CHANGE POINT DETECTION IN THE CONDITIONAL CORRELATION STRUCTURE OF MULTIVARIATE VOLATILITY MODELS

MARCO BARASSI, LAJOS HORVÁTH, AND YUQIAN ZHAO

Abstract. We propose semi-parametric CUSUM tests to detect a change point in the correlation structures of non–linear multivariate models with dynamically evolving volatilities. The asymptotic distributions of the proposed statistics are derived under mild conditions. We discuss the applicability of our method to the most often used models, including constant conditional correlation (CCC), dynamic conditional correlation (DCC), BEKK, corrected DCC and factor models. Our simulations show that, our tests have good size and power properties. Also, even though the near–unit root property distorts the size and power of tests, de–volatizing the data by means of appropriate multivariate volatility models can correct such distortions. We apply the semi–parametric CUSUM tests in the attempt to date the occurrence of financial contagion from the U.S. to emerging markets worldwide during the great recession.

JEL Classification: C12, C14, C32, G10, G15

Keywords: Change point detection, Time varying correlation structure, Volatility processes, Monte Carlo simulation, Contagion effect

1. Introduction

Multivariate models are widely used in financial applications. The development of technology and the increased computational ability, together with the availability of data at higher frequencies, have made more feasible modeling and estimating

systems of larger dimensions. The second moment dynamics of multivariate processes play a crucial role in the understanding of the relationship between economic and especially financial observations. Hence the literature on multivariate volatilities, especially on GARCH-type models has become rich. The BEKK model (Engle and Kroner, 1995), and generalizations of the constant conditional correlation–CCC model of Bollerslev (1990), including the dynamic conditional correlation–DCC model (Engle, 2002), and their extensions (cf. Cappiello et al., 2006; Aielli, 2013), are often used in econometrics. For reviews refer to Bauwens et al. (2006), Engle (2009), Silvennoinen and Teräsvirta (2009) and Francq and Zakoian (2010).

However, all these popular models, like every empirical model in econometrics, must account for changes in their parameters which might arise as a result of sudden shocks occurring in the economy, such as, market crashes, financial crises or intervention of policy markers. As a result, both parametric and non–parametric tests for change point detection have been developed to test the stability of the mean of independent observations and their asymptotic distributions have been derived (cf. Csörgő and Horváth, 1997). Aue and Horváth (2013) and Horváth and Rice (2014) reviewed several methods on how to derive asymptotic properties of popular methods when dependence between the observations cannot be neglected and the data structure is high dimensional. From the statistical point of view, likelihood–based parametric tests have been widely used due to their optimality properties. Nonetheless, non-parametric, especially CUSUM–based approaches have become popular since they are easy to apply and usually robust to model specifications.

Indeed, non–parametric methods have been developed in the literature and found their natural application to financial time series. Inclan and Tiao (1994) made the first attempt on change point detection in the variance of independent observations using the cumulative sum of the squares of the residuals. De Pooter and Van Dijk (2004) used a CUSUM test to detect a permanent change in the variance of a heteroscedastic process. Lee et al. (2003) also used the CUSUM statistics to test for changes in the variances of non-stationary AR(q) sequences. In the context of financial data, second moments are usually modeled by ARCH or GARCH–type models. Kokoszka and Leipus (2000) and Ling (2007) examined the behavior of change point tests in processes with dependent volatility. Their findings showed that CUSUM tests are valid when applied to short memory ARCH/GARCH model making feasible to detect changes within certain types of ARCH models in financial data (cf. also Andreou and Ghysels, 2002; Fryzlewicz and Rao, 2011). Andreou and Ghysels (2002) applied the test of Kokoszka and Leipus (2000) to detect multiple changes in the volatility of high frequency stock and foreign exchange data, where the conditional variance is captured by a GARCH model. Change points detection in the second moment is not limited to univariate cases, but it can be extended to the covariance and correlation structure of multivariate models. For an example of parametric likelihood ratio type tests applied to a context similar to ours, see Qu and Perron (2007). Early studies on change points detection in the covariance structure were focused on using model selection criteria and standard stability tests on the parameters of GARCH models. For example, Lavielle and Teyssiere (2006) proposed a penalized contrast function to detect simultaneous multiple changes in covariance structures. Andreou

and Ghysels (2003) modeled multivariate data by DCC model and then detected parameter changes through the test of Bai and Perron (1998). More recently, Aue et al. (2009), constructed CUSUM statistics for detecting changes in the covariance structure of multivariate stationary sequences, e.g. CCC sequences, and derived their asymptotics. Their tests were designed to examine the stability of cross-volatilities, however, studying just the pure correlation relationships sometimes is an issue to assets or other financial variables. To this end, Wied et al. (2012) extended the work of Aue et al. (2009) to study the stability of the correlation matrix.

The present paper aims to contribute to the literature by proposing semi–parametric tests for the stability of the conditional correlations in multivariate GARCH models. Compared with the existing works, we show that the asymptotics of non-parametric CUSUM tests in Aue et al. (2009) and Wied et al. (2012) are still valid in multivariate GARCH models with dynamically evolving conditional correlations, such as the BEKK (Engle and Kroner, 1995) and corrected DCC (Aielli, 2013) processes, and that therefore, the tests can be applied to detect correlation change–points in the pervasive framework often used in financial econometrics. Our Monte Carlo simulations show that the proposed semi–parametric tests are reasonably sized and display good power even in relatively small samples. We also apply the proposed test to detect the occurrence of financial contagion (Forbes and Rigobon, 2002), from the U.S. to emerging markets worldwide. Specifically, using data on Latin American, Central East European and East Asian stock markets, we find evidence of contagion from U.S. to these three regions during the

Great Recession. However, the transmission from U.S. to the East Asian markets is not as strong as that found towards the two other regions.

This paper is structured as follows. Section 2 introduces the semi–parametric CUSUM tests and their properties. Section 3 provides examples of models for which the assumptions of our theoretical framework are satisfied. In Section 4, we assess the finite sample performances of the proposed tests. Section 5 provides an empirical application in the context of tests for global financial contagion, and some concluding remarks are offered in Section 6. More discussions on the verification of regularity conditions are documented in the online supplementary material. Further examples of the models and Monte Carlo simulations are provided in the Online Supplement.

2. Test for the stability of time–varying correlation structures

In this section, we modify the test of Aue et al. (2009) and extend it to the cases where the correlation structure of observations evolves according to popular specifications of multivariate GARCH models. To detect changes in the correlation structure, this paper uses de–volatilized data to remove the influence from volatilities. Let $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T$ denote the observations, and write $\mathbf{y}_t = (y_t(1), y_t(2), \ldots, y_t(d))^\top$. The conditional variance of $y_t(j)$ given the past is denoted by $\tau_t^2(j)$, i.e. $\tau_t^2(j) = E(y_t^2(j)|\mathcal{F}_{t-1})$, where the σ–algebra $\mathcal{F}_{t-1}$ is generated by $\{\mathbf{y}_s, s \leq t-1\}$. The de–volatilized observations are denoted by

$$\mathbf{y}_t^* = (y_t^*(1), y_t^*(2), \ldots, y_t^*(d))^\top \text{ with } y_t^*(j) = \frac{y_t(j)}{\tau_t(j)}, \quad 1 \leq t \leq T, 1 \leq j \leq d.$$

Our paper follows the methodology of the most often used multivariate volatility models

$$\mathbf{y}_t = \mathbf{\Sigma}_t^{1/2}\mathbf{e}_t, \tag{2.1}$$

where the following conditions hold:

Assumption 2.1. $\{\mathbf{e}_t, -\infty < t < \infty\}$ *are independent and identically distributed random vectors in* R^d *with* $E\mathbf{e}_t = \mathbf{0}$ *and* $E\mathbf{e}_t\mathbf{e}_t^\top = \mathbf{I}_d$, *where* $\mathbf{I}_d$ *is the* $d \times d$ *identity matrix,*

Assumption 2.2. $\mathbf{\Sigma}_t \in \mathcal{F}_{t-1}$ *and* $\{\mathbf{\Sigma}_t, -\infty < t < \infty\}$ *is a stationary and ergodic sequence.*

Hence the conditional covariance matrix of $\mathbf{y}_t$ with respect to its past is $E(\mathbf{y}_t\mathbf{y}_t^\top|\mathcal{F}_{t-1}) = \mathbf{\Sigma}_t$. To avoid degenerate cases we assume that

Assumption 2.3. *There exists a positive definite lower bound matrix* $\mathbf{\Sigma}^0$ *such that* $\mathbf{\Sigma}_t - \mathbf{\Sigma}^0$ *is non–negative definite for all* t.

If $\mathbf{\Sigma}_t = \{\sigma_t(k,j), 1 \le k, j \le d\}$, then $\tau_t(j) = \sigma_t^{1/2}(j,j)$. It follows from Assumption 2.3 that there is a positive constant τ_0 such that $\tau_t(j) \ge \tau_0$ for all t and $1 \le j \le d$. It is an immediate consequence of Assumptions 2.1 and 2.2 that $\mathbf{y}_t$ is a stationary and ergodic sequence. The next condition is on the dependence structure of the observations. Let $\|\cdot\|$ denote the Euclidean norm of vectors and matrices.

Assumption 2.4. $E\|\mathbf{y}_t\|^r$ *with some* $r > 4$ *and* $\{\mathbf{y}_t, -\infty < t < \infty\}$ *is* β*–mixing with rate* $t^{-\delta - r/(r-2)}$ *with some* $\delta > 0$.

The mixing condition is very mild since in the examples we discuss in this paper, the rate of mixing is exponential. We note that Assumption 2.4 can be replaced with the conditions that $E\|\mathbf{e}_t\|^r < \infty$, $E\|\mathbf{\Sigma}_t\|^{r/2} < \infty$ and $\{\mathbf{\Sigma}_t, -\infty < t < \infty\}$ is β–mixing.

Let $\rho_t(i,j) = Ey_t^*(i)y_t^*(j), 1 \le t \le T, 1 \le i,j \le d$ be the covariance of the de–volatized observations $y_t^*(i)$ and $y_t^*(j)$. The objective of this paper is to test the null hypothesis that

$$H_0: \quad \rho_1(i,j) = \rho_2(i,j) = \ldots = \rho_T(i,j) \ \text{ for all } \ 1 \le i,j \le d$$

against the alternative

$H_A:$ there are $1 < t^* < T$ and $1 \le i^*, j^* \le d$ such that

$$\rho_1(i^*,j^*) = \rho_2(i^*,j^*) = \ldots = \rho_{t^*}(i^*,j^*) \ne \rho_{t^*+1}(i^*,j^*) = \ldots = \rho_T(i^*,j^*).$$

Under the null hypothesis the covariance matrix of the vector $(y_t^*(1), y_t^*(2), \ldots, y_t^*(d))^\top$ does not depend on the time t while under the alternative at least one of the elements of the covariance matrix changes at an unknown time t^*.

Let vech be the operator which stacks the columns of a symmetric matrix starting with the diagonals into a vector. Our procedure is based on two functionals of the CUSUM of the vectors $\mathfrak{r}_t = \text{vech}\left(y_t^*(i)y_t^*(j), 1 \le i,j \le d\right), \mathfrak{r}_0 = \mathbf{0}$. Define the partial sum process

$$\mathfrak{s}(t) = \sum_{s=1}^{t} \mathfrak{r}_s, \quad \text{and} \quad \mathfrak{s}(0) = 0.$$

Assuming that H_0 holds, i.e. the data are stationary we define the long run covariance matrix

$$\mathfrak{D} = \sum_{s=-\infty}^{\infty} E\mathfrak{r}_0\mathfrak{r}_s^\top.$$

The normalization in our procedures requires

Assumption 2.5. $\mathfrak{D}$ *is a nonsingular matrix.*

Following Aue et al. (2009) and Wied et al. (2012) we define two statistics

$$M_T^{(1)} = \frac{1}{T}\max_{1\le t\le T}\left(\mathfrak{s}(t) - \frac{t}{T}\mathfrak{s}(T)\right)^\top \mathfrak{D}^{-1}\left(\mathfrak{s}(t) - \frac{t}{T}\mathfrak{s}(T)\right)$$

and

$$M_T^{(2)} = \frac{1}{T^2}\sum_{t=1}^{T}\left(\mathfrak{s}(t) - \frac{t}{T}\mathfrak{s}(T)\right)^\top \mathfrak{D}^{-1}\left(\mathfrak{s}(t) - \frac{t}{T}\mathfrak{s}(T)\right).$$

Theorem 2.1. *If H_0 and Assumptions 2.1–2.5 hold, then*

$$M_T^{(1)} \xrightarrow{\mathcal{D}} M^{(1)} \quad \textit{and} \quad M_T^{(2)} \xrightarrow{\mathcal{D}} M^{(2)}, \tag{2.2}$$

where

$$M^{(1)} = \sup_{0\le u\le 1}\sum_{i=1}^{\bar{d}} B_i^2(u) \quad \textit{and} \quad M^{(2)} = \sum_{i=1}^{\bar{d}}\int_0^1 B_i^2(u)du \quad \textit{with } \bar{d} = d(d+1)/2,$$

and $B_1, B_2, \ldots, B_{\bar{d}}$ denote independent Brownian bridges.

The proof is given in Appendix A. The limiting random variables $M^{(1)}$ and $M^{(2)}$ already appeared in Aue et al. (2009), where selected critical values and approximations for moderate and large values of $\bar{d}$ can also be found. The applicability of Theorem 2.1 requires the estimation of $\mathfrak{D}$ which will be discussed before Theorem 2.2.

The conditional covariance matrices $\mathbf{\Sigma}_t$ can be written as functionals of the random vectors $\mathbf{y}_s, s \leq t-1$. However, since we can observe only $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T$, first we replace $\tau_t(i)$ with $\bar{\tau}_t(i)$, where $\bar{\tau}_t(i)$ is a function of $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}$ only. In parametric models, $\tau_t(i)$ as well as $\bar{\tau}_t(i)$ depend on unknown parameters which will be denoted by $\boldsymbol{\theta} \in R^p$. We require that $\bar{\tau}_t(i;\boldsymbol{\theta})$ and $\tau_t(i;\boldsymbol{\theta})$ are close, if t is large. This requirement is standard in the estimation of GARCH and similar volatility processes (cf. Francq and Zakoian, 2010):

Assumption 2.6. *There is a ball $\mathbf{\Theta}_0 \subset R^p$ with center $\boldsymbol{\theta}_0$ and a sequence $a(t)$ satisfying $t \cdot a(t) \to 0$ $(t \to \infty)$ such that $\max_{1\leq i\leq d} \sup_{\boldsymbol{\theta}\in\mathbf{\Theta}_0} |\tau_t(i;\boldsymbol{\theta}) - \bar{\tau}_t(i;\boldsymbol{\theta})| = O(a(t))$ a.s. as $t \to \infty$.*

Assumption 2.6 means that the difference between the stationary $\tau_t(i;\boldsymbol{\theta})$ and the nonstationary $\bar{\tau}_t(i;\boldsymbol{\theta})$ is small, i.e. there is a negligible effect that either the estimation is based on information $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}$ or $\{\mathbf{y}_s, s \leq t-1\}$ when t is large. We estimate $\boldsymbol{\theta}_0$ with $\hat{\boldsymbol{\theta}}_T$ which is consistent with rate $T^{-1/2}$:

Assumption 2.7. $\|\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\| = O_P(T^{-1/2})$, *where $\boldsymbol{\theta}_0$ denotes the value of the parameter under H_0.*

The random functions $\tau_t(i) = \tau_t(i;\boldsymbol{\theta}), 1 \leq i \leq p$, are smooth functions of $\boldsymbol{\theta}$ in a neighbourhood of $\boldsymbol{\theta}_0$:

Assumption 2.8. *There is a ball $\mathbf{\Theta}_0 \subset R^p$ with center $\boldsymbol{\theta}_0$ such that*

$$\left|\tau_t(i;\boldsymbol{\theta}) - \tau_t(i;\boldsymbol{\theta}_0) - \mathbf{g}_t^\top(i)(\boldsymbol{\theta} - \boldsymbol{\theta}_0)\right| \leq \bar{g}_t \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|^2$$

for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$*, where* $\{\mathbf{g}_t(i), 1 \le i \le p, \bar{g}_t, -\infty < t < \infty\}$ *is a stationary and ergodic sequence with* $E\|\mathbf{g}_0(i)\|^2 < \infty$ *and* $E|\bar{g}_0|^2 < \infty$*.*

The quasi maximum likelihood method (QMLE hereafter) is the most often used technique to estimate the parameters of a multivariate GARCH model. In the examples discussed in this paper, the QMLE satisfies Assumptions 2.6–2.8. Now the de–volatized variables

$$\hat{y}_t(i) = \frac{y_t(i)}{\bar{\tau}_t(i; \hat{\boldsymbol{\theta}}_T)}$$

can be computed from the sample. Let $\hat{\mathfrak{r}}_s = \mathrm{vech}(\hat{y}_s(i)\hat{y}_s(j), 1 \le i, j \le d)$.

The long run covariance matrix $\mathfrak{D}$ is estimated from the sample by $\hat{\mathfrak{D}}_T$ which satisfies

Assumption 2.9. $\|\hat{\mathfrak{D}}_T - \mathfrak{D}\| = o_P(1)$.

We propose the kernel estimators

$$\hat{\mathfrak{D}}_T = \sum_{\ell=-T}^{T} K\left(\frac{\ell}{h}\right) \hat{\boldsymbol{\gamma}}_\ell,$$

where

$$\hat{\boldsymbol{\gamma}}_\ell = \begin{cases} \dfrac{1}{T}\displaystyle\sum_{t=1}^{T-\ell} (\hat{\mathfrak{r}}_t - \bar{\mathfrak{r}}_T)(\hat{\mathfrak{r}}_{t+\ell} - \bar{\mathfrak{r}}_T)^\top, & \text{if } \ 0 \le \ell < T \\ \dfrac{1}{T}\displaystyle\sum_{t=-\ell+1}^{T} (\hat{\mathfrak{r}}_t - \bar{\mathfrak{r}}_T)(\hat{\mathfrak{r}}_{t+\ell} - \bar{\mathfrak{r}}_T)^\top, & \text{if } \ -T < \ell < 0. \end{cases}$$

where

$$\bar{\mathfrak{r}}_T = \frac{1}{T}\sum_{s=1}^{T} \hat{\mathfrak{r}}_s.$$

There are several choices for the kernel K, including the Bartlett, truncated, Parzen, Tukey–Hanning and quadratic spectral kernels (cf. Andrews, 1991 for a

review of the properties of kernel functions). The window (smoothing parameter) satisfies $h = h(T), h/T \to \infty$ and $h/T \to 0$. Following Wu and Zaffaroni (2018), Assumption 2.9 can be established.

Similarly to $M_T^{(1)}$ and $M_T^{(2)}$ we define

$$\hat{M}_T^{(1)} = \frac{1}{T} \max_{1 \le t \le T} \left(\hat{\mathfrak{s}}(t) - \frac{t}{T} \hat{\mathfrak{s}}(T) \right)^\top \hat{\mathfrak{D}}_T^{-1} \left(\hat{\mathfrak{s}}(t) - \frac{t}{T} \hat{\mathfrak{s}}(T) \right)$$

and

$$\hat{M}_T^{(2)} = \frac{1}{T^2} \sum_{t=1}^{T} \left(\hat{\mathfrak{s}}(t) - \frac{t}{T} \hat{\mathfrak{s}}(T) \right)^\top \hat{\mathfrak{D}}_T^{-1} \left(\hat{\mathfrak{s}}(t) - \frac{t}{T} \hat{\mathfrak{s}}(T) \right),$$

where

$$\hat{\mathfrak{s}}(t) = \sum_{s=1}^{t} \hat{\mathfrak{r}}_s.$$

Theorem 2.2. *If H_0 and Assumptions 2.1–2.9 hold, then*

$$\hat{M}_T^{(1)} \xrightarrow{\mathcal{D}} M^{(1)} \quad \text{and} \quad \hat{M}_T^{(2)} \xrightarrow{\mathcal{D}} M^{(2)}, \tag{2.3}$$

where $M^{(1)}$ and $M^{(2)}$ are defined in Theorem 2.1.

The proof is given in Appendix A. It follows from (2.1) and Assumptions 2.1 and 2.2 that $E\mathbf{y}_t = \mathbf{0}$. If the mean of the observations is not $\mathbf{0}$, i.e. $\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\Sigma}_t^{1/2}\mathbf{e}_t$, the results in Theorems 2.1 and 2.2 remained valid when $\boldsymbol{\mu}$ is removed, i.e. the analysis is based on $\mathbf{y}_t - \hat{\mathbf{y}}_T$ with $\hat{\mathbf{y}}_T = \sum_{t=1}^{T} \mathbf{y}_t/T$. It has been observed for a long time in the literature that demeaning does not change the asymptotic distribution of residual based tests (cf., for example, Kulperger and Yu, 2005 and Demetrescu and Wied, 2016 for de–meaning in time series). Besides, if the conditional mean is introduced and which is removed by suitable estimators, this will change the

asymptotic distribution and the test statistic will depend on the values of some unknown parameters.

3. Examples of time dependent conditional volatilities

Here we briefly describe how our test is valid when applied to two typical examples of multivariate GARCH models, as they are of interest for practitioners. More examples with other parameterizations such as the CCC, DCC and Factor-GARCH are discussed in the online supplement.

Example 3.1. (BEKK model) Baba, Engle, Kraft and Kroner (cf. Engle and Kroner, 1995) introduced the model where the conditional covariance matrix satisfies the recursion

$$\mathbf{\Sigma}_t = \mathbf{C} + \sum_{j=1}^{q} \mathbf{A}_j \mathbf{y}_{t-j} (\mathbf{A}_j \mathbf{y}_{t-j})^\top + \sum_{k=1}^{p} \mathbf{B}_k \mathbf{\Sigma}_{t-k} \mathbf{B}_k^\top, \tag{3.1}$$

where $\mathbf{C}$, $\mathbf{A}_j$, $1 \le j \le q$, and $\mathbf{B}_k$, $1 \le k \le p$, are $d \times d$ matrices, and $\mathbf{C}$ is positive definite. The parameters of the BEKK sequences can be estimated by the QMLE and the variance targeting QMLE (cf. Comte and Lieberman, 2003, Hafner and Preminger, 2009, Pedersen and Rahbek, 2014 and Francq et al., 2016). In the Online Supplement we discuss the BEKK models and how the QML type estimators satisfy the conditions of Theorem 2.2. For further details on the BEKK model, we refer to Francq and Zakoian (2010).

Example 3.2. (Corrected dynamic conditional correlation) Following Aielli (2013) we introduce the corrected DCC (cDCC) model:

$$\mathbf{\Sigma}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t, \tag{3.2}$$

where $\mathbf{D}_t$ is a diagonal matrix, $\mathbf{D}_t = \text{diag}(\tau_t(1), \tau_t(2), \ldots, \tau_t(d))$. It is assumed that $y_t(i)$ is modeled as a univariate GARCH process, $\tau_t^2(i) = h_i(\boldsymbol{\zeta}_i, y_{t-1}(i), y_{t-2}(i), \ldots)$, $i = 1, 2, \ldots, d$, where h_i is a known function and $\boldsymbol{\zeta}_i, 1 \leq i \leq d$ are unknown parameters. The conditional correlation of $\mathbf{y}_t$ satisfies

$$\mathbf{R}_t = (\text{diag}(\mathbf{Q}_t))^{-1/2}\mathbf{Q}_t(\text{diag}(\mathbf{Q}_t))^{-1/2}, \tag{3.3}$$

and

$$\mathbf{Q}_t = \theta_1\mathbf{C}+\theta_2[(\text{diag}(\mathbf{Q}_{t-1}))^{1/2}\mathbf{y}^*_{t-1}(\mathbf{y}^*_{t-1})^\top(\text{diag}(\mathbf{Q}_{t-1}))^{1/2}] + \theta_3\mathbf{Q}_{t-1}, \tag{3.4}$$

where $\mathbf{C}$ is a positive definite matrix, $\theta_1 > 0, \theta_i \geq 0, i = 2, 3$ satisfy $\theta_1+\theta_2+\theta_3 = 1$. The parameters of the process are $\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_d, \theta_2, \theta_3$ and $\mathbf{C}$. In principle, the QMLE method could be used, but due to the large number of parameters it is infeasible. To overcome the problem, Aielli (2013) suggested a three–step procedure. Following Aielli (2013), we show in the Online Supplement that the conditions of Theorem 2.2 hold. Since there are several univariate asymmetric GARCH models (cf. Francq and Zakoian, 2010), the cDCC model accounts for possible asymmetry of the returns.

4. The Monte Carlo simulations

To assess the performance of the statistics $\hat{M}_T^{(1)}$ and $\hat{M}_T^{(2)}$ under the conditions of Examples 3.1 and 3.2, we conduct a Monte Carlo simulation to study the rejection rates under the null and alternative hypotheses in finite samples. We only report our findings for $\hat{M}_T^{(2)}$ since the results for $\hat{M}_T^{(1)}$ are essentially the same. We first consider bivariate observations $\mathbf{y}_t = (y_t(1), y_t(2))^\top$. In the data

generating process (DGP) $\mathbf{e}_t$ is a bivariate standard normal vector and $\mathbf{\Sigma}_t^{1/2}$ of (2.1) is in Cholesky form. For each model, we set the initial value $\mathbf{\Sigma}_0$ to be the 2×2 identity matrix and simple iterations give $\mathbf{\Sigma}_t$ for the specified parameter values. The Bartlett kernel $K_B(x) = (1 - |x|)I\{|x| \leq 1\}$ and the Newey–West optimal window (smoothing parameter) are used in the definition of $\hat{\mathfrak{D}}_T$. The observations are first demeaned, i.e. the sample mean is removed from the observations. Assuming that a change occurred, we estimate the time of change with $\hat{t}_T = \text{argmax}\{\hat{\mathfrak{s}}(t) - (t/T)\hat{\mathfrak{s}}(T), 1 \leq t \leq T\}$. In our simulations the time of change is $t^* = T/2$. In each experiment, we set $T = 300$ for a small sample, roughly the number of trading days in 14 months, and $T = 1000$ for a large sample, trading days in four years. Each simulation is replicated 5000 times. The warming up parameter is 0.2, so the simulation will burn 200 observations if sample size is 1000.

We generate bivariate full–BEKK sequences of Example 3.1 ($p = q = 1$) with coefficient matrices

$$\mathbf{C} = \begin{bmatrix} 1 & \delta \\ \delta & 1 \end{bmatrix} \quad \mathbf{A}_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \mathbf{B}_1 = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}.$$

Keeping financial applications in mind, we choose $a_{11} = a_{22} = a = 0.1$ or 0.2 standing for relatively lower or higher ARCH effect, respectively. Coefficients $b_{11} = b_{22} = b = 0.8$ or 0.9 for relatively lower or higher persistence. We always set $a_{12} = a_{21} = b_{12} = b_{21} = 0.001$.

We also simulate bivariate cDCC sequences of Example 3.2, where $\mathbf{C}$ is the same as above. We set $\theta_2 = 0.005$ or 0.01 (relatively lower and higher ARCH effect in

quasi conditional correlation process), $\theta_3 = 0.9$ or 0.95 (relatively low and high persistence). The variances follow univariate GJR(1,1,1) with intercept 0.01, ARCH and GARCH coefficients 0.01 and 0.94, and the coefficient for the asymmetric term 0.01, respectively. The model is estimated by the 3–step estimation procedure (cf. Aielli, 2013).

We compute the empirical rejection rates for the BEKK and cDCC when δ of $\mathbf{C}$ changes from 0 to $\delta = 0.2, 0.4, 0.6$ and 0.8 at $t^* = T/2$ ($\delta = 0$ corresponds to the empirical rejection under the null hypothesis). Figure 4.1 shows the empirical rejections for $\hat{M}_T^{(2)}$ under Examples 3.1 and 3.2 for both small and large samples. Similar results can be obtained if δ is negative. Empirical and asymptotic critical values are given in the online supplement. Both tests are well sized under the null hypothesis. The powers are close to 1 when $\delta = 0.2$ in large samples and $\delta = 0.4$ in small samples. We also note that high ARCH and persistence show limited impact on the empirical size and power of our tests. Table 4.1 summarizes the results for the estimation of t^*. The results show that along with the change magnitude increasing, the standard deviations or the differences between two quantiles of the change point estimators are decreasing, thereby producing more accurate estimators.

FIGURE 4.1. Graphs of the power functions of $\hat{M}_T^{(2)}$ in the BEKK (left panel) and cDCC (right panel) models in case of $d = 2$, $T = 1000$ ($*$'s) and $T = 300$ (lines) at 95% significance level

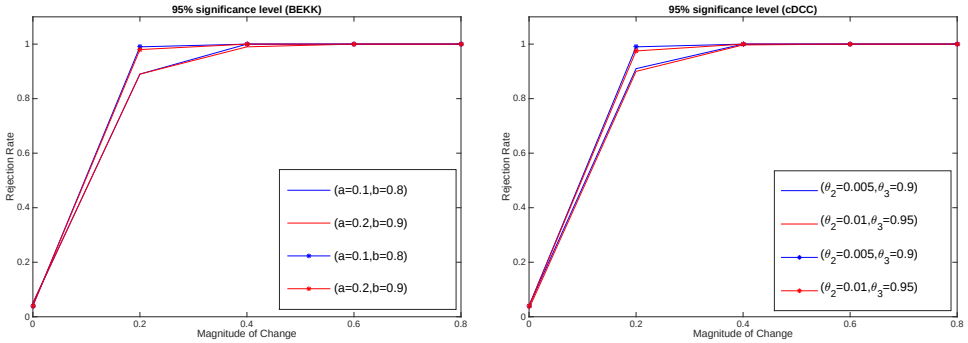
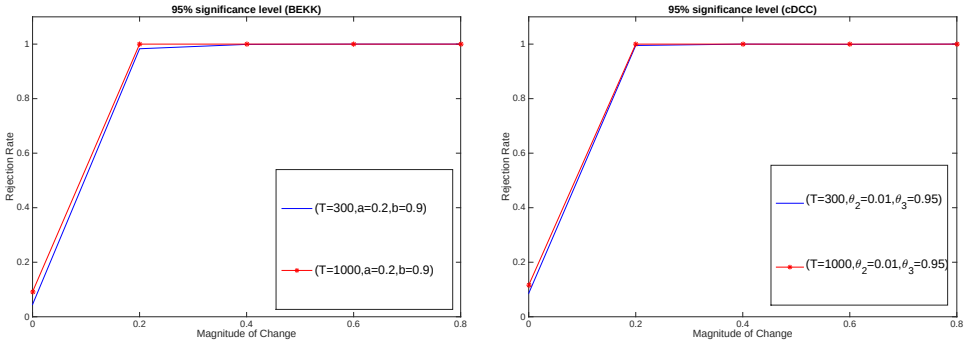


TABLE 4.1. Empirical performance of t_T^*, the estimator for $t^* = T/2$, when $d = 2$

	BEKK($a = 0.2\&b = 0.9$)								cDCC ($\theta_2 = 0.01\&\theta_3 = 0.95$)							
	T=300				T=1000				T=300				T=1000			
δ	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
Median	0.50	0.50	0.49	0.50	0.50	0.50	0.50	0.50	0.51	0.49	0.49	0.49	0.50	0.50	0.50	0.50
Quantile 0.1	0.31	0.39	0.44	0.46	0.38	0.46	0.48	0.49	0.31	0.38	0.43	0.46	0.37	0.45	0.48	0.49
Quantile 0.9	0.71	0.58	0.53	0.50	0.60	0.52	0.51	0.50	0.72	0.61	0.54	0.51	0.62	0.53	0.51	0.50

FIGURE 4.2. Graphs of the power functions of $\hat{M}_T^{(2)}$ in the BEKK (left panel) and cDCC (right panel) models in case of $d = 9$, $T = 1000$ ($*$'s) and $T = 300$ (lines) at 95% significance level



Next, in order to assess the validity of our tests in a higher dimensional environment, we simulate data with dimension $d = 9$ in concordance with the data set used in the application section. With regard to the BEKK model, we set coefficient matrices A_1 and B_1 with elements $a_{11} = a_{22} = \cdots = a_{99} = a = 0.2$, $b_{11} = b_{22} = \cdots = b_{99} = b = 0.9$, and all other off-diagonal elements are 0.001. In the cDCC model, we set parameters $\theta_2 = 0.01$ and $\theta_3 = 0.95$. Other settings are the same with those studied in the bivariate case except for the replications, which are reduced to 2000. Figure 4.2 plots the empirical rejection rates of $\hat{M}_T^{(2)}$ for small and large samples. There are two nontrivial observations. First, the test gains more powers even in small samples. This makes sense as the order of CUSUM statistics depends on $\bar{d}$ according to Remark 2.1 in Aue et al. (2009). Consequently, Table 4.2 reports the more accurate estimation of t^*. Second, the test looks slightly over–sized. We attribute this distortion to the finite sample bias of the Gaussian QMLE estimator in multivariate GARCH models. Note that the consistency of Gaussian QMLE works under strict stationarity condition, the

near–integrate higher dimensional processes generated in our simulations might produce more outliers. Hence the QMLE estimators might not be accurate for small sample sizes. A similar issue has been discussed in Boudt and Croux (2010).

TABLE 4.2. Empirical performance of t_T^*, the estimator for $t^* = T/2$, when $d = 9$

	BEKK($a = 0.2\&b = 0.9$)								cDCC ($\theta_2 = 0.01\&\theta_3 = 0.95$)							
	T=300				T=1000				T=300				T=1000			
δ	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
Median	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
Quantile 0.1	0.43	0.48	0.49	0.49	0.49	0.50	0.50	0.50	0.43	0.46	0.48	0.49	0.48	0.50	0.50	0.50
Quantile 0.9	0.56	0.51	0.50	0.50	0.52	0.50	0.50	0.50	0.57	0.51	0.50	0.50	0.52	0.51	0.50	0.50

5. An empirical application: testing for financial contagion

Forbes and Rigobon (2002) indicated that a financial contagion effect occurs if the inter–linkages across markets experienced a significant increase after some market events. Actual change dates in conditional correlations are unknown and need to be detected through statistical methods (cf. Dimitriou et al. 2013, Blatt et al. 2015 and Dungey et al. 2015).

We collect three groups of emerging stock market price indexes in three regions: six Latin American markets including Argentina (Argentina MERVAL), Brazil (Brazil BOVESPA), Chile (Chile Santiago SE General), Mexico (Mexico IPC), Colombia (Colombia IGBC), Peru (BVL General); seven Central East European (CEE hereafter) markets including Czech (Prague SEPX), Estonia (OMX Tallin),

Hungary (Budapest), Poland (Warsaw General), Romania (Romania BET), Slovakia (Slovakia SAX 16), Slovenia (Slovenian blue chip); nine East Asian markets including Hong Kong (Hang Seng), Indonesia (IDX composite), South Korea (Korea SE composite), Malaysia (Malaysia KLCI), Philippines (Philippine SE), Singapore (Straits Times), Taiwan (Taiwan SE weighted), Thailand (Bangkok S.E.T), China (Shanghai S.E. A share). The S&P 500 index of the United States is used as the eye of the storm for each group. The Germany index (DAX 40) and the Japan index (Nikki 225) are also collected due to their important influence on CEE and East Asian countries, respectively. The data are taken from the Datastream database and cover the period going from the 1st of September 2006 to the 1st of September 2010. We calculate log returns for each index to achieve the mean stationarity.

To find changes in the correlation structures of these three data sets, we use $\hat{M}_T^{(2)}$ in the BEKK as well as in the cDCC models. If a change is detected, we estimate the time of change and split the data into two subsets at the estimated time of change. Then we look for changes in both subsets (binary segmentation). Thus we segment the data into 6 homogeneous subsets. The change–point detection results are displayed in Figure 5.1. Overall, both models show consistent patterns. The correlation structures initially changed around February 2007 (Chinese stock bubble) and then changed around August 2007 (ceasing activities in the U.S. mortgage debt market), the third change happened close to September 2008 (the bankruptcy of Lehman Brothers), and the fourth and fifth changes occurred in the second half of 2009 (bailout decision made by G20 summit) and April 2010 (European debt crisis), respectively. Basically, these five dates split

the whole sample into six periods: boom, panic, bubble, bust, recovery and new crisis. The reactions to crises are faster in the Latin American region because of the higher level of integration with U.S. The capital flow from the U.S. has less impact on CEE markets due to the regional economic dominance by Germany. The banking systems in the CEE countries are largely dominated by U.S. and Western European banks/financial institutions, mainly German banks. Beside the direct falling in their capital flows with U.S. banks, the German authority, as a regional dominance, would implement appropriate policies to resist market risks during the crisis, thereby providing an indirect buffer zone to CEE countries. East Asian markets are relatively less connected with the U.S. and tend to have higher resistance, which might be explained with their closer relation with the large economies in the area, such as Japan and China.

For each of the six segments we compute $\bar{\delta}$, the level of regional integration and $\bar{\delta}_{US}$, regional correlation with U.S. market. We measure $\bar{\delta}$ and $\bar{\delta}_{US}$ by averaging off diagonal elements and U.S. related elements in the (empirical) correlation matrix, respectively, where the correlation matrix is computed via the estimated parameters of the underlying volatility model. Table 5 reports the results. Although the BEKK model gives relatively lower correlations, both models present similar features. Firstly, in case of regional integration level, the Latin American and East Asian regions are more integrated than the CEE regions. Secondly, the U.S. market has less impact on the CEE and least impact on the East Asian region. Finally, the integration levels in all regions keep increasing with some fluctuations, and the regional linkages with U.S. climb to a high point after September 2008, then decrease slightly and reboot to the peak again during the

European debt crisis. These results imply that contagion effects are significant in all data sets, resulting a higher integrated but more fragile global capital market.

FIGURE 5.1. Plots of the conditional correlations between the U.S. and Latin American (left column), CEE (middle column), Asian (right column) in the BEKK (first row) and cDCC (second row) models. The vertical lines are the estimated times of changes.

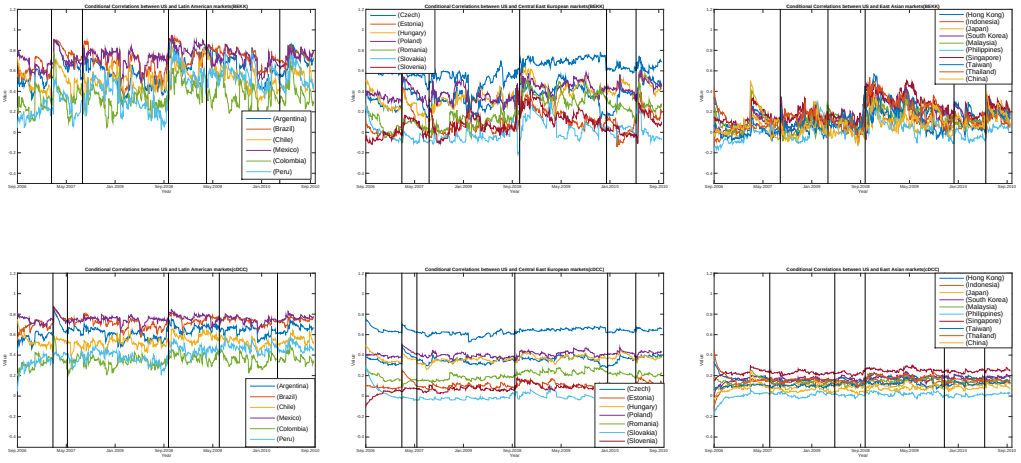


TABLE 5.1. The regional correlation levels and correlation levels with the U.S. market between 2006 and 2010

	Latin American Markets				Central East European Markets				East Asian Markets			
	BEKK		cDCC		BEKK		cDCC		BEKK		cDCC	
	$\bar{\delta}$	$\bar{\delta}_{US}$	$\bar{\delta}$	$\bar{\delta}_{US}$	$\bar{\delta}$	$\bar{\delta}_{US}$	$\bar{\delta}$	$\bar{\delta}_{US}$	$\bar{\delta}$	$\bar{\delta}_{US}$	$\bar{\delta}$	$\bar{\delta}_{US}$
Phase 1	0.26	0.21	0.38	0.42	0.15	0.17	0.21	0.28	0.20	0.03	0.34	0.16
Phase 2	0.28	0.28	0.52	0.65	0.15	0.04	0.28	0.21	0.20	0.01	0.44	0.04
Phase 3	0.28	0.23	0.46	0.47	0.19	0.13	0.25	0.16	0.19	0.05	0.32	0.05
Phase 4	0.33	0.36	0.64	0.65	0.20	0.24	0.34	0.33	0.19	0.07	0.46	0.20
Phase 5	0.33	0.31	0.55	0.61	0.18	0.15	0.26	0.28	0.21	0.02	0.40	0.14
Phase 6	0.41	0.39	0.61	0.70	0.17	0.17	0.40	0.35	0.23	0.09	0.43	0.23

6. CONCLUSION

In this paper, we suggested a semi-parametric CUSUM type test to detect a change point in the correlation structure of non–linear multivariate dynamically evolving volatility models, e.g. the BEKK and cDCC models, where the regularity conditions are satisfied. Simulations showed that the limit results work well for finite samples. We apply the test to date global financial contagion from the U.S market to three regions, including Latin American, Central East European and East Asian markets, between 1 September, 2006 and 1 September, 2010. Our tests allowed us to obtain the dates when contagion from the U.S. hit three sets of markets and noted that these dates are consistent with the dates when

particular events took place in the U.S.. The findings indicated that there were global contagion effects and resulted a more fragile global capital market.

It is worth noticing that, although our test is valid for models which asymmetry in the dynamics of conditional variance such as the cDCC model, a generalization to all asymmetric multivariate GARCH processes is not made here, but it will definitely be an object of future research. The main issue to overcome is that so far, there are only few theoretical results available on the consistency of estimators for stationary asymmetric multivariate GARCH processes. The general methods in Meyn and Tweedie (1993) can be in theory used, but it is clear from Boussama et al. (2011) and Fermanian and Malongo (2017) that the calculations will be lengthy using methods and results from probability theory and algebraic geometry. The parameters could be estimated by the QMLE. If dimension d is large, then a large number of parameters need to be estimated, however the variance targeting estimators could help to overcome numerical issues. As in Aue et al. (2009), the limits in Theorem 2.2 might be approximated well in case of moderate and large d. Also, detecting change points in conditional correlation structure with the non-zero conditional mean might be another subject of further research.

Appendix A. Proofs of Theorems 2.1 and 2.2

We start with the weak convergence of the process $\mathfrak{s}(t), 0 \leq t \leq T$.

Lemma A.1. *If H_0 and Assumptions 2.1–2.5 are satisfied, then we have*

$$T^{-1/2}(\mathfrak{s}(Tu) - E\mathfrak{s}(Tu)) \overset{\mathcal{D}^{\bar{d}}[0,1]}{\longrightarrow} \mathbf{W}_{\mathfrak{D}}(u),$$

where $\mathbf{W}_{\mathfrak{D}}(u), 0 \leq u \leq 1$ *is a Brownian motion in* $R^{\bar{d}}$ *with covariance matrix* $\mathfrak{D}$*, i.e.* $\mathbf{W}(u)$ *is Gaussian with* $E\mathbf{W}(u) = \mathbf{0}$ *and* $E\mathbf{W}_{\mathfrak{D}}(u)\mathbf{W}_{\mathfrak{D}}^{\top}(v) = \min(u, v)\mathfrak{D}$.

Proof. It follows from Assumptions 2.1–2.4 that $y_t^*(i)y^*(j)$ is also stationary and β–mixing with the same rate as of $\mathbf{y}_t$. Also, since Assumption 2.3 implies that $\tau_t(i) \geq \tau_0$ we get that

$$E|y_t^*(i)y_t^*(j)|^{r/2} \leq \frac{1}{\tau_0^2}(E|y_t^*(i)|^r E|y_t^*(j)|^r)^{1/2} < \infty$$

via the Cauchy–Schwartz inequality and the moment condition in Assumption 2.4. Hence the result of Ibragimov (1962) (cf. also Rio, 2000) implies the lemma. □

Proof of Theorem 2.1. Lemma A.1 implies that

$$T^{-1/2}\left(\mathfrak{s}(Tu) - \frac{\lfloor Tu \rfloor}{T}\mathfrak{s}(T)\right) \overset{\mathcal{D}^{\bar{d}}[0,1]}{\longrightarrow} \mathbf{W}_{\mathfrak{D}}(u) - u\mathbf{W}_{\mathfrak{D}}(1). \tag{A.1}$$

Checking the covariance structure, one can easily verify that

$$\left\{\mathfrak{D}^{-1/2}\left(\mathbf{W}_{\mathfrak{D}}(u) - u\mathbf{W}_{\mathfrak{D}}(1)\right), 0 \leq u \leq 1\right\} \overset{\mathcal{D}}{=} \left\{(B_1(u), B_2(u), \ldots, B_{\bar{d}}(u)), 0 \leq u \leq 1\right\}, \tag{A.2}$$

where $B_1, B_2, \ldots, B_{\bar{d}}$ are independent Brownian bridges. Hence Theorem 2.1 follows from (A.1) and (A.2) via the continuous mapping theorem. □

Proof of Theorem 2.2. It follows from the definition of $\hat{y}_t(i)$ that

$$\hat{y}_t(i)\hat{y}_t(j) - y_t^*(i)y_t^*(j) = a_{t,1}(i,j) + \ldots + \ldots + a_{t,8}(i,j),$$

$$a_{t,1}(i,j) = y_t(i)y_t(j)\left(\frac{1}{\bar{\tau}_t(i;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(i;\hat{\boldsymbol{\theta}}_T)}\right)\left(\frac{1}{\bar{\tau}_t(j;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_T)}\right),$$

$$a_{t,2}(i,j) = y_t(i)y_t(j)\left(\frac{1}{\bar{\tau}_t(i,\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(i,\hat{\boldsymbol{\theta}}_T)}\right)\left(\frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(j;\boldsymbol{\theta}_0)}\right),$$

$$a_{t,3}(i,j) = y_t(i)\left(\frac{1}{\bar{\tau}_t(i,\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(i,\hat{\boldsymbol{\theta}}_T)}\right)\frac{y_t(j)}{\tau_t(i)},$$

$$a_{t,4}(i,j) = y_t(i)y_t(j)\left(\frac{1}{\tau_t(i;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(i;\boldsymbol{\theta}_0)}\right)\left(\frac{1}{\bar{\tau}_t(j;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_T)}\right),$$

$$a_{t,5}(i,j) = y_t(i)y_t(j)\left(\frac{1}{\tau_t(i;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(i;\boldsymbol{\theta}_0)}\right)\left(\frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(j;\boldsymbol{\theta}_0)}\right),$$

$$a_{t,6}(i,j) = \frac{y_t(i)}{\tau_t(i)}y_t(j)\left(\frac{1}{\bar{\tau}_t(j;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_T)}\right),$$

$$a_{t,7}(i,j) = y_t(i)\left(\frac{1}{\tau_t(i;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(i;\boldsymbol{\theta}_0)}\right)\frac{y_t(j)}{\tau_t(j)},$$

$$a_{t,8}(i,j) = \frac{y_t(i)}{\tau_t(i)}y_t(j)\left(\frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(j;\boldsymbol{\theta}_0)}\right).$$

Since $\bar{\tau}_t(i,\hat{\boldsymbol{\theta}}_T)\tau_0 > 0$, by Assumptions 2.7 and 2.6 we have on account of the mean value theorem that

$$T^{-1/2}\max_{1\le t\le T}\sum_{s=1}^{t}|a_{s,1}| = O_P(1)T^{-1/2}\sum_{t=1}^{T}|y_t(i)y_t(j)|a^2(t).$$

We can assume without loss of generality that $a(t)$ is non increasing as $t\to\infty$. Using again Assumption 2.6 we can find a sequence a_T such that $T^{-1/2}a_T \to 0$

and $T^{1/2}a(a_T) \to 0$ and therefore

$$
\begin{aligned}
(\text{A.3}) \qquad & T^{-1/2}\sum_{t=1}^{T}|y_t(i)y_t(j)|a^2(t) \\
& \leq T^{-1/2}\sum_{t=1}^{a_T}|y_t(i)y_t(j)|a^2(t) + T^{-1/2}\sum_{t=a_T+1}^{T}|y_t(i)y_t(j)|a^2(t) \\
& = O_P(T^{-1/2}a_T + T^{1/2}a^2(a_T)) = o_P(1),
\end{aligned}
$$

where we used the ergodic theorem that

$$
\frac{1}{L}\sum_{t=1}^{L}|y_t(i)y_t(j)| \to E|y_0(i)y_0(j)| \quad \text{a.s. } (L \to \infty),
$$

since by Assumption 2.4 $E|y_0(i)y_0(j)| \leq (Ey_0^2(i)Ey_0^2(j))^{1/2} < \infty$. Putting together Assumptions 2.6–2.8 we conclude via two term Taylor expansion and the mean value theorem that

$$
T^{-1/2}\max_{1\leq t\leq T}\sum_{s=1}^{t}|a_{s,2}| = O_P(T^{-1/2})\sum_{t=1}^{T}|y_t(i)y_t(j)|a(t)\left[\|\mathbf{g}_t(j)\|\|\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}\| + \bar{g}_t\|\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\|^2\right].
$$

Following the proof of (A.3) one can show that

$$
T^{-1/2}\sum_{t=1}^{T}|y_t(i)y_t(j)|a(t)\|\mathbf{g}_t(j)\|\|\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\| = O_P(1)\frac{1}{T}\sum_{t=1}^{T}|y_t(i)y_t(j)|a(t)\|\mathbf{g}_t(j)\| = o_P(1),
$$

since $E|y_t(i)y_t(j)|\|\mathbf{g}_t(j)\| \leq (E|y_t(i)y_t(j)|^2E\|\mathbf{g}_t(j)\|^2)^{1/2} \leq (Ey_t^4(i)Ey_t^4(j))^{1/4}(E\|\mathbf{g}_t(j)\|^2)^{1/2} < \infty$. The same arguments give

$$
\begin{aligned}
T^{-1/2}\sum_{t=1}^{T}|y_t(i)y_t(j)|a(t)\bar{g}_t\|\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\|^2 &= O_P(1)\frac{1}{T^{3/2}}\sum_{t=1}^{T}|y_t(i)y_t(j)|a(t)\bar{g}_t \\
&= O_P(1)\left(\frac{1}{T^{1/2}}\max_{1\leq t\leq T}\bar{g}_t\right)\frac{1}{T}\sum_{t=1}^{T}|y_t(i)y_t(j)|a(t) = o_P(1),
\end{aligned}
$$

since $E\bar{g}_0^2 < \infty$ implies $\max_{1\le t\le T} \bar{g}_t = o_P(T^{1/2})$. Similarly,

$$T^{-1/2} \max_{1\le t\le T} \sum_{s=1}^{t} |a_{s,3}| = O_P(1) T^{-1/2} \sum_{t=1}^{T} |y_t(i) y_t(j)| a(t) = o_P(1)$$

and by symmetry, $T^{-1/2} \max_{1\le t\le T} \sum_{s=1}^{t} |a_{s,\ell}| = o_P(1), \ell = 4, 5, 6$. Assumption 2.8 implies that

$$T^{-1/2} \max_{1\le t\le T} \left| \sum_{s=1}^{t} a_{s,7} - \sum_{s=1}^{t} \frac{y_t(i)}{\tau_t^2(i)} \frac{y_t(j)}{\tau_t(j)} \mathbf{g}_s(i)^\top (\boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_T) \right|$$

$$= O_P(1) T^{-1/2} \sum_{s=1}^{T} |y_t(i) y_t(j)| \bar{g}_t \|\boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_T\|^2 = O_P(1) \left(T^{-1/2} \max_{1\le t\le T} \bar{g}_t \right) \frac{1}{T} \sum_{s=1}^{T} |y_t(i) y_t(j)| = o_P(1).$$

Using again the ergodic theorem and Assumption 2.4, we conclude

$$T^{-1/2} \max_{1\le t\le T} \left| \left(\sum_{s=1}^{t} \frac{y_s(i)}{\tau_s^2(i)} \frac{y_s(j)}{\tau_s(j)} \mathbf{g}_s(i) - \frac{t}{T} \sum_{s=1}^{T} \frac{y_s(i)}{\tau_s^2(i)} \frac{y_s(j)}{\tau_s(j)} \mathbf{g}_s(i) \right)^\top (\boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_T) \right|$$

$$= O_P(1) \frac{1}{T} \max_{1\le t\le T} \left\| \sum_{s=1}^{t} \frac{y_s(i)}{\tau_s^2(i)} \frac{y_s(j)}{\tau_s(j)} \mathbf{g}_s(i) - \frac{t}{T} \sum_{s=1}^{T} \frac{y_s(i)}{\tau_s^2(i)} \frac{y_s(j)}{\tau_s(j)} \mathbf{g}_s(i) \right\| = o_P(1)$$

since $E|y_0(i) y_0(j)| \|\mathbf{g}_0(i)\| < \infty$. Hence we obtain that

$$T^{-1/2} \max_{1\le t\le T} \left| \sum_{s=1}^{t} a_{s,7} - \frac{t}{T} \sum_{s=1}^{T} a_{s,7} \right| = o_P(1)$$

and by the same arguments

$$T^{-1/2} \max_{1\le t\le T} \left| \sum_{s=1}^{t} a_{s,8} - \frac{t}{T} \sum_{s=1}^{T} a_{s,8} \right| = o_P(1).$$

Thus we proved that

$$T^{-1/2} \max_{1\le t\le T} \left| \left(\sum_{s=1}^{t} \hat{y}_s(i) \hat{y}_s(j) - \frac{\lfloor t \rfloor}{T} \sum_{s=1}^{T} \hat{y}_s(i) \hat{y}_s(j) \right) \right.$$

$$\left. - \left(\sum_{s=1}^{t} y_s^*(i) y_s^*(j) - \frac{t}{T} \sum_{s=1}^{T} y_s^*(i) y_s^*(j) \right) \right| = o_P(1),$$

and therefore the result follows from Theorem 2.1. □

Supplementary Materials. The supplemental appendix contains the verifications of the conditions of the main theorem for our examples, additional examples and further Monte Carlo simulations.

Acknowledgements. We are grateful to Professors Christian Francq and Jean–Michel Zakoian for their comments on the first version of this paper and for useful references. We also thank the editor, Todd Clark, the associate editor and two anonymous referees, whose detailed and constructive comments helped to improve the quality of the paper.

References

[1] Aielli, G.P.: Dynamic conditional correlation: on properties and estimation. *Journal of Business and Economic Statistics* **31**(2013), 282–299.

[2] Andreou, E. and Ghysels, E.: Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics* **17**(2002), 579–600.

[3] Andreou, E. and Ghysels, E.: Tests for breaks in the conditional co-movements of asset returns. *Statistica Sinica* (2003), 1045–1073.

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

[5] Aue, A. and Horváth, L.: Structural breaks in time series. *Journal of Time Series Analysis* **34**(2013), 1–16.

[6] Bai, J. and Perron, P.: Estimating and testing linear models with multiple structural changes. *Econometrica* **66**(1998), 47–78.

[7] Bauwens, L., Laurent, S. and Rombouts, J.V.K.: Multivariate GARCH models: A survey.*Journal of Applied Econometrics* **21**(2006), 79-109.

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

[9] Bollerslev, T.: Modelling the coherence in short–run nominal exchange rates: a multivariate generalized ARCH model. *The Review of Economics and Statistics* (1990), 498–505.

[10] Boudt, K., Croux, C.: Robust M-estimation of multivariate GARCH models. *Computational Statistics & Data Analysis* (2010), 2459-2469.

[11] Boussama, F., Fuchs, F. and Stelzer, R.: Stationarity and geometric ergodicity of BEKK multivariate GARCH models. *Stochastic Processes and Their Applications* **121**(2011), 2331–2360.

[12] Cappiello, L, Engle, R.F. and Sheppard, K.: Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics* **4**(2006), 537–572.

[13] Comte, F. and Lieberman, O.: Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis* **84**(2003), 61-84.

[14] Csörgő, M. and Horváth, L.: *Limit Theorems in Change–point Analysis.* Wiley, 1997.

[15] Demetrescu, M. and Wied, D.: *Residual–based inference on moment hypotheses, with an application to testing for constant correlation.* Discussion paper, Nr. 37/2016, University of Dortmund.

[16] De Pooter, M. and Van Dijk, D.: Testing for changes in volatility in heteroskedastic time series–a further examination. *Technical report, Econometrics Institute Research Papers* (2004).

[17] Dimitriou, D., Kenourgios, D. and Simos, T.: Global financial crisis and emerging stock market contagion: a multivariate FIAPARCH–DCC approach. *International Review of Financial Analysis* **30** (2013), 46–56.

[18] Dungey, M., Milunovich, G., Thorp, S. and Yang, M.: Endogenous crisis dating and contagion using smooth transition structural GARCH. *Journal of Banking & Finance* **58** (2015), 71–79.

[19] Engle, R.F.: Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* **20**(2002), 339–350.

[20] Engle, R.F: *Anticipating Correlations: A New Paradigm for Risk Management.* Princeton University Press, 2009.

[21] Engle, R.F. and Kroner, K.F.: Multivariate simultaneous generalized ARCH. *Econometric Theory* **11**(1995), 122-150.

[22] Fermanian, J.-D. and Malongo, H.: On the stationarity of dynamical conditional correlation models. *Econometric Theory* **33**(2017), 636–663.

[23] Forbes, K.J. and Rigobon, R.: No contagion, only interdependence: measuring stock market comovements. *The journal of Finance* **57**(2002), 2223–2261.

[24] Francq, C. and Horváth, L. and Zakoian, J.M.: Variance targeting estimation of multivariate GARCH models. *Journal of Financial Econometrics* **14**(2016), 353–382.

[25] Francq, C. and Zakoian, J.M.: *GARCH Models.* Wiley, 2010.

[26] Fryzlewicz, P. and Rao, S.S.: Mixing properties of ARCH and time–varying ARCH processes, *Bernoulli* **17**(2011), 320–346.

[27] Hafner, C.M. and Preminger, A.: On asymptotic theory for multivariate GARCH models. *Journal of Multivariate Analysis* **100**(2009), 2044–2054.

[28] Horváth, L. and Rice, G.: Extensions of some classical methods in change point analysis (with discussions). *TEST* **23**(2014), 219–290.

[29] Inclan, C. and Tiao, G.C.: Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association* **89**(1994), 913–923.

[30] Kokoszka, P. and Leipus, R.: Change–point estimation in ARCH models *Bernoulli* **6**(2000), 513–539.

[31] Kulperger, R. and Yu, H.: High moment partial sum processes of residuals in GARCH models and their applications. *Annals of Statistics* **33**(2005), 2395–2422.

[32] Lavielle, M. and Teyssiére, G.: Detection of multiple change–points in multivariate time series. *Lithuanian Mathematical Journal* **46**(2006), 287–306.

[33] Lee, S., Na, O. and Na, S.: On the cusum of squares test for variance change in nonstationary and nonparametric time series models. *The Annals of the Institute of Statistical Mathematics* **55**(2003), 467–485.

[34] Ling, S.: Testing for change points in time series models and limiting theorems for NED sequences. *Annals of Statistics* **35**(2007), 1213–1237.

[35] Meyn, S.P. and Tweedie, R.L.: *Markov Chains and Stochastic Stability.* Springer, London, 1993.

[36] Pedersen, R.S. and Rahbek, A.: Multivariate variance targeting in the BEKK–GARCH model. *Econometrics Journal* **17**(2014), 24–55.

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

[38] Silvennoinen, A. and Teräsvirta, T.: Multivariate GARCH models. In: *Handbook of Financial Time Series* (Davis, R.A., Kreiss, J.-P. and Mikosch, T., Eds.), Springer, New York, 2009.

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

[40] Wu, W.B. and Zaffaroni, P.: Asymptotic theory for spectral density estimates of general multivariate time series. *Econometric Theory* **34**(2018), 1–22.

Marco Barassi, Department of Economics, University of Birmingham, Birmingham, B15 2TT UK

Lajos Horváth, Department of Mathematics, University of Utah, Salt Lake City, UT 84112–0090 USA

Yuqian Zhao, Department of Statistics and Actuarial Science, University of Waterloo, N2L 3G1, Canada