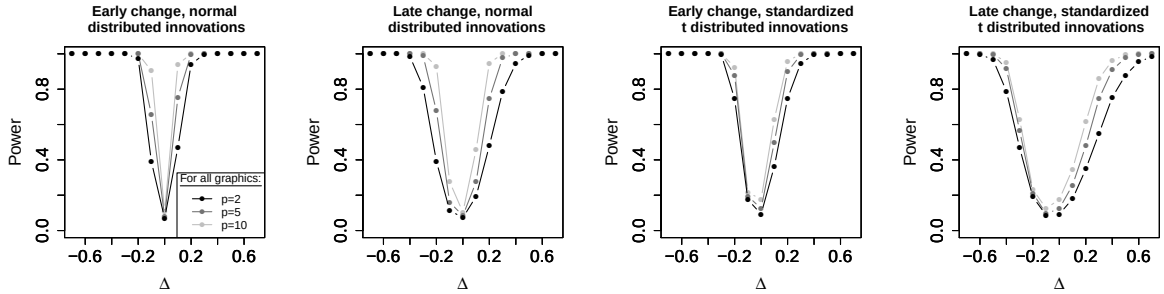


Figure 11: Power: scalar $BEKK$ time series with $N\left(0, \sigma^2 I_p\right)$ or $t_\nu\left(0, \frac{\nu-2}{\nu}\sigma^2 I_p\right)$ distributed innovations and variance shifts of varying magnitude Δ.

In line with the previous results, the power approaches 1 with increasing absolute magnitude of the variance change and dimension. Besides, it is noticeable that for smaller absolute values of Δ the change is detected more frequently in the case of increasing variances compared to a decrease of the same amount.

4. Real data example

Finally, we use the proposed procedure to monitor a time series of log returns, namely those of the DAX listed assets of $Allianz$, $Bayer$, $DeutscheBank$, RWE and $Siemens$ from 1979 to 2014. Fitting a scalar $BEKK$ model to the observations suggests that the parameters α and β are considerably higher than allowed by Lemma 1 such that the process does not fulfill Assumption 3. To circumvent this problem, we use GARCH residuals as inputs in the detector. Additional simulations show that filtering multivariate $GARCH$ time series and monitoring the residual

vectors leads to empirical sizes close to those in the i.i.d. case indicating that the time series of *GARCH* residuals show similar characteristics as the underlying innovation vectors. A detailed investigation of this behavior that was considered for the univariate case in Kulperger and Yu (2005) is the object of future research.

The significance level for all following applications is $\alpha = 0.05$. Our approach is as follows: The parameters α and β as well as the unconditional covariance matrix H are estimated from a historical data set of length $m \in \{500, 1.000\}$ via two stage quasi maximum likelihood estimation as described in Pedersen and Rahbek (2014). Since longer historical periods rather tend to be affected by variance change, m is limited to a maximum of 1.000 observations. After the estimation, the monitoring procedure is applied to the resulting series of residuals. Since the parameters are estimated from the historical period, it must be ensured that it is free from variance changes. To avoid missing a changepoint in the historical period, we first perform a retrospective version of the procedure to $X_1, \ldots, X_m$. This procedure is similar to the method in Aue et al. (2009a) or a multivariate variant of Wied et al. (2012a) or Wied et al. (2012b) with detector

$$Q_k = \frac{k}{\sqrt{m}} \hat{D}_p^{-\frac{1}{2}} \left(\left[\hat{\sigma}^2 \right]_1^k - \left[\hat{\sigma}^2 \right]_1^m \right),$$

where

$$\lim_{m \to \infty} \mathsf{P} \left(\sup_{b \in [0,1]} \frac{\left\| Q_{\lfloor m \cdot b \rfloor + 2} \right\|_2}{w(b)} > c \right) = \mathsf{P} \left(\sup_{b \in [0,1]} \frac{\left\| B_p(b) \right\|_2}{w(b)} > c \right)$$

and $B_p(\cdot)$ is a p dimensional Brownian bridge whose component processes are p independent Brownian bridges. According to Aue et al. (2009a) the location of a detected changepoint can be estimated by

$$\hat{k}_r := \underset{2 \leq k \leq m}{sup}\, Q_k.$$

If this retrospective inspection indicates the presence of a changepoint, all observations before the estimated changepoint $\hat{k}_r$ are cut off. Then, a new historical data set consisting of $X_{\hat{k}_r+1}, \ldots, X_{\hat{k}_r+m}$ is created and tested for another changepoint using the retrospective procedure. If no changepoint is detected in the historical period, the model parameters are estimated

from the historical data and used to gain a sequence of residual vectors to which the monitoring procedure can be applied. If a changepoint is detected in the monitoring period and located at $\hat{k}$, the following m data points form the new historical period. This practice is reiterated until no more changepoint can be found in the monitoring period or until there are less than m data points left after the last detected changepoint. Unfortunately, performing several retrospective tests on only partially exchanged observations leads to an increased probability to commit a type I error. However, we neglect this problem as we need a changepoint-free historical period to be able to perform the procedure properly. Choosing $\gamma = \{0, 0.25\}$ and B as the number of remaining data points after the historical period divided by m, we obtain changepoints that are presented in Table 2.

$m = 500$							
$\gamma = 0$				$\gamma = 0.25$			
τ_m	$\hat{k}$	τ_m	$\hat{k}$	τ_m	$\hat{k}$	τ_m	$\hat{k}$
1983-03-24	1981-05-12		2002-06-13	1982-06-24	1981-05-13		2002-06-13
1984-10-22	1984-08-13		2003-04-15	1984-10-22	1984-08-13		2003-04-15
	1985-09-16		2003-11-24		1985-09-16		2003-11-24
1992-03-12	1988-07-22	2007-05-22	2006-06-14	1992-02-10	1988-07-22	2006-06-15	2006-03-03
1991-03-07	1990-10-24	2008-09-19	2008-09-10	1990-10-09	1990-09-12	2008-03-18	2008-03-14
1996-11-08	1996-02-12		2009-04-02	1997-09-23	1996-06-11	2013-01-08	2011-08-31
1999-12-24	1999-03-15	2012-07-23	2011-12-05	1998-09-30	1998-07-13	2014-05-30	2013-11-18
2001-10-15	2001-08-07	2014-11-20	2014-09-17	2003-02-26	2001-08-07		
$m = 1.000$							
1985-04-12	1984-08-13		2004-05-19	1984-10-30	1984-08-13	2003-12-22	2003-04-11
1992-03-26	1991-01-17	2008-11-25	2008-09-05	1991-12-24	1989-10-17		2004-05-19
1997-01-27	1996-06-11		2009-05-19		1991-03-20	2008-10-10	2008-09-05
	1998-07-13		2011-08-08	1996-12-04	1996-06-11		2009-05-19
2004-06-30	2003-04-16				1998-07-13		2011-08-08

Table 2: First hitting times and estimated changepoint locations when applying the monitoring procedure to asset returns of *Allianz*, *Bayer*, *Deutsche Bank*, *RWE* and *Siemens*.

Along with the log returns of the *Allianz* and *Siemens* assets, Figure 12 illustrates the changepoints that are detected using $\gamma = 0$ and $m = 1.000$. The time series are divided effectively in parts of higher or lower volatility by the procedure. The remaining time series show a similar behavior and will not be illustrated here for the sake of clarity. The reported changepoints are used to split the time series in parts of constantly higher and lower variance. The sample standard deviations between two succeeding changepoints are presented in Table 3 and illustrated in Figure 13 for the *Allianz* and the *Siemens* asset.

The results illustrated in Figures 12 and 13 can be associated with distinctive events in the last

25 years. The late eighties were influenced strongly by the stock market crash and the Chernobyl catastrophe. The latter one is of interest since the asset of RWE, an energy generating company that relies on nuclear power since the seventies, is included in our sample. By the end

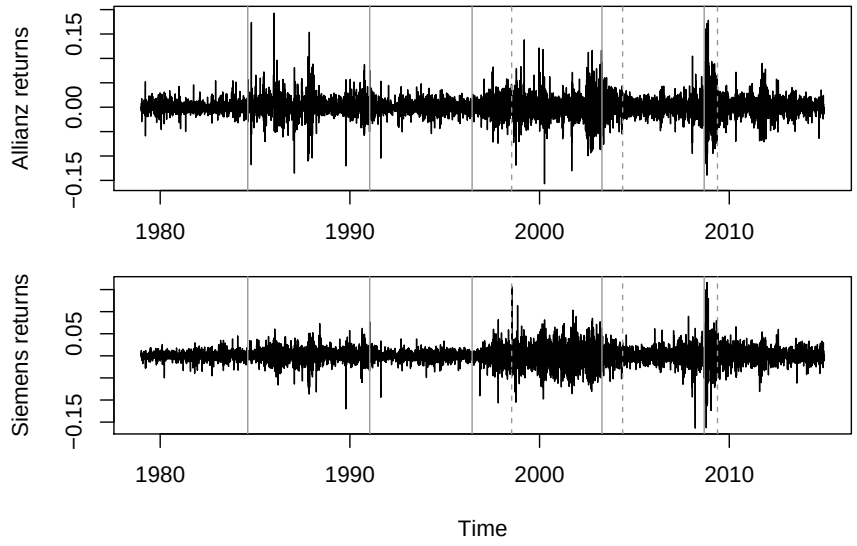


Figure 12: Returns of the *Allianz* and *Siemens* assets with detected changepoints ($\gamma = 0$ and $m = 1.000$). —— indicates that a changepoint was detected in the monitoring period; - - - indicates that a changepoint was detected in the historical period.

of the nineties the volatilities increased in the course of the financial crises in Southeast Asia and Russia, a trend that was reinforced around the turn of the millennium by the bursting of the dotcom bubble and the beginning of the Iraq war. The following years of sinking volatility

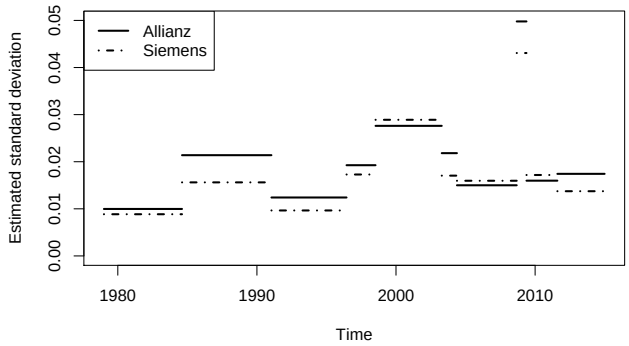


Figure 13: Sample standard deviations of the returns of the *Allianz* and *Siemens* assets between succeeding detected changepoints.

were interrupted by the Lehman bankruptcy and the following finance and debt crisis. Also, especially the asset of *RWE* was strongly influenced by the consequences of the nuclear incident in Fukushima in 2011.

Estimation period	Allianz	Bayer	Deutsche Bank	RWE	Siemens
1979-01-02 to 1984-08-12	0.0100	0.0103	0.0096	0.0084	0.0089
1984-08-13 to 1991-01-16	0.0214	0.0154	0.0167	0.0167	0.0156
1991-01-17 to 1996-06-10	0.0124	0.0116	0.0105	0.0102	0.0097
1996-06-11 to 2003-03-20	0.0193	0.0187	0.0173	0.0179	0.0173
1998-07-13 to 2003-04-15	0.0276	0.0255	0.0261	0.0218	0.0289
2003-04-16 to 2004-05-18	0.0218	0.0199	0.0171	0.0155	0.0170
2004-05-19 to 2008-09-04	0.0150	0.0148	0.0146	0.0130	0.0160
2008-09-04 to 2009-05-18	0.0498	0.0311	0.0627	0.0325	0.0431
2009-05-19 to 2011-08-07	0.0160	0.0157	0.0213	0.0133	0.0172
2011-08-08 to 2014-12-31	0.0174	0.0166	0.0236	0.0197	0.0137

Table 3: Sample standard deviations calculated from the time periods between detected changepoints

5. Conclusion

We propose a multivariate monitoring procedure to detect changes in the vector of variances of a sequence of random vectors and analyzed its size and power properties. An application to a group of asset returns reported plausible changepoints that could be associated to past events that actually showed strong influence on the stock market.

In the paper, we refrain from monitoring the whole covariance matrix as proposed by Aue et al. (2009b) in the retrospective case and only focus on the variances instead. From a practitioner's point of view an application of the proposed procedure extended to the covariances to time series of higher dimension is problematic. Even for a moderate number of observation units, D_p is of unpropitious high dimension. The matrix has to be estimated and the quality of the estimate declines with p which shows strong influence on the performance of the procedure. To circumvent this problem, one should pursue different approaches, e.g., one could monitor the largest eigenvalue of covariance matrices. We leave this task for future research.

References

Aue, A., S. Hörmann, L. Horváth, M. Hušková, and J. Steinebach (2011): "Sequential Testing for the Stability of High Frequency Portfolio Betas," *Econometric Theory*, 28(4), 804–837.

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

Aue, A., L. Horváth, M. Hušková, and P. Kokoszka (2006): "Change-point Monitoring in Linear Models," *Econometrics Journal*, 9, 373–403.

Aue, A., L. Horváth, and M. Reimherr (2009b): "Delay Times of Sequential Procedures for Multiple Time Series Regression Models," *Journal of Econometrics*, 149(2), 174–190.

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.

Bissantz, N., D. Ziggel, and K. Bissantz (2011): "An Empirical Study of Correlation and Volatility Changes of Stock Indices and their Impact on Risk Figures," *Acta Universitatis Danubis. Oeconomica*, 7(4), 127–141.

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

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.

Davidson, J. (1994): *Stochastic Limit Theory*, Oxford University Press.

Ding, Z. and R. Engle (2001): "Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing," *Academia Economic Papers*, 29(2), 157–184.

Engle, R. and K. Kroner (1995): "Multivariate Simultaneaous Generalized Arch," *Econometric Theory*, 11, 122–150.

Engle, R. and J. Mezrich (1996): "GARCH For Groups," *Risk Magazine*, 9, 36–40.

Groen, J., G. Kapetanios, and S. Price (2013): "Multivariate Methods for Monitoring Structural Change," *Journal of Applied Econometrics*, 28(2), 250–274.

Hafner, C. (2003): "Fourth Moment Structure of Multivariate GARCH Models," *Journal of Financial Econometrics*, 1(1), 26–54.

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

Kulperger, R. and H. Yu (2005): "High Moment Partial Sum Processes of Residuals in GARCH Models and Their Applications," *Journal of Business and Economic Statistics*, 33(5), 2395–2422.

Pedersen, R. and A. Rahbek (2014): "Multivariate Variance Tarjeting in the BEKK-GARCH Model," *Econometrics Journal*, 17(1), 24–55.

Schwert, G. (2011): "Stock Volatility during the Recent Financial Crisis," *European Financial Management*, 17(5), 789–805.

Wied, D. (2014): "A Nonparametric Test for a Constant Correlation Matrix," *submitted for publication*.

Wied, D., M. Arnold, N. Bissantz, and D. Ziggel (2012a): "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.

Wied, D., W. Krämer, and H. Dehling (2012b): "Testing for a Change in Correlation at an Unknown Point in Time Using an Extended Functional Delta Method," *Econometric Theory*, 68(3), 570–589.

A. Appendix: Proofs

The proofs of Theorem 1-3 are along the lines of Wied and Galeano (2013).

Proof of Theorem 1

Let $D[d_1, d_2]$ be the space of càdlàg functions on the interval $[d_1, d_2]$ equipped with the supremum norm. Denote the time invariant vector of variances by $\sigma^2 = \left(\sigma_1^2, \ldots, \sigma_p^2\right)'$ and define $\{P_m(d),\ d \in [c, B]\}$ by

$$P_m(d) = \hat{D}^{-\frac{1}{2}} \frac{[m \cdot d] - [m \cdot c]}{\sqrt{m}} \left(\left[\hat{\sigma}^2\right]_{[m\cdot c]}^{[m\cdot d]} - \sigma^2 \right) = \hat{D}^{-\frac{1}{2}} \frac{[m \cdot d] - [m \cdot c]}{\sqrt{m}} \begin{pmatrix} \left[\hat{\sigma}_1^2\right]_{[m\cdot c]}^{[m\cdot d]} - \sigma_1^2 \\ \vdots \\ \left[\hat{\sigma}_p^2\right]_{[m\cdot c]}^{[m\cdot d]} - \sigma_p^2 \end{pmatrix}.$$

Then, it has to be shown for fixed $c \geq 0$, assuming $m \to \infty$ that $\{P_m(d),\ d \in [c, B]\}$ converges in distribution to $\{W_p(d) - W_p(c),\ d \in [c, B]\}$ on $D[c, B]$ with $W_p(\cdot)$ being a p dimensional Brownian Motion. This leads to

$$S_m(b) := \begin{pmatrix} \hat{D}^{-\frac{1}{2}} \frac{[m\cdot b]+2}{\sqrt{m}} \left(\left[\hat{\sigma}^2\right]_{m+1}^{m+[m\cdot b]+2} - \sigma^2 \right) \\ \hat{D}^{-\frac{1}{2}} \sqrt{m} \left(\left[\hat{\sigma}^2\right]_1^m - \sigma^2 \right) \end{pmatrix} \Rightarrow_d \begin{pmatrix} W_p(b+1) - W_p(1) \\ W_p(1) \end{pmatrix}, \text{ for } b \in [0, B].$$

Consequently,

$$V_{[m\cdot b]+2} = \hat{D}^{-\frac{1}{2}} \frac{[m \cdot b] + 2}{\sqrt{m}} \left(\left[\hat{\sigma}^2\right]_{m+1}^{m+[m\cdot b]+2} - \sigma^2 \right) - \hat{D}^{-\frac{1}{2}} \frac{[m \cdot b] + 2}{\sqrt{m}} \left(\left[\hat{\sigma}^2\right]_1^m - \sigma^2 \right)$$

converges to the process $\{W_p(b+1) - (b+1)W_p(1),\ b \in [0, B]\}$. Applying the continuous mapping theorem and calculating the covariance structure of the limit process proves the result. ■

Proof of Theorem 2

The proof uses the same arguments as the one of Theorem 1 and is mainly based on the fact that for fixed $c \geq 0$, and $m \to \infty$ the process $\{P_m(d),\ d \in [c, B]\}$ converges in distribution to

$$\left\{ W_p(d) - W_p(c) + H \begin{pmatrix} \int_c^d g_1(z)dz \\ \vdots \\ \int_c^d g_p(z)dz \end{pmatrix},\ d \in [c, B] \right\}$$

on $D[c, B]$. The constant H is, up to a constant, the limit of $\hat{D}$ under the null hypothesis, see the proof of Theorem 2 in Wied et al. (2012a). This result is a generalization of arguments used in Theorem 2 in Wied et al. (2012a), executed along the lines of to the proof of Theorem 1. ■

Proof of Theorem 3

Assume w.l.o.g. $g_1(\cdot) := Mh(\cdot)$. Then, the detector converges in the following way:

$$\sup_{b\in[0,B]} \left\| V_{\lfloor mb \rfloor + 2} \right\|_2 \overset{d}{\Rightarrow} \sup_{b\in[0,B]} \left\| \begin{pmatrix} G_1(b) \\ \vdots \\ G_p(b) \end{pmatrix} + D^{-\frac{1}{2}} \begin{pmatrix} M \cdot \int_1^{b+1} h(u)du \\ \int_1^{b+1} g_2(u)du \\ \vdots \\ \int_1^{b+1} g_p(u)du \end{pmatrix} \right\|_2 . \tag{9}$$

Denote $D^{-\frac{1}{2}} := \left(d_{ij}\right)_{i,j=1,\ldots,p}$ and define constants $c_1(b) := \int_1^{b+1} h(u)du$ and $c_i(b) := \int_1^{b+1} g_i(u)du$, $i = 2, \ldots, p$. Thus, (9) has asymptotically the same distribution as

$$\sup_{b\in[0,B]} \sqrt{\sum_{j=1}^{p} \left(G_j(b) + Md_{j1}c_1(b) + \sum_{i=2}^{p} d_{ji}c_i(b) \right)^2} . \tag{10}$$

Since D_p is positive definite there exists $j \in \{1, \ldots, p\}$ with $d_{j1} \neq 0$. Assuming $M \to \infty$, we have $\left| G_j(b) + Md_{j1}c_1(b) \right| \to \infty$. Thus, Jensen's inequality implies that for all $b \in [0, B]$ the square root of the sum in (10) tends to ∞. This implies that (10) will exceed every quantile of the asymptotic null distribution for $M \to \infty$. ■

Proof of Lemma 1

Hafner (2003) provides conditions to establish the existence of the matrix of fourth moments and cross moments of a multivariate $GARCH(1,1)$ model in *vech* representation:

$$vech\left(H_t\right) = C_0 + A_1 vech\left(X_{t-1}X_{t-1}'\right) + B_1 vech\left(H_{t-1}\right), \tag{11}$$

where C_0 is a d dimensional parameter vector and A_1 and B_1 are parameter matrices of dimension $d \times d$. The closely related *vec* representation is given as

$$vec\left(H_t\right) = C_0^* + A_1^* vec\left(X_{t-1}X_{t-1}'\right) + B_1^* vec\left(H_{t-1}\right) \tag{12}$$

and contains several redundant equations that lead to inflated parameter matrices A_1^* and B_1^* of dimension $p^2 \times p^2$ and a parameter vector C_0^* of dimension p^2. Following Engle and Kroner (1995), the model in (7) can be given in *vec* representation by choosing

$$C_0^* = (1-\alpha-\beta)\left[H^{\frac{1}{2}} \otimes H^{\frac{1}{2}}\right] vec\left(I_p\right), \quad A_1^* = \alpha I_{p^2} \quad \text{and} \quad B_1^* = \beta I_{p^2}. \tag{13}$$

in (12). Thus, (7) can be given in *vech* representation by transforming it first to its *vec* and then to its *vech* representation. Substituting (13) in (12) and multiplying D_p and L_p gives

$$C_0 = (1-\alpha-\beta)\, L_p \cdot H^{\frac{1}{2}} \otimes H^{\frac{1}{2}} \cdot vec\left(I_p\right), \quad A_1 = \alpha L_p D_p \quad \text{and} \quad B_1 = \beta L_p D_p \tag{14}$$

in model (11). Using (14) and G_p from Hafner (2003) to construct Z allows to check the existence of Γ according to Hafner (2003). ■

B. Appendix: Tables

			$p = 2$			$p = 5$			$p = 10$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000	$m = 500$	1000	2000
D is known	0	0.5	0.0476	0.0431	0.0451	0.0493	0.0518	0.0535	0.0500	0.0502	0.0515
		1	0.0432	0.0465	0.0496	0.0485	0.0501	0.0500	0.0487	0.0488	0.0474
		2	0.0459	0.0435	0.0507	0.0507	0.0520	0.0536	0.0527	0.0533	0.0493
	0.25	0.5	0.0522	0.0489	0.0434	0.0550	0.0480	0.0518	0.0499	0.0496	0.0499
		1	0.0487	0.0473	0.0498	0.0470	0.0474	0.0457	0.0473	0.0485	0.0487
		2	0.0516	0.0462	0.0478	0.0528	0.0506	0.0539	0.0529	0.0522	0.0517
$\Sigma = I_p$	0	0.5	0.0659	0.0558	0.0574	0.0740	0.0641	0.0607	0.1041	0.0726	0.0652
		1	0.0608	0.0605	0.0528	0.0802	0.0659	0.0588	0.1080	0.0781	0.0650
		2	0.0689	0.0605	0.0598	0.0852	0.0657	0.0573	0.1016	0.0826	0.0628
	0.25	0.5	0.0639	0.0613	0.0545	0.0839	0.0677	0.0613	0.1034	0.0752	0.0633
		1	0.0674	0.0623	0.0573	0.0879	0.0668	0.0584	0.1144	0.0884	0.0755
		2	0.0702	0.0564	0.0539	0.0820	0.0747	0.0605	0.1132	0.0813	0.0666
$\Sigma = H_p$	0	0.5	0.0636	0.0588	0.0569	0.0788	0.0652	0.0607	0.1005	0.0741	0.0645
		1	0.0632	0.0611	0.0580	0.0845	0.0669	0.0607	0.1083	0.0814	0.0696
		2	0.0662	0.0588	0.0589	0.0844	0.0735	0.0644	0.1024	0.0787	0.0642
	0.25	0.5	0.0620	0.0672	0.0576	0.0844	0.0732	0.0646	0.1134	0.0818	0.0680
		1	0.0646	0.0570	0.0518	0.0900	0.0800	0.0652	0.1128	0.0850	0.0760
		2	0.0696	0.0640	0.0508	0.0926	0.0664	0.0688	0.1190	0.0946	0.0690

Table 4: Size when monitoring a sequence of realizations of i.i.d. $N\left(0, \Sigma_p\right)$ distributed random vectors.

			$p = 2$			$p = 5$			$p = 10$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000	$m = 500$	1000	2000
D is known	0	0.5	0.0492	0.0509	0.0520	0.0713	0.0614	0.0610	0.0777	0.0719	0.0695
		1	0.0520	0.0512	0.0523	0.0633	0.0634	0.0641	0.0767	0.0696	0.0690
		2	0.0533	0.0526	0.0555	0.0634	0.0631	0.0639	0.0746	0.0771	0.0724
	0.25	0.5	0.0631	0.0590	0.0582	0.0831	0.0754	0.0657	0.1073	0.0911	0.0907
		1	0.0636	0.0601	0.0586	0.0810	0.0762	0.0706	0.0945	0.0912	0.0813
		2	0.0625	0.0621	0.0528	0.0845	0.0749	0.0706	0.1033	0.0914	0.0872
$\Sigma = I_p$	0	0.5	0.0914	0.0796	0.0711	0.1356	0.1066	0.0830	0.1943	0.1430	0.1053
		1	0.0942	0.0821	0.0664	0.1405	0.1045	0.0832	0.2037	0.1386	0.1069
		2	0.1027	0.0800	0.0701	0.1475	0.1085	0.0836	0.2013	0.1384	0.0961
	0.25	0.5	0.1095	0.0906	0.0775	0.1632	0.1239	0.0993	0.2329	0.1635	0.1169
		1	0.1199	0.0906	0.0815	0.1713	0.1279	0.1020	0.2487	0.1810	0.1324
		2	0.1079	0.0967	0.0738	0.1741	0.1259	0.0961	0.2490	0.1670	0.1210
$\Sigma = H_p$	0	0.5	0.0966	0.0818	0.0717	0.1382	0.1067	0.0900	0.2008	0.1382	0.1006
		1	0.1003	0.0792	0.0686	0.1413	0.1101	0.0848	0.2107	0.1459	0.1057
		2	0.0978	0.0791	0.0745	0.1413	0.1112	0.0860	0.1962	0.1405	0.1013
	0.25	0.5	0.1140	0.0918	0.0836	0.1622	0.1212	0.0924	0.2472	0.1706	0.1162
		1	0.1138	0.0878	0.0712	0.1850	0.1316	0.1026	0.2522	0.1662	0.1224
		2	0.1174	0.0918	0.0718	0.1796	0.1334	0.1036	0.2702	0.1832	0.1214

Table 5: Size when monitoring a sequence of realizations of i.i.d. $t_\nu\left(0, \frac{\nu-2}{\nu}\Sigma_p\right)$ distributed random vectors with ν=8 degrees of freedom.

			$k^* = 0.05$			$k^* = 0.5$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000
$p = 2$	0	0.5	0.7775 (0.6665)	0.9597 (0.8941)	0.9995 (0.9938)	0.3398(0.2276)	0.5127 (0.3389)	0.7670 (0.5243)
		1	0.8923 (0.7960)	0.9939 (0.9742)	1.0000 (0.9998)	0.4174 (0.2798)	0.6592 (0.4258)	0.9101 (0.6819)
		2	0.9549 (0.8972)	0.9992 (0.9929)	1.0000 (1.0000)	0.5121 (0.3336)	0.7742 (0.5253)	0.9703 (0.8163)
	0.25	0.5	0.7664 (0.6482)	0.9552 (0.8818)	0.9995 (0.9929)	0.3137 (0.2073)	0.4846 (0.3123)	0.7483 (0.4987)
		1	0.8876 (0.7848)	0.9917 (0.9644)	1.0000 (0.9997)	0.3930 (0.2599)	0.6095 (0.3750)	0.8885 (0.6337)
		2	0.9432 (0.8689)	0.9987 (0.9901)	1.0000 (1.0000)	0.4585 (0.2890)	0.7364 (0.4771)	0.9604 (0.7841)
$p = 5$	0	0.5	0.9568 (0.8929)	0.9987 (0.9911)	1.0000 (1.0000)	0.5393 (0.3564)	0.7563 (0.5098)	0.9642 (0.7866)
		1	0.9926 (0.9671)	1.0000 (0.9997)	1.0000 (1.0000)	0.6677 (0.4431)	0.8977 (0.6687)	0.9967 (0.9264)
		2	0.9991 (0.9934)	1.0000 (1.0000)	1.0000 (1.0000)	0.7540 (0.5245)	0.9590 (0.7791)	0.9994(0.9782)
	0.25	0.5	0.9553 (0.8852)	0.9987 (0.9900)	1.000 (0.9999)	0.5155 (0.3344)	0.7431 (0.4934)	0.9546 (0.7501)
		1	0.9919 (0.9630)	1.0000 (0.9997)	1.0000 (1.0000)	0.6409 (0.4137)	0.8752 (0.6294)	0.9951 (0.9044)
		2	0.9990 (0.9918)	1.0000 (1.0000)	1.0000 (1.0000)	0.7286 (0.4893)	0.9508 (0.7509)	0.9994 (0.9725)
$p = 10$	0	0.5	0.9998 (0.9977)	1.0000 (1.0000)	1.0000 (1.0000)	0.8464 (0.6361)	0.9821 (0.8518)	1.0000 (0.9870)
		1	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)	0.9491 (0.7749)	0.9989 (0.9529)	1.0000 (0.9994)
		2	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)	0.9820 (0.8546)	1.0000 (0.9885)	1.0000 (1.0000)
	0.25	0.5	0.9997 (0.9969)	1.0000 (1.0000)	1.0000 (1.0000)	0.8270 (0.6084)	0.9736 (0.8174)	1.000 (0.9840)
		1	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)	0.9349 (0.7433)	0.9986 (0.9434)	1.0000 (0.9994)
		2	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)	0.9774 (0.8499)	0.9999 (0.9868)	1.0000 (1.0000)

Table 6: Power when monitoring a sequence of realizations of i.i.d. $N\left(0, \Sigma_p\right)$ distributed random vectors when all of the variances are affected by a change.

			$k^* = 0.05$			$k^* = 0.5$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000
$p = 2$	0	0.5	0.6063 (0.5079)	0.7991 (0.7059)	0.9636 (0.9083)	0.2868 (0.2222)	0.3797 (0.2635)	0.5689 (0.3786)
		1	0.7377 (0.6359)	0.9112 (0.8363)	0.9921 (0.9746)	0.3566 (0.2595)	0.4842 (0.3329)	0.7015 (0.4742)
		2	0.8173 (0.7291)	0.9568 (0.9032)	0.9984 (0.9933)	0.4045 (0.2834)	0.5500 (0.3715)	0.8021 (0.5726)
	0.25	0.5	0.6154 (0.5066)	0.7883 (0.6833)	0.9547 (0.8912)	0.2821 (0.2216)	0.3584 (0.2454)	0.5247 (0.3401)
		1	0.7339 (0.6217)	0.8980 (0.8100)	0.9902 (0.9670)	0.3479 (0.2564)	0.4371 (0.2967)	0.6646 (0.4337)
		2	0.8026 (0.7045)	0.9513 (0.8896)	0.9980 (0.9916)	0.3775 (0.2677)	0.5176 (0.3450)	0.7730 (0.5302)
$p = 5$	0	0.5	0.7924 (0.7084)	0.9349 (0.8636)	0.9971 (0.9836)	0.4229 (0.3303)	0.5270 (0.3801)	0.7286 (0.5106)
		1	0.8999 (0.8297)	0.9860 (0.9547)	0.9996 (0.9981)	0.5232 (0.3931)	0.6607 (0.4719)	0.8745 (0.6661)
		2	0.9515 (0.8975)	0.9965 (0.9827)	1.0000 (0.9996)	0.5944 (0.4377)	0.7575 (0.5446)	0.9353 (0.7727)
	0.25	0.5	0.8056 (0.7119)	0.9358 (0.8603)	0.9962 (0.9803)	0.4245 (0.3355)	0.5147 (0.3700)	0.7008 (0.4799)
		1	0.8995 (0.8250)	0.9841 (0.9503)	0.9996 (0.9977)	0.5094 (0.3898)	0.6412 (0.4536)	0.8501 (0.6338)
		2	0.9493 (0.8881)	0.9963 (0.9807)	1.0000 (0.9995)	0.5757 (0.4201)	0.7412 (0.5294)	0.9227 (0.7483)
$p = 10$	0	0.5	0.9326 (0.8790)	0.9889 (0.9687)	0.9999 (0.9990)	0.6205 (0.4834)	0.7074 (0.5246)	0.8861 (0.6824)
		1	0.9763 (0.9506)	0.9989 (0.9939)	1.0000 (1.000)	0.7016 (0.5559)	0.8513 (0.6592)	0.9639 (0.8228)
		2	0.9929(0.9806)	0.9999 (0.9990)	1.0000 (1.0000)	0.7831 (0.6117)	0.9144 (0.7545)	0.9899 (0.9103)
	0.25	0.5	0.9355 (0.8776)	0.9879 (0.9651)	0.9999 (0.9988)	0.6117 (0.4798)	0.6832 (0.4983)	0.8769 (0.6645)
		1	0.9781 (0.9510)	0.9990 (0.9919)	1.0000 (1.0000)	0.6956 (0.5564)	0.8293 (0.6292)	0.9581 (0.8033)
		2	0.9951 (0.9771)	0.9999 (0.9986)	1.0000 (1.0000)	0.7683 (0.6135)	0.8965 (0.7175)	0.9907 (0.8983)

Table 7: Power when monitoring a sequence of realizations of i.i.d. $t_\nu\left(0, \frac{\nu-2}{\nu}\Sigma_p\right)$ distributed random vectors with ν=8 degrees of freedom when all of the variances are affected by a change.

			$k^* = 0.05$			$k^* = 0.5$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000
$p = 2$	0	0.5	0.9695 (0.9298)	0.9991 (0.9975)	1.0000 (1.0000)	0.6006 (0.4040)	0.8514 (0.6364)	0.9853 (0.8923)
		1	0.9972 (0.9843)	1.0000 (1.0000)	1.0000 (1.0000)	0.7483 (0.5150)	0.9573 (0.7892)	0.9995 (0.9784)
		2	0.9997 (0.9982)	1.0000 (1.0000)	1.0000 (1.0000)	0.8618 (0.6383)	0.9894 (0.8994)	1.0000 (0.9970)
	0.25	0.5	0.9656 (0.9196)	0.9989 (0.9967)	1.0000 (1.0000)	0.5672 (0.3673)	0.8248 (0.5966)	0.9825 (0.8792)
		1	0.9967 (0.9809)	1.0000 (1.0000)	1.0000 (1.0000)	0.7167 (0.4821)	0.9520 (0.7731)	0.9994 (0.9712)
		2	0.9997 (0.9968)	1.0000 (1.0000)	1.0000 (1.0000)	0.8280 (0.5847)	0.9861 (0.8746)	1.0000 (0.9953)
$p = 5$	0	0.5	0.9310 (0.8578)	0.9972 (0.9877)	1.0000 (0.9996)	0.4833 (0.3255)	0.7406 (0.4878)	0.9518 (0.7774)
		1	0.9883 (0.9588)	1.0000 (0.9997)	1.0000 (1.0000)	0.6319 (0.4153)	0.8966 (0.6611)	0.9950 (0.9283)
		2	0.9995 (0.9908)	1.0000 (1.0000)	1.0000 (1.0000)	0.7430 (0.5007)	0.9621 (0.7900)	0.9995 (0.9792)
	0.25	0.5	0.9244 (0.8443)	0.9968 (0.9853)	1.0000 (0.9996)	0.4566 (0.2945)	0.7144 (0.4595)	0.9444 (0.7568)
		1	0.9874 (0.9533)	1.0000(0.9996)	1.0000 (1.0000)	0.6135 (0.3952)	0.8707 (0.6152)	0.9936 (0.9140)
		2	0.9988 (0.9890)	1.0000 (1.0000)	1.0000 (1.0000)	0.7191 (0.4722)	0.9530 (0.7582)	0.9994 (0.9720)
$p = 10$	0	0.5	0.8507 (0.7481)	0.9878 (0.9524)	1.0000 (0.9996)	0.3891 (0.2557)	0.5842 (0.3674)	0.8753 (0.6334)
		1	0.9600 (0.8996)	0.9993 (0.9941)	1.0000 (1.0000)	0.5139 (0.3417)	0.7631 (0.5013)	0.9753 (0.8131)
		2	0.9896 (0.9530)	1.0000 (0.9996)	1.0000 (1.0000)	0.6127 (0.4093)	0.8885 (0.6395)	0.9963 (0.9210)
	0.25	0.5	0.8541 (0.7444)	0.9879 (0.9504)	1.0000 (0.9996)	0.3814 (0.2543)	0.5727 (0.3550)	0.8574 (0.5955)
		1	0.9575 (0.8892)	0.9993 (0.9934)	1.0000 (1.0000)	0.4898 (0.3244)	0.7453 (0.4769)	0.9702 (0.7898)
		2	0.9886 (0.9477)	1.0000 (0.9995)	1.0000 (1.0000)	0.5843 (0.3821)	0.8730 (0.6110)	0.9950 (0.9082)

Table 8: Power when monitoring a sequence of realizations of i.i.d. $N\left(0, \Sigma_p\right)$ distributed random vectors when only one of the variances is affected by a change.

			$k^* = 0.05$			$k^* = 0.5$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000
$p = 2$	0	0.5	0.8771 (0.7992)	0.9861 (0.9606)	0.9994 (0.9989)	0.4695 (0.3303)	0.6917 (0.4828)	0.9132 (0.7289)
		1	0.9597 (0.9155)	0.9976 (0.9921)	0.9998 (0.9997)	0.5990 (0.4133)	0.8311 (0.6079)	0.9802 (0.8709)
		2	0.9873 (0.9641)	0.9998 (0.9976)	0.9999 (0.9999)	0.7014 (0.4957)	0.9167 (0.7339)	0.9965 (0.9512)
	0.25	0.5	0.8717 (0.7878)	0.9836 (0.9544)	0.9993 (0.9986)	0.4522 (0.3206)	0.6599 (0.4538)	0.8969 (0.6931)
		1	0.9564 (0.9064)	0.9972 (0.9904)	0.9998 (0.9997)	0.5784 (0.3974)	0.8124 (0.5769)	0.9731 (0.8427)
		2	0.9848 (0.9551)	0.9995 (0.9972)	0.9999 (0.9998)	0.6687 (0.4624)	0.9015 (0.7018)	0.9945 (0.9395)
$p = 5$	0	0.5	0.8261 (0.7345)	0.9710 (0.9238)	0.9995 (0.9968)	0.4155 (0.3161)	0.5889 (0.3971)	0.8428 (0.6091)
		1	0.9327 (0.8623)	0.9955 (0.9846)	0.9999 (0.9997)	0.5220 (0.3706)	0.7521 (0.5402)	0.9508 (0.7864)
		2	0.9748 (0.9352)	0.9993 (0.9958)	0.9999 (0.9998)	0.6300 (0.4340)	0.8600 (0.6618)	0.9851 (0.8855)
	0.25	0.5	0.8295 (0.7370)	0.9686 (0.9165)	0.9996 (0.9961)	0.4205 (0.3272)	0.5665 (0.3805)	0.8277 (0.5871)
		1	0.9307 (0.8552)	0.9947 (0.9801)	0.9999 (0.9997)	0.5102 (0.3687)	0.7196 (0.5018)	0.9423 (0.7646)
		2	0.9728 (0.9267)	0.9993 (0.9950)	0.9999 (0.9998)	0.6135 (0.4242)	0.8380 (0.6273)	0.9832 (0.8746)
$p = 10$	0	0.5	0.7508 (0.6610)	0.9238 (0.8528)	0.9961 (0.9843)	0.3973 (0.3184)	0.4922 (0.3394)	0.7157 (0.4896)
		1	0.875 (0.8071)	0.9815 (0.9522)	0.9998 (0.9989)	0.4815 (0.3773)	0.6395 (0.4323)	0.8718 (0.6423)
		2	0.9469 (0.8809)	0.9958 (0.9849)	1.0000 (0.9994)	0.5613 (0.4238)	0.7520 (0.5393)	0.9498 (0.7700)
	0.25	0.5	0.7605 (0.6705)	0.9207 (0.8456)	0.9954 (0.9806)	0.4150 (0.3410)	0.4838 (0.3398)	0.6851 (0.4607)
		1	0.8764 (0.8014)	0.9819 (0.9512)	0.9998 (0.9983)	0.4830 (0.3883)	0.6348 (0.4341)	0.8552 (0.6176)
		2	0.9462 (0.8787)	0.9950 (0.9822)	1.0000 (0.9993)	0.5660 (0.4325)	0.7358 (0.5250)	0.9416 (0.7451)

Table 9: Power when monitoring a sequence of realizations of i.i.d. $t_\nu\left(0, \frac{\nu-2}{\nu}\Sigma_p\right)$ distributed random vectors with ν=8 degrees of freedom when only one of the variances is affected by a change.

			$p = 2$			$p = 5$			$p = 10$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000	$m = 500$	1000	2000
multivariate procedure	0	0.5	0.0733	0.0694	0.0612	0.0984	0.0778	0.0662	0.1325	0.1031	0.0834
		1	0.0788	0.0706	0.0602	0.1056	0.0789	0.0739	0.1408	0.1051	0.0876
		2	0.0762	0.0716	0.0692	0.1058	0.0861	0.0723	0.1326	0.0973	0.0785
	0.25	0.5	0.0751	0.0704	0.0578	0.1103	0.0822	0.0734	0.1362	0.0987	0.0798
		1	0.0879	0.0693	0.0666	0.1147	0.0888	0.0787	0.1540	0.1128	0.0925
		2	0.0836	0.0694	0.0596	0.1149	0.0810	0.0731	0.1477	0.1113	0.0908
univariate procedures	0	0.5	0.0645	0.0596	0.0524	0.0912	0.0789	0.0683	0.1031	0.0812	0.0682
		1	0.0705	0.0634	0.0607	0.0905	0.0803	0.0694	0.1101	0.0897	0.0733
		2	0.0653	0.0602	0.0504	0.0961	0.0759	0.0738	0.0953	0.0732	0.0639
	0.25	0.5	0.0706	0.0624	0.0555	0.1065	0.0884	0.0757	0.1226	0.0884	0.0699
		1	0.0753	0.0701	0.0567	0.1001	0.0767	0.0687	0.1306	0.0992	0.0852
		2	0.0844	0.0743	0.0644	0.1048	0.0860	0.0689	0.1428	0.1068	0.0870

Table 10: Size when monitoring scalar $BEKK$ time series with parameters $\alpha = 0.03$, $\beta = 0.45$ and $N\left(0, \Sigma_p\right)$ distributed innovations.

			$p = 2$			$p = 5$			$p = 10$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000	$m = 500$	1000	2000
multivariate procedure	0	0.5	0.1091	0.0877	0.0798	0.1533	0.1207	0.0987	0.2430	0.1674	0.1288
		1	0.1045	0.0871	0.0754	0.1703	0.1294	0.0978	0.2492	0.1816	0.1319
		2	0.1133	0.0938	0.0747	0.1684	0.1323	0.1001	0.2434	0.1663	0.1194
	0.25	0.5	0.1278	0.1005	0.0831	0.2012	0.1462	0.1173	0.2771	0.1974	0.1417
		1	0.1336	0.1082	0.0894	0.2101	0.1567	0.1091	0.3112	0.2143	0.1568
		2	0.1361	0.1018	0.0862	0.2017	0.1540	0.1160	0.3068	0.2131	0.1525
univariate procedures	0	0.5	0.0910	0.0713	0.0663	0.1443	0.1163	0.0924	0.1825	0.1369	0.1060
		1	0.1004	0.0812	0.0700	0.1448	0.1144	0.0896	0.2025	0.1468	0.1158
		2	0.0988	0.0853	0.0661	0.1508	0.1217	0.0895	0.1941	0.1369	0.1079
	0.25	0.5	0.1069	0.0919	0.0785	0.1761	0.1475	0.1107	0.2785	0.2102	0.1589
		1	0.1199	0.1022	0.0825	0.1769	0.1392	0.1046	0.2782	0.2147	0.1570
		2	0.1237	0.1015	0.0842	0.1932	0.1481	0.1157	0.2714	0.1964	0.1417

Table 11: Size when monitoring scalar $BEKK$ time series with parameters $\alpha = 0.03$, $\beta = 0.45$ and $t_\nu\left(0, \frac{\nu-2}{\nu}\Sigma_p\right)$ distributed innovations with ν=8 degrees of freedom.

			$k^* = 0.05$			$k^* = 0.5$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000
$p = 2$	0	0.5	0.7516 (0.6546)	0.9384 (0.8732)	0.9982 (0.9900)	0.3388 (0.2451)	0.4866 (0.3405)	0.7274 (0.5131)
		1	0.8748 (0.7820)	0.9908 (0.9599)	1.0000 (0.9990)	0.4066 (0.2876)	0.6120 (0.4159)	0.8716 (0.6522)
		2	0.9390 (0.8707)	0.9978 (0.9879)	1.0000 (1.0000)	0.4944 (0.3413)	0.7308 (0.5197)	0.9436 (0.7794)
	0.25	0.5	0.7500 (0.6408)	0.9312 (0.8602)	0.9972 (0.9869)	0.3172 (0.2340)	0.4578 (0.3120)	0.6946 (0.4760)
		1	0.8708 (0.7608)	0.9888 (0.9517)	1.0000 (0.9988)	0.3828 (0.2639)	0.5762 (0.3816)	0.8522 (0.6150)
		2	0.9274 (0.8518)	0.9978 (0.9847)	1.0000 (1.0000)	0.4586 (0.3109)	0.6936 (0.4796)	0.9322 (0.7428)
$p = 5$	0	0.5	0.9416 (0.8757)	0.9974 (0.9891)	1.0000 (0.9999)	0.5108 (0.3782)	0.7194 (0.5140)	0.9412 (0.7645)
		1	0.9902 (0.9630)	1.0000 (0.9992)	1.0000 (1.0000)	0.6532 (0.4592)	0.8756 (0.6433)	0.9904 (0.9010)
		2	0.9978 (0.9878)	1.0000 (1.0000)	1.0000 (1.0000)	0.7626 (0.5333)	0.9464 (0.7625)	0.9990 (0.9622)
	0.25	0.5	0.9402 (0.8654)	0.9968 (0.9869)	1.0000 (0.9999)	0.4838 (0.3544)	0.6826 (0.4795)	0.9248 (0.7298)
		1	0.9900 (0.9586)	1.0000 (0.9991)	1.0000 (1.0000)	0.6250 (0.4360)	0.8610 (0.6146)	0.9884 (0.8840)
		2	0.9970 (0.9852)	1.0000 (1.0000)	1.0000 (1.0000)	0.7284 (0.5008)	0.9314 (0.7305)	0.9986 (0.9536)
$p = 10$	0	0.5	0.9996 (0.9964)	1.0000 (1.0000)	1.0000 (1.0000)	0.8292 (0.6333)	0.9688 (0.8231)	0.9998 (0.9795)
		1	0.9998 (0.9999)	1.0000 (1.0000)	1.0000 (1.0000)	0.9344 (0.7639)	0.9968 (0.9412)	1.0000 (0.9990)
		2	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)	0.9786 (0.8492)	0.9996 (0.9802)	1.0000 (1.0000)
	0.25	0.5	0.9996 (0.9957)	1.0000 (1.0000)	1.0000 (1.0000)	0.8106 (0.6086)	0.9624 (0.8032)	0.9998 (0.9743)
		1	0.9998 (0.9999)	1.0000 (1.0000)	1.0000 (1.0000)	0.9234 (0.7415)	0.9962 (0.9288)	1.0000 (0.9983)
		2	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)	0.9734 (0.8321)	0.9996 (0.9765)	1.0000 (1.0000)

Table 12: Power when monitoring scalar $BEKK$ time series and all of the variances increase ($N\left(0, \Sigma_p\right)$ distributed innovations).

			$k^* = 0.05$			$k^* = 0.5$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000
$p = 2$	0	0.5	0.5999 (0.5036)	0.7869 (0.6778)	0.9509 (0.8787)	0.3110 (0.2225)	0.3778 (0.2525)	0.5388 (0.3506)
		1	0.7255 (0.6175)	0.8954 (0.8044)	0.9885 (0.9580)	0.3634 (0.2602)	0.4692 (0.3140)	0.6691 (0.4451)
		2	0.8115 (0.7052)	0.9474 (0.8835)	0.9973 (0.9881)	0.4022 (0.2922)	0.5604 (0.3749)	0.7846 (0.5585)
	0.25	0.5	0.6115 (0.5168)	0.7820 (0.6812)	0.9462 (0.8763)	0.3078 (0.2418)	0.3609 (0.2598)	0.5063 (0.3347)
		1	0.7241 (0.6213)	0.8873 (0.7956)	0.9858 (0.9558)	0.3517 (0.2743)	0.4434 (0.3116)	0.6342 (0.4272)
		2	0.8035 (0.6948)	0.9412 (0.8755)	0.9966 (0.9847)	0.3815 (0.2835)	0.5256 (0.3575)	0.7514 (0.5266)
$p = 5$	0	0.5	0.7884 (0.7018)	0.9277 (0.8550)	0.9924 (0.9762)	0.4446 (0.3476)	0.5299 (0.3916)	0.7123 (0.5033)
		1	0.8906 (0.8171)	0.9796 (0.9466)	0.9998 (0.9971)	0.5303 (0.4188)	0.6547 (0.4707)	0.8440 (0.6412)
		2	0.9399 (0.8823)	0.9946 (0.9774)	1.0000 (0.9997)	0.5997 (0.4534)	0.7450 (0.5464)	0.9270 (0.7491)
	0.25	0.5	0.7994 (0.7103)	0.9285 (0.8517)	0.9915 (0.9716)	0.4467 (0.3645)	0.5117 (0.3797)	0.6918 (0.4745)
		1	0.8917 (0.8188)	0.9777 (0.9348)	0.9995 (0.9953)	0.5206 (0.4077)	0.6295 (0.4593)	0.8207 (0.6193)
		2	0.9402 (0.8838)	0.9942 (0.9751)	0.9999 (0.9992)	0.5874 (0.4555)	0.7219 (0.5378)	0.9138 (0.7240)
$p = 10$	0	0.5	0.9279 (0.8772)	0.9856 (0.9609)	0.9995 (0.9988)	0.6379 (0.5284)	0.7077 (0.5673)	0.8697 (0.6859)
		1	0.9784 (0.9403)	0.9979 (0.9928)	1.0000 (0.9998)	0.7284 (0.5768)	0.8357 (0.6494)	0.9590 (0.8109)
		2	0.9902 (0.9798)	0.9999 (0.9979)	1.0000 (1.0000)	0.7953 (0.6480)	0.9071 (0.7476)	0.9877 (0.9011)
	0.25	0.5	0.9328 (0.8755)	0.9852 (0.9549)	0.9996 (0.9977)	0.6318 (0.5190)	0.6924 (0.5357)	0.8549 (0.6618)
		1	0.9780 (0.9463)	0.9977 (0.9932)	1.0000 (1.0000)	0.7121 (0.5770)	0.8171 (0.6474)	0.9485 (0.7978)
		2	0.9903 (0.9730)	0.9999 (0.9981)	1.0000 (1.0000)	0.7855 (0.6399)	0.8948 (0.7114)	0.9852 (0.8835)

Table 13: Power when monitoring scalar $BEKK$ time series and all of the variances increase ($t_\nu\left(0, \frac{\nu-2}{\nu}\Sigma_p\right)$ distributed innovations with ν=8 degrees of freedom).

			$k^* = 0.05$			$k^* = 0.5$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000
$p = 2$	0	0.5	0.9582 (0.9063)	0.9986 (0.9950)	1.0000 (1.0000)	0.5642 (0.3937)	0.8276 (0.6026)	0.9776 (0.8623)
		1	0.9938 (0.9794)	1.0000 (0.9995)	1.0000 (1.0000)	0.7120 (0.5065)	0.9364 (0.7684)	0.9990 (0.9646)
		2	0.9992 (0.9954)	1.0000 (1.0000)	1.0000 (1.0000)	0.8358 (0.6141)	0.9808 (0.8814)	1.0000 (0.9938)
	0.25	0.5	0.9552 (0.8979)	0.9986 (0.9926)	1.0000 (1.0000)	0.5410 (0.3747)	0.8000 (0.5614)	0.9728 (0.8442)
		1	0.9918 (0.9721)	1.0000 (0.9996)	1.0000 (1.0000)	0.6862 (0.4712)	0.9248 (0.7320)	0.9990 (0.9583)
		2	0.9988 (0.9948)	1.0000 (1.0000)	1.0000 (1.0000)	0.8102 (0.5855)	0.9752 (0.8482)	1.0000 (0.9906)
$p = 5$	0	0.5	0.9120 (0.8341)	0.9954 (0.9784)	1.0000 (1.0000)	0.4640 (0.3264)	0.6984 (0.4779)	0.9378 (0.7516)
		1	0.9804 (0.9454)	0.9996 (0.9985)	1.0000 (1.0000)	0.6216 (0.4172)	0.8746 (0.6438)	0.9914 (0.9058)
		2	0.9962 (0.9862)	1.0000 (1.0000)	1.0000 (1.0000)	0.7292 (0.5058)	0.9454 (0.7637)	0.9990 (0.9696)
	0.25	0.5	0.9028 (0.8256)	0.9950 (0.9760)	1.0000 (0.9999)	0.4368 (0.3021)	0.6646 (0.4463)	0.9226 (0.7299)
		1	0.9782 (0.9364)	0.9996 (0.9985)	1.0000 (1.0000)	0.5964 (0.3930)	0.8588 (0.6079)	0.9894 (0.8873)
		2	0.9960 (0.9811)	1.0000 (1.0000)	1.0000 (1.0000)	0.7004 (0.4869)	0.9330 (0.7326)	0.9986 (0.9614)
$p = 10$	0	0.5	0.8280 (0.7390)	0.9804 (0.9396)	1.0000 (0.9985)	0.3942 (0.2859)	0.5632 (0.3695)	0.8462 (0.6011)
		1	0.9480 (0.8844)	0.9990 (0.9919)	1.0000 (1.0000)	0.5062 (0.3666)	0.7442 (0.5067)	0.9612 (0.7968)
		2	0.9878 (0.9499)	1.0000 (0.9995)	1.0000 (1.0000)	0.6328 (0.4210)	0.8622 (0.6228)	0.9936 (0.8966)
	0.25	0.5	0.8266 (0.7231)	0.9772 (0.9302)	0.9998 (0.9980)	0.3790 (0.2779)	0.5384 (0.3554)	0.8232 (0.5619)
		1	0.9464 (0.8729)	0.9988 (0.9882)	1.0000 (1.0000)	0.4914 (0.3485)	0.7208 (0.4888)	0.9528 (0.7668)
		2	0.9860 (0.9483)	1.0000 (0.9982)	1.0000 (1.0000)	0.6102 (0.3947)	0.8396 (0.5893)	0.9902 (0.8758)

Table 14: Power when monitoring scalar $BEKK$ time series and just one of the variances increases ($N\left(0, \Sigma_p\right)$ distributed innovations).

			k^* = 0.05			k^* = 0.5		
	γ	B	m = 500	1000	2000	m = 500	1000	2000
p = 2	0	0.5	0.8604 (0.7823)	0.9778 (0.9444)	0.9998 (0.9982)	0.4666 (0.3340)	0.6556 (0.4609)	0.8922 (0.7014)
		1	0.9487 (0.9004)	0.9965 (0.9895)	0.9999 (0.9997)	0.5838 (0.4263)	0.8107 (0.6036)	0.9716 (0.8481)
		2	0.9842 (0.9540)	0.9988 (0.9972)	1.0000 (1.0000)	0.6880 (0.4919)	0.8969 (0.7163)	0.9913 (0.9296)
	0.25	0.5	0.8574 (0.7725)	0.9756 (0.9365)	0.9997 (0.9977)	0.4532 (0.3186)	0.6304 (0.4463)	0.8750 (0.6724)
		1	0.9458 (0.8842)	0.9960 (0.9870)	0.9999 (0.9996)	0.5645 (0.4001)	0.7889 (0.5738)	0.9643 (0.8206)
		2	0.9818 (0.9450)	0.9988 (0.9970)	1.0000 (0.9998)	0.6656 (0.4827)	0.8810 (0.6907)	0.9892 (0.9206)
p = 5	0	0.5	0.8043 (0.7142)	0.9612 (0.9085)	0.9989 (0.9929)	0.4229 (0.3175)	0.5673 (0.3925)	0.8130 (0.5881)
		1	0.9135 (0.8439)	0.9933 (0.9759)	0.9996 (0.9997)	0.5342 (0.3904)	0.7229 (0.5086)	0.9314 (0.7587)
		2	0.9673 (0.9233)	0.9989 (0.9946)	1.0000 (1.0000)	0.6248 (0.4487)	0.8317 (0.6239)	0.9774 (0.8684)
	0.25	0.5	0.8046 (0.7168)	0.9599 (0.8983)	0.9988 (0.9940)	0.4274 (0.3304)	0.5554 (0.3863)	0.7947 (0.5621)
		1	0.9109 (0.8507)	0.9921 (0.9773)	0.9996 (0.9994)	0.5316 (0.4003)	0.7030 (0.5138)	0.9199 (0.7378)
		2	0.9656 (0.9207)	0.9986 (0.9925)	1.0000 (0.9999)	0.6135 (0.4460)	0.8176 (0.5954)	0.9744 (0.8456)
p = 10	0	0.5	0.7439 (0.6705)	0.9084 (0.8354)	0.9925 (0.9750)	0.4308 (0.3587)	0.4926 (0.3673)	0.7025 (0.4872)
		1	0.8722 (0.7969)	0.9784 (0.9419)	0.9996 (0.9984)	0.5129 (0.3983)	0.6390 (0.4642)	0.8607 (0.6375)
		2	0.9304 (0.8736)	0.9932 (0.9747)	1.0000 (0.9993)	0.5776 (0.4425)	0.7384 (0.5159)	0.9330 (0.7415)
	0.25	0.5	0.7505 (0.6736)	0.9033 (0.8235)	0.9915 (0.9700)	0.4482 (0.3761)	0.4862 (0.3648)	0.6792 (0.4611)
		1	0.8728 (0.7939)	0.9749 (0.9367)	0.9996 (0.9960)	0.5220 (0.4322)	0.6207 (0.4526)	0.8381 (0.6152)
		2	0.9326 (0.8691)	0.9924 (0.9741)	1.0000 (0.9993)	0.5859 (0.4655)	0.7270 (0.5185)	0.9238 (0.7241)

Table 15: Power when monitoring scalar $BEKK$ time series and just one of the variances increases ($t_\nu\left(0, \frac{\nu-2}{\nu}\Sigma_p\right)$ distributed innovations with ν=8 degrees of freedom).

			normal distribution						t distribution					
			k^* = 0.05			k^* = 0.5			k^* = 0.05			k^* = 0.5		
	γ	B	m = 500	1000	2000	m = 500	1000	2000	m = 500	1000	2000	m = 500	1000	2000
p = 2	0	0.5	0.7570	0.9416	0.9987	0.3344	0.4931	0.7526	0.5714	0.7736	0.9479	0.2615	0.3506	0.5175
		1	0.8772	0.9886	1.0000	0.4108	0.6381	0.8932	0.7161	0.8947	0.9892	0.3477	0.4684	0.6811
		2	0.9477	0.9978	1.0000	0.5164	0.7538	0.9583	0.8052	0.9508	0.9960	0.4051	0.5672	0.7915
	0.25	0.5	0.7495	0.9351	0.9984	0.3119	0.4591	0.7204	0.5846	0.7773	0.9474	0.2691	0.3455	0.5025
		1	0.8716	0.9870	1.0000	0.3797	0.5985	0.8686	0.7056	0.8849	0.9875	0.3338	0.4399	0.6480
		2	0.9404	0.9972	1.0000	0.4822	0.7214	0.9491	0.7843	0.9414	0.9952	0.3752	0.5244	0.7544
p = 5	0	0.5	0.9400	0.9966	1.0000	0.5042	0.7155	0.9250	0.7913	0.9293	0.9957	0.4323	0.5455	0.7252
		1	0.9850	0.9997	1.0000	0.6220	0.8525	0.9851	0.8905	0.9839	0.9992	0.5141	0.6556	0.8657
		2	0.9962	1.0000	1.0000	0.6988	0.9099	0.9960	0.9375	0.9932	1.0000	0.5767	0.7372	0.9157
	0.25	0.5	0.9385	0.9962	1.0000	0.4826	0.6860	0.9078	0.8023	0.9288	0.9952	0.4331	0.5283	0.7040
		1	0.9834	0.9995	1.0000	0.5850	0.8201	0.9799	0.8859	0.9820	0.9989	0.4984	0.6240	0.8391
		2	0.9955	1.0000	1.0000	0.6542	0.8836	0.9942	0.9334	0.9925	0.9999	0.5514	0.7052	0.8952
p = 10	0	0.5	0.9953	1.0000	1.0000	0.7139	0.8957	0.9934	0.9227	0.9872	0.9999	0.6011	0.7131	0.8720
		1	0.9996	1.0000	1.0000	0.7678	0.9476	0.9990	0.9599	0.9977	1.0000	0.6367	0.7810	0.9346
		2	1.0000	1.0000	1.0000	0.8812	0.9903	0.9999	0.9895	0.9999	1.0000	0.7353	0.8737	0.9812
	0.25	0.5	0.9947	1.0000	1.0000	0.6850	0.8697	0.9896	0.9223	0.9868	0.9996	0.5964	0.6914	0.8481
		1	0.9993	1.0000	1.0000	0.7346	0.9293	0.9988	0.9620	0.9973	1.0000	0.6287	0.7620	0.9199
		2	1.0000	1.0000	1.0000	0.8425	0.9802	0.9999	0.9888	0.9997	1.0000	0.7054	0.8446	0.9732

Table 16: Power when using univariate monitoring procedures: scalar $BEKK$ time series and all of the variances increase.

			normal distribution						t distribution					
			$k^* = 0.05$			$k^* = 0.5$			$k^* = 0.05$			$k^* = 0.5$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000	$m = 500$	1000	2000	$m = 500$	1000	2000
$p = 2$	0	0.5	0.9095	0.9958	1.0000	0.4507	0.7100	0.9453	0.7367	0.9279	0.9962	0.3457	0.5112	0.7653
		1	0.9785	1.0000	1.0000	0.6033	0.8653	0.9945	0.8710	0.9813	0.9996	0.4473	0.6616	0.8949
		2	0.9957	1.0000	1.0000	0.7160	0.9497	0.9999	0.9285	0.9932	0.9995	0.5292	0.7599	0.9532
	0.25	0.5	0.8975	0.9944	1.0000	0.4309	0.6846	0.9381	0.7297	0.9196	0.9947	0.3366	0.4876	0.7357
		1	0.9750	0.9999	1.0000	0.5588	0.8394	0.9902	0.8592	0.9770	0.9994	0.4299	0.6328	0.8790
		2	0.9955	1.0000	1.0000	0.6964	0.9356	0.9993	0.9224	0.9919	0.9995	0.5083	0.7353	0.9455
$p = 5$	0	0.5	0.8684	0.9908	1.0000	0.3904	0.6333	0.9047	0.6736	0.8966	0.9909	0.3282	0.4417	0.6878
		1	0.9695	0.9999	1.0000	0.5594	0.8340	0.9863	0.8340	0.9715	0.9978	0.4315	0.6155	0.8659
		2	0.9910	1.0000	1.0000	0.6351	0.9132	0.9980	0.8934	0.9879	0.9987	0.4846	0.6865	0.9221
	0.25	0.5	0.8554	0.9867	1.0000	0.3653	0.5873	0.8846	0.6773	0.8923	0.9898	0.3387	0.4318	0.6681
		1	0.9585	0.9994	1.0000	0.5095	0.7948	0.9796	0.8219	0.9664	0.9973	0.4203	0.5829	0.8399
		2	0.9907	0.9999	1.0000	0.6262	0.9021	0.9980	0.8898	0.9872	0.9986	0.4856	0.6712	0.9126
$p = 10$	0	0.5	0.8388	0.9879	1.0000	0.3676	0.5919	0.8820	0.6612	0.8711	0.9903	0.3401	0.4302	0.6469
		1	0.9599	0.9993	1.0000	0.5167	0.7915	0.9787	0.8167	0.9611	0.9984	0.4340	0.5860	0.8226
		2	0.9854	1.0000	1.0000	0.5763	0.8748	0.9953	0.8709	0.9800	0.9994	0.4640	0.6514	0.8976
	0.25	0.5	0.8393	0.9857	1.0000	0.3716	0.5752	0.8677	0.6824	0.8723	0.9901	0.3845	0.4463	0.6354
		1	0.9577	0.9989	1.0000	0.5168	0.7788	0.9756	0.8277	0.9602	0.9983	0.4735	0.5954	0.8143
		2	0.9838	1.0000	1.0000	0.5572	0.8563	0.9942	0.8696	0.9783	0.9993	0.4795	0.6418	0.8846

Table 17: Power when using univariate monitoring procedures: scalar $BEKK$ time series and only one of the variances increases.

			normal distribution						t distribution					
			$k^* = 0.05$			$k^* = 0.5$			$k^* = 0.05$			$k^* = 0.5$		
	γ	B	$m = 500$	1000	2000	$m = 500$	1000	2000	$m = 500$	1000	2000	$m = 500$	1000	2000
$p = 2$	0	0.5	0.7807	0.9559	0.9990	0.3618	0.5348	0.7816	0.6241	0.8071	0.9614	0.3136	0.3901	0.5716
		1	0.8966	0.9933	1.0000	0.4497	0.6823	0.9162	0.7451	0.9112	0.9931	0.3738	0.5050	0.7194
		2	0.9556	0.9994	1.0000	0.5335	0.7911	0.9715	0.8385	0.9619	0.9980	0.4442	0.6063	0.8187
	0.25	0.5	0.7766	0.9521	0.9989	0.3408	0.5036	0.7546	0.6279	0.8001	0.9563	0.3127	0.3699	0.5420
		1	0.8895	0.9910	1.0000	0.4205	0.6440	0.8970	0.7410	0.9043	0.9911	0.3607	0.4759	0.6876
		2	0.9498	0.9988	1.0000	0.4972	0.7612	0.9651	0.8313	0.9570	0.9975	0.4228	0.5738	0.7890
$p = 5$	0	0.5	0.9589	0.9994	1.0000	0.5760	0.7803	0.9645	0.8156	0.9416	0.9962	0.4771	0.5696	0.7556
		1	0.9914	0.9999	1.0000	0.7058	0.9143	0.9969	0.9115	0.9866	0.9997	0.5624	0.6920	0.8862
		2	0.9989	1.0000	1.0000	0.8065	0.9653	0.9997	0.9570	0.9969	1.0000	0.6372	0.7955	0.9505
	0.25	0.5	0.9568	0.9994	1.0000	0.5514	0.7541	0.9564	0.8212	0.9397	0.9957	0.4742	0.5493	0.7296
		1	0.9913	0.9999	1.0000	0.6816	0.8998	0.9964	0.9144	0.9861	0.9997	0.5556	0.6740	0.8684
		2	0.9986	1.0000	1.0000	0.7865	0.9581	0.9997	0.9568	0.9968	1.0000	0.6228	0.7747	0.9403
$p = 10$	0	0.5	0.9998	1.0000	1.0000	0.8720	0.9845	1.0000	0.9453	0.9906	0.9997	0.6657	0.7525	0.9024
		1	1.0000	1.0000	1.0000	0.9604	0.9994	1.0000	0.9858	0.9998	1.0000	0.7544	0.8717	0.9721
		2	1.0000	1.0000	1.0000	0.9888	1.0000	1.0000	0.9955	0.9998	1.0000	0.8266	0.9287	0.9936
	0.25	0.5	0.9997	1.0000	1.0000	0.8578	0.9814	1.0000	0.9487	0.9902	0.9997	0.6690	0.7395	0.8920
		1	1.0000	1.0000	1.0000	0.9535	0.9991	1.0000	0.9860	0.9998	1.0000	0.7472	0.8574	0.9665
		2	1.0000	1.0000	1.0000	0.9854	0.9999	1.0000	0.9958	0.9998	1.0000	0.8191	0.9190	0.9920

Table 18: Power when the variance of the innovations increases.