



**HAL**  
open science

# Tests of the Constancy of Conditional Correlations of Unknown Functional Form in Multivariate GARCH Models

Anne Peguin-Feissolle, Bilel Sanhaji

► **To cite this version:**

Anne Peguin-Feissolle, Bilel Sanhaji. Tests of the Constancy of Conditional Correlations of Unknown Functional Form in Multivariate GARCH Models. *Annals of Economics and Statistics*, 2016, 123/124, pp.77. 10.15609/annaeconstat2009.123-124.0077 . hal-04218472

**HAL Id: hal-04218472**

**<https://hal.science/hal-04218472>**

Submitted on 27 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GENES

ADRES

---

Tests of the Constancy of Conditional Correlations of Unknown Functional Form in Multivariate GARCH Models

Author(s): Anne Péguin-Feissolle and Bilel Sanhaji

Source: *Annals of Economics and Statistics*, No. 123/124, SPECIAL ISSUE ON RECENT DEVELOPMENTS IN FINANCIAL ECONOMETRICS (December 2016), pp. 77-101

Published by: GENES on behalf of ADRES

Stable URL: <https://www.jstor.org/stable/10.15609/annaconstat2009.123-124.0077>

**REFERENCES**

Linked references are available on JSTOR for this article:

[https://www.jstor.org/stable/10.15609/annaconstat2009.123-124.0077?seq=1&cid=pdf-reference#references\\_tab\\_contents](https://www.jstor.org/stable/10.15609/annaconstat2009.123-124.0077?seq=1&cid=pdf-reference#references_tab_contents)

You may need to log in to JSTOR to access the linked references.

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



GENES and ADRES are collaborating with JSTOR to digitize, preserve and extend access to *Annals of Economics and Statistics*

JSTOR

# Tests of the Constancy of Conditional Correlations of Unknown Functional Form in Multivariate GARCH Models

Anne PÉGUIN-FEISSOLLE

*Aix-Marseille School of Economics,  
CNRS and EHESS, France*

Bilel SANHAJI

*Université Paris 8, LED, Paris, France*

We introduce two tests for the constancy of conditional correlations of unknown functional form in multivariate GARCH models. The first test is based on artificial neural networks and the second on a Taylor expansion of each unknown conditional correlation. They can be seen as general misspecification tests for a large set of multivariate GARCH-type models. We investigate their size and their power through Monte Carlo experiments. Moreover, we study the robustness of these tests to nonnormality by simulating some models, such as the *GARCH - t* and *Beta - t - EGARCH*. We give some illustrative empirical examples based on financial data.\*

## I. Introduction

During the last decade, there has been a rapid increase in the amount of literature on theoretical and empirical developments of multivariate GARCH-type modeling (see, for example, the extensive surveys in BAUWENS, L., S. LAURENT, and J. V. K. ROMBOUITS [2006] and SILVENNOINEN, A., and T. TERÄSVIRTA [2009a]). This is an important econometric modeling issue, particularly in finance where, for instance, the correlation structure among different national stock returns is one of the most useful instruments to evaluate the gains from international portfolio diversification. Estimating multivariate GARCH (MGARCH) models is time-consuming, therefore introducing new tests specific to multivariate models to check *ex ante* the properties of the data is important. For example, in the presence of constant conditional correlations some parameters of multivariate GARCH-type models might not be identified, leading to estimation issues which could have been avoided by pretesting the constancy of the conditional correlations.

Most of the tests for the constancy of conditional correlations are based on the Lagrange Multiplier (LM) procedure. As noted by TSE, Y. K. [2000], computing the LM test statistics requires only the estimation of the constant correlation model and is thus generally computationally convenient (see, among others, the LM constant conditional correlation tests of TSE, Y. K. [2000]; BERA, A. K., and S. KIM [2002]; BERBEN, R. P., and W. J. JANSEN [2005];

\*JEL: C22, C45, C58 / KEY WORDS: Multivariate GARCH, Neural Network, Taylor Expansion.

SILVENNOINEN, A., and T. TERÄSVIRTA [2009b] and SILVENNOINEN, A., and T. TERÄSVIRTA [2015]; HARVEY, A., and S. THIELE [2014]).

In this paper, we introduce two LM tests for constancy of the conditional correlations. The first test is based on artificial neural networks (ANN). This ANN-based test relies on a statistical technique proposed by LEE, T. H., H. WHITE, and C. W. J. GRANGER [1993]. The ANN framework has already been used for some tests (see, for example, LEE, T. H., H. WHITE, and C. W. J. GRANGER [1993] and TERÄSVIRTA, T., C. F. LIN, and C. W. J. GRANGER [1993] for linearity tests, KAMSTRA, M. [1993], CAULET, R., and A. PÉGUIN-FEISSOLLE [2000] and PÉGUIN-FEISSOLLE, A. [1999] for conditional heteroscedasticity tests, LEBRETON, M., and A. PÉGUIN-FEISSOLLE [2007] for heteroscedasticity tests and PÉGUIN-FEISSOLLE, A., and T. TERÄSVIRTA [1999] for causality tests).

The second test is based on the linearization by Taylor expansion of each unknown conditional correlation around a given point in a sample space. This approach has already been applied to test causality (PÉGUIN-FEISSOLLE, A., and T. TERÄSVIRTA [1999]; PÉGUIN-FEISSOLLE, A., B. STRIKHOLM, and T. TERÄSVIRTA [2013]), heteroscedasticity (LEBRETON, M., and A. PÉGUIN-FEISSOLLE [2007]) and conditional heteroscedasticity (CAULET, R., and A. PÉGUIN-FEISSOLLE [2000]; PÉGUIN-FEISSOLLE, A. [1999]).

Both tests present similar fundamental characteristics; they require little knowledge of the functional relationship determining the correlations; they are easy to implement and perform well in our small-sample simulations; and they generalize well to a high number of endogenous series. Both the ANN and Taylor expansion permit the approximation of unknown relationships, therefore a rejection of the null hypothesis of constancy of conditional correlations does not imply that the data have been generated from a model where the conditional correlations are specified as neural functions or specific functions. These tests can be seen as general misspecification tests of very different multivariate GARCH-type models. It is also worth noting that the tests presented here can be applied easily to test partially constant correlations.

Finite-sample properties of these two new tests are examined using Monte Carlo methods by comparing them to two alternative conditional correlation tests: those of TSE, Y. K. [2000] and SILVENNOINEN, A., and T. TERÄSVIRTA [2015]. Using a variety of different specifications for the MGARCH model, we show that they perform well in small-sample simulations, even in the nonnormality case. Empirical applications show that these tests reject the constant correlation hypothesis when the tests of both TSE, Y. K. [2000] and SILVENNOINEN, A., and T. TERÄSVIRTA [2015] do not.

The paper is organized as follows. In SECTION II, we introduce the tests. SECTION III reports the results of the Monte Carlo simulation study. In SECTION IV, we study the robustness of the tests to nonnormality. SECTION V describes some illustrative examples based on financial data. SECTION VI contains some conclusions.

## II. Testing the Constancy of Conditional Correlation

We consider the following general MGARCH model:

$$\begin{aligned} y_t &= E[y_t | \Omega_{t-1}] + \varepsilon_t \\ \text{Var}[y_t | \Omega_{t-1}] &= H_t \end{aligned} \quad t = 1, \dots, T. \quad (1)$$

where  $y_t$  is a  $N$ -dimensional vector process,  $E[y_t | \Omega_{t-1}]$  and  $Var[y_t | \Omega_{t-1}]$  are respectively the conditional expectation and the conditional covariance of  $y_t$  with respect to  $\Omega_{t-1}$ , the sigma-field generated by all the information until time  $t - 1$ . The process  $y_t$  is strictly stationary and ergodic. To simplify the discussion, we assume that  $E_{t-1}[y_t] = E[y_t | \Omega_{t-1}] = 0$  or  $y_t = \varepsilon_t$  where  $\varepsilon_t$  is defined by  $\varepsilon_t = H_t^{1/2}\eta_t$  with  $\eta_t \sim iid(0, I_N)$ . Note that the tests presented in this paper can be generalized easily to a large variety of time-varying structures of the conditional expectation of  $y_t$ . We assume that the conditional variances follow a *GARCH*(1,1) process:

$$h_{iit} = \zeta_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{iit-1} \quad i = 1, \dots, N \quad (2)$$

where the standard positivity and covariance stationarity constraints are imposed, i.e.  $\zeta_i > 0$ ,  $\alpha_i > 0$ ,  $\beta_i \geq 0$ , and  $\alpha_i + \beta_i < 1$  for  $i = 1, \dots, N$ . Let  $P_t = (\rho_{ijt})_{i,j=1,\dots,N}$  be the  $N \times N$  conditional correlation matrix for the  $\varepsilon_t$ ; we can write  $H_t = S_t P_t S_t$  where  $S_t$  is the  $N \times N$  matrix given by  $S_t = \text{diag}(\sqrt{h_{11t}}, \dots, \sqrt{h_{NNt}})$ . We assume that the conditional correlation matrix  $P_t$  is positive definite at each  $t$ ; this assumption guarantees the positive definiteness of  $H_t$ . Moreover, the conditional normality of  $\varepsilon_t$ ,  $\varepsilon_t | \Omega_{t-1} \sim N(0, H_t)$ , implies that  $z_t | \Omega_{t-1} \sim N(0, P_t)$  with  $z_t = S_t^{-1} \varepsilon_t$ .

### II.1. The ANN-Based Test

The first test we present is an ANN based LM-type test. In order to construct the statistic, we extend the ideas of the tests of constant correlation hypothesis from TSE, Y. K. [2000], SILVENNOINEN, A., and T. TERÄSVIRTA [2015], SILVENNOINEN, A., and T. TERÄSVIRTA [2009b] and BERBEN, R. P., and W. J. JANSEN [2005] to a model where the conditional correlation  $\rho_{ijt}$  is a neural function. ANN models are useful in situations where the functional form of a potential relationship between two variables is not known in advance. As noted in PÉGUIN-FEISSOLLE, A. [1999], the use of ANN models in the present context is validated by the universal mapping theorem, which states that, under mild regularity conditions, ANN models provide arbitrarily accurate approximations to nonlinear mappings (see HORNIK, K., M. STINCHCOMBE, and H. WHITE [1989] and HORNIK, K., M. STINCHCOMBE, and H. WHITE [1990]; HORNIK, K. [1991]; STINCHCOMBE, M., and H. WHITE [1989]; CYBENKO, G. [1989]; CARROLL, S. M., and B. W. DICKINSON [1989], among others).

We specify the time-varying structure of the conditional correlations as follows: we assume that the conditional correlations  $\rho_{ijt}$  are changing smoothly over time according to a neural function:

$$\rho_{ijt} = \rho_{ij} + \sum_{m=1}^p \delta_{ijm} (1 + \exp\{-w'_{ijt} \gamma_{ijm}\})^{-1} \quad (3)$$

where  $1 \leq i < j \leq N$ ,  $p < \infty$ ,  $w_{ijt}$  and  $\gamma_{ijm}$  are  $(2q + 1) \times 1$  vectors.  $w_{ijt}$  is given by

$$w_{ijt} = (1, \tilde{w}'_{ijt})' = (1, \varepsilon_{i,t-1}, \dots, \varepsilon_{i,t-q}, \varepsilon_{j,t-1}, \dots, \varepsilon_{j,t-q})'. \quad (4)$$

The conditional correlations  $\rho_{ijt}$  should satisfy  $|\rho_{ijt}| \leq 1$ ,  $1 \leq i < j \leq N$  and  $t = 1, \dots, T$ , and the corresponding correlation matrix should be positive semidefinite. However, we need only to assume that these restrictions are satisfied in the neighborhood of the null hypothesis of constant conditional correlation, i.e.  $\rho_{ijt} = \rho_{ij}$ , as noted by TSE, Y. K. [2000], p. 111.

For each pair of indices  $(i, j)$ , the  $2q + 1$  input units of the network send signals, amplified or attenuated by weighting factors  $\gamma_{ijm}$ , to  $p$  hidden units (or hidden nodes) that sum up the signals and generate a squashing function, which is assumed to be a logistic function. As previously mentioned, neural functions of the form (3) accurately approximate arbitrary functions, given a sufficiently large number  $p$  of hidden units and a suitable choice of the parameters  $\delta_{ijm}$  and  $\gamma_{ijm}$ .

Given (3), the null hypothesis of constant conditional correlation, i.e.  $\rho_{ijt} = \rho_{ij}$ , can be formulated as:

$$H_{01} : \delta_{ijm} = 0, 1 \leq i < j \leq N, 1 \leq m \leq p. \tag{5}$$

Under  $H_{01}$ ,  $\gamma_{ijm}$ , for  $i = 1, \dots, N - 1, j = i + 1, \dots, N$  and  $m = 1, \dots, p$ , are not identified. For this reason, conventional maximum likelihood theory is not applicable to derive the test procedure, as noted in LEBRETON, M., and A. PÉGUIN-FEISSOLLE [2007]. To solve this problem, we use the method given by LEE, T. H., H. WHITE, and C. W. J. GRANGER [1993], i.e. we generate the hidden unit weights  $\gamma_{ijm}$  randomly from the uniform distribution on  $[-\mu, \mu]$  with  $\mu = 2$ .

The null hypothesis can be tested using LM procedure. Let  $\theta$  be the vector of the parameters in the model, i.e. the conditional variances  $\omega_i = (\zeta_i, \alpha_i, \beta_i)'$ ,  $i = 1, \dots, N$ , the conditional correlations  $\rho_{ij}$  with  $1 \leq i < j \leq N$  and the parameters of the neural function,  $\delta_{ijm}$  for  $1 \leq i < j \leq N, 1 \leq m \leq p$  and  $p < \infty$ . Under standard regularity conditions, the test statistic, denoted by *NEURAL*, is given by:

$$\frac{1}{T} \left( \sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \theta} \right)' \mathfrak{I}(\theta)^{-1} \left( \sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \theta} \right) \tag{6}$$

where  $\mathfrak{I}(\theta) = \frac{1}{T} \sum_{t=1}^T E_{t-1} \left[ \frac{\partial l_t(\theta)}{\partial \theta} \frac{\partial l_t(\theta)}{\partial \theta'} \right]$  and  $l_t(\theta)$  is the log-likelihood for the observation  $t$ :

$$l_t(\theta) = -\frac{N}{2} \ln(2\pi) - \frac{1}{2} \ln(|P_t|) - \frac{1}{2} \sum_{i=1}^N \ln(h_{iit}) - \frac{1}{2} z_t' P_t^{-1} z_t. \tag{7}$$

The *NEURAL* statistic has an asymptotic  $\chi^2$  distribution under the null with  $p \frac{N(N-1)}{2}$  degrees of freedom (see the Appendix for details of the technical derivations of the test statistic, and PÉGUIN-FEISSOLLE, A., and B. SANHAJI [2015]).

As LEE, T. H., H. WHITE, and C. W. J. GRANGER [1993] and PÉGUIN-FEISSOLLE, A., and T. TERÄSVIRTA [1999] point out, the elements of some matrices used to build the ANN-based test statistic can lead to collinearity problems when the number of hidden units  $p$  is large. Therefore, in our case, the  $p \times 1$  vectors  $g_{ijt}$  defined as:

$$g_{ijt} = \begin{pmatrix} g_{ijt,1} \\ \vdots \\ g_{ijt,p} \end{pmatrix} = \begin{pmatrix} (1 + \exp\{-w'_{ijt} \gamma_{ij1}\})^{-1} \\ \vdots \\ (1 + \exp\{-w'_{ijt} \gamma_{ijp}\})^{-1} \end{pmatrix} \quad 1 \leq i < j \leq N, p < \infty, \tag{8}$$

tend to be collinear among themselves, especially when  $p$  is large. The conditional correlations  $\rho_{ijt}$  given by (3) can be written as

$$\rho_{ijt} = \rho_{ij} + \sum_{m=1}^p \delta_{ijm} g_{ijt,m}.$$

Let us define  $G_{ij}$  the  $p \times T$  matrix given by:  $G_{ij} = (g_{ij1}, \dots, g_{ijT})$ . To solve the collinearity problem, we conduct the test using the main principal components of each  $G_{ij}$  matrix. We use only the largest principal components explaining together at least 90% of the variation in this matrix (see PÉGUIN-FEISSOLLE, A., and T. TERÄSVIRTA [1999]; LEBRETON, M., and A. PÉGUIN-FEISSOLLE [2007]; PÉGUIN-FEISSOLLE, A., B. STRIKHOLM, and T. TERÄSVIRTA [2013]; see also CASTLE, J. L., and D. F. HENDRY [2010] for discussions on using principal components to solve the collinearity problem). More precisely, instead of  $G_{ij}$ , we build the  $p_{ij}^* \times T$  matrix  $G_{ij}^*$  of the  $p_{ij}^*$  principal components chosen according to the preceding rule, i.e.  $G_{ij}^* = (g_{ij1}^*, \dots, g_{ijT}^*)$ . The null hypothesis is now that the parameters associated with the main principal components are equal to zero in the following specification of the conditional correlations:

$$\rho_{ijt} = \rho_{ij} + \sum_{m=1}^{p_{ij}^*} \delta_{ijm} g_{ijt,m}^*.$$

The *NEURAL* statistic will have in this case an asymptotic  $\chi^2$  distribution under the null with  $\sum_{i=1}^{N-1} \sum_{j=i+1}^N p_{ij}^*$  degrees of freedom.

## II.2. The Taylor Expansion-Based Test

We assume that the functional form  $f_{ij}$  determining  $\rho_{ijt}$ , for  $i, j = 1, \dots, N$  and  $t = 1, \dots, T$ , is unknown and is adequately represented by the following equation:

$$\rho_{ijt} = f_{ij}(\tilde{w}_{ijt}, \theta_{ij}^*) \quad (9)$$

where, for each pair of indices  $(i, j)$ ,  $\tilde{w}_{ijt}$  is a  $2q \times 1$  vector given by (4) and  $\theta_{ij}^*$  is a  $r_{\theta_{ij}^*} \times 1$  unknown parameter vector. The conditional correlations  $\rho_{ijt}$  should fulfill the conditions  $|\rho_{ijt}| \leq 1$ ,  $1 \leq i < j \leq N$  and  $t = 1, \dots, T$ , and the corresponding correlation matrix has to be positive semidefinite; but, as in the ANN-based test, we will simply assume that these restrictions are verified in the neighborhood of the null hypothesis.

The test is based on a finite-order Taylor expansion. Following PÉGUIN-FEISSOLLE, A., and T. TERÄSVIRTA [1999] and PÉGUIN-FEISSOLLE, A., B. STRIKHOLM, and T. TERÄSVIRTA [2013], we assume that all the functions  $f_{ij}$  have a convergent Taylor expansion at any arbitrary point of the sample space for every  $\theta_{ij}^* \in \Theta_{ij}$  (the parameter spaces). When the order of the Taylor expansion increases, the remainder of the Taylor expansion converges to zero, for each pair  $(i, j)$ .

In order to linearize  $f_{ij}$  in (9), we expand the function into a  $k$ th-order Taylor series around an arbitrary fixed point in the sample space. After approximating  $f_{ij}$ , merging terms and reparametrizing, we obtain:

$$\begin{aligned}
 \rho_{ijt} = & \rho_{ij} + \sum_{m=1}^q \lambda_m \varepsilon_{i,t-m} + \sum_{m=1}^q \phi_m \varepsilon_{j,t-m} + \sum_{m_1=1}^q \sum_{m_2=m_1}^q \lambda_{m_1 m_2} \varepsilon_{i,t-m_1} \varepsilon_{i,t-m_2} \\
 & + \sum_{m_1=1}^q \sum_{m_2=1}^q \psi_{m_1 m_2} \varepsilon_{i,t-m_1} \varepsilon_{j,t-m_2} + \sum_{m_1=1}^q \sum_{m_2=m_1}^q \phi_{m_1 m_2} \varepsilon_{j,t-m_1} \varepsilon_{j,t-m_2} \\
 & + \dots + \sum_{m_1=1}^q \sum_{m_2=m_1}^q \dots \sum_{m_k=m_{k-1}}^q \lambda_{m_1 \dots m_k} \varepsilon_{i,t-m_1} \dots \varepsilon_{i,t-m_k} \\
 & + \dots + \sum_{m_1=1}^q \sum_{m_2=m_1}^q \dots \sum_{m_k=m_{k-1}}^q \phi_{m_1 \dots m_k} \varepsilon_{j,t-m_1} \dots \varepsilon_{j,t-m_k} + R_t^{(k)}
 \end{aligned} \tag{10}$$

where  $R_t^{(k)}$  is the remainder. In expansion (10), we find all possible combinations of lagged values of  $\varepsilon_{i,t}$  and  $\varepsilon_{j,t}$  up to order  $q$ .

The assumption that the conditional correlation  $\rho_{ijt}$  is constant means that all terms involving functions of elements of lagged values of  $\varepsilon_{i,t}$  and  $\varepsilon_{j,t}$  in (10) must have zero coefficients, i.e. all the parameters except  $\rho_{ij}$  are equal to zero. Therefore, the null hypothesis of interest is:

$$H_{02} : \begin{cases} \lambda_m = 0, m = 1, \dots, q \\ \phi_m = 0, m = 1, \dots, q \\ \lambda_{m_1 m_2} = 0, m_1 = 1, \dots, q, m_2 = m_1, \dots, q \\ \psi_{m_1 m_2} = 0, m_1 = 1, \dots, q, m_2 = 1, \dots, q \\ \phi_{m_1 m_2} = 0, m_1 = 1, \dots, q, m_2 = m_1, \dots, q \\ \vdots \\ \phi_{m_1 \dots m_k} = 0, m_1 = 1, \dots, q, m_2 = m_1, \dots, q, \dots, m_k = m_{k-1}, \dots, q. \end{cases}$$

The number of parameters to be tested under the null hypothesis is:

$$N^* = 2 \left[ \sum_{r=1}^k \binom{2q+r-1}{r} - \sum_{r=1}^k \binom{q+r-1}{r} \right].$$

The test is based on the LM procedure as for the *NEURAL* statistic. Under standard regularity conditions, the test statistic, denoted *TAYLOR*, is given by formula (6) where  $\theta$  is the vector of the parameters in the model, i.e. the conditional variances  $\omega_i = (\zeta_i, \alpha_i, \beta_i)'$  for  $i = 1, \dots, N$ ,  $\rho_{ij}$  with  $1 \leq i < j \leq N$  and the  $N^* \times 1$  parameters equal to zero under the null hypothesis. If the remainder  $R_t^{(k)} \equiv 0$ , the *TAYLOR* statistic has an asymptotic  $\chi^2$  distribution with  $N^*$  degrees of freedom under the null hypothesis (see the Appendix for details of the technical derivations of the test statistic).

Following PÉGUIN-FEISSOLLE, A., B. STRIKHOLM, and T. TERÄSVIRTA [2013], there are two practical difficulties when the order of the Taylor expansion  $k$  increases: the regressors tend to be highly collinear and the dimension of the null hypothesis may become large because the number of regressors increases rapidly with  $k$ . More precisely, the conditional correlations  $\rho_{ijt}$  given by (10) can be written as (assuming  $R_t^{(k)} \equiv 0$ )



$$\rho_{ijt} = \rho_{ij} + \sum_{m=1}^{N^*} \delta_{ijm} d_{ijt,m}$$

where the  $N^* \times 1$  vectors  $d_{ijt}$  correspond to terms depending on lagged values of  $\epsilon_{i,t}$  and  $\epsilon_{j,t}$ . We replace the matrix  $D_{ij} = (d_{ij1}, \dots, d_{ijT})$  for  $1 \leq i < j \leq N$  by its largest principal components. As in the case of the ANN-based test, only the largest principal components explaining together at least 90% of the variation in the matrix are used. Instead of  $D_{ij}$  we build the  $N_{ij}^* \times T$  matrix  $D_{ij}^*$  of the  $N_{ij}^*$  chosen principal components:  $D_{ij}^* = (d_{ij1}^*, \dots, d_{ijT}^*)$ . The null hypothesis is that the parameters associated with the main principal components are equal to zero in the relationship determining  $\rho_{ijt}$ :

$$\rho_{ijt} = \rho_{ij} + \sum_{m=1}^{N_{ij}^*} \delta_{ijm} d_{ijt,m}^*$$

Under the null, the *TAYLOR* statistic follows an asymptotic  $\chi^2$  distribution with  $\sum_{i=1}^{N-1} \sum_{j=i+1}^N N_{ij}^*$  degrees of freedom.

### III. Monte Carlo Experiments

In this SECTION, we investigate the small-sample performances of the two new tests. Using Monte Carlo experiments, we compare both the sizes and the powers of these tests to the tests of TSE, Y. K. [2000] and SILVENNOINEN, A., and T. TERÄSVIRTA [2015] whose statistics are denoted *TSE* and *STCC*, respectively (see TSE, Y. K. [2000] and SILVENNOINEN, A., and T. TERÄSVIRTA [2015] for further details).

The data generating process (DGP) is based on the general multivariate GARCH model introduced earlier where the  $N \times 1$  vector of residuals is given by  $\epsilon_t = H_t^{1/2} \eta_t$  with  $\eta_t \sim \text{nid}(0, I_N)$ ,  $t = 1, \dots, T$ . We generate several MGARCH-type models commonly used in financial time series analysis under the normality assumption and relax this assumption in SECTION IV. The individual GARCH models considered in the simulations have to be (weakly) stationary.

For all the Monte Carlo simulations, we consider nominal sizes of 1%, 5% and 10%, sample sizes  $T = 1000$ , 1500 and 2500 and  $N = 2$  endogenous variables (see p. 27 and TABLE VI of PÉGUIN-FEISSOLLE, A., and B. SANHAJI [2015] for  $N > 2$ ). Moreover, we remove the first observations in order to eliminate initialization effects. The number of replications is  $S = 2000$ , and in every replication the true values of the parameters are used as starting values for the iterations so that the rate of convergence is fast. For the ANN-based tests, following LEE, T. H., H. WHITE, and C. W. J. GRANGER [1993], the number of hidden units is  $p = 20$ . Moreover, for both new tests, we use the largest principal components-based statistics presented above, the number of principal components being determined automatically.

#### III.1. Size Simulations

Based on simulations performed with different models (see p. 21 and TABLES I and II of PÉGUIN-FEISSOLLE, A., and B. SANHAJI [2015], for details) we choose  $p = 20$  hidden units and

$q = 3$  lagged residuals for the *NEURAL* test statistic, and  $k = 3$  for the order of the Taylor expansion and  $q = 2$  lagged residuals for the *TAYLOR* test statistic. Following SILVENOINEN, A., and T. TERÄSVIRTA [2015], the transition variable for the *STCC* test is generated from an exogenous *GARCH*(1,1) process such that  $s_t = h_t^{1/2} z_t$  where  $z_t \sim N(0,1)$  and  $h_t = 0.02 + 0.03s_{t-1}^2 + 0.94h_{t-1}$ .

We consider four models with constant conditional correlations. The first is an extended constant conditional correlations *GARCH* model (*ECCC – GARCH*) defined by JEANTHEAU, T. [1998]. The individual conditional variance equations depend on the past squared returns and conditional variances of all the series. Therefore, the elements of the conditional variance matrix are characterized for  $N = 2$  by:

$$= 3pt \begin{pmatrix} h_{1t} \\ h_{2t} \end{pmatrix} = \begin{pmatrix} \alpha_{10} \\ \alpha_{20} \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{2,t-1}^2 \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} h_{1,t-1} \\ h_{2,t-1} \end{pmatrix} \tag{11}$$

and  $h_{12t} = \rho_{12} \sqrt{h_{1t} h_{2t}}$  where  $\rho_{12}$  is the constant conditional correlation. This model will be denoted by *ECCC*. In the second model, called *CCC*, the off-diagonal parameters in (11) are equal to zero i.e.  $\alpha_{12} = \alpha_{21} = 0$  and  $\beta_{12} = \beta_{21} = 0$ . The third model, called *GARCH*, will be composed of two univariate independent *GARCH*(1,1) models, i.e.  $\alpha_{12} = \alpha_{21} = 0$ ,  $\beta_{12} = \beta_{21} = 0$  and  $\rho_{12} = 0$  in (11). We also design an asymmetric model where (11) is

completed by leverage terms  $\begin{pmatrix} \gamma_{11} & 0 \\ 0 & \gamma_{22} \end{pmatrix} \begin{pmatrix} v_{1,t-1}^2 \\ v_{2,t-1}^2 \end{pmatrix}$  where  $v_{i,t} = I(\varepsilon_{i,t} < 0)\varepsilon_{i,t}$  for  $i = 1, \dots, N$

and  $I(\cdot)$  is the indicator function. This fourth model, denoted *GJR*, is composed of two univariate independent *GJR – GARCH* models, i.e.  $\alpha_{12} = \alpha_{21} = 0$ ,  $\beta_{12} = \beta_{21} = 0$  and  $\rho_{12} = 0$ .

TABLE I shows the empirical sizes of the different tests, i.e. the rejection probabilities under the null hypothesis of constant conditional correlations. We consider models with different constant correlations:  $\rho_{12} = 0.60$  for the *ECCC – GARCH* model,  $\rho_{12} = 0.30$  for the *CCC – GARCH* model, and  $\rho_{12} = 0$  for the bivariate *GARCH* and *GJR – GARCH* models. The *DGP* models correspond to different persistences, for example  $\alpha_{11} + \beta_{11} = \alpha_{22} + \beta_{22} = 0.99$  for the *CCC – GARCH* model, and  $\alpha_{11} + \beta_{11} = 0.95$  and  $\alpha_{22} + \beta_{22} = 0.90$  in the bivariate *GARCH* model.

We observe that the empirical sizes converge towards the nominal sizes when the number of observations  $T$  increases, for most of the tests. More specifically, in the *ECCC*, *GARCH* and *GJR* cases, for the different sample sizes, the *NEURAL*, *TAYLOR* and *STCC* statistics show similar performances and clearly outperform the *TSE* test, highlighting their good size properties by showing weak distortions. In the *CCC* case, the empirical sizes are close to the nominal sizes for all the test statistics when the sample size reaches  $T = 2500$ . Concerning the *ECCC* and *GJR – GARCH* models, the conditional variances are not exactly *GARCH*(1,1). Nevertheless, we observe that the performances of the tests are similar for all the *DGPs*.

### III.2. Power Simulations

To evaluate the power of the tests, we generate four different time-varying conditional correlations multivariate *GARCH*-type models, chosen to represent a variety of situations.

TABLE I. – Small sample sizes of the different constant conditional correlation tests

<i>T</i>	1,000			1,500			2,500		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
<i>ECCC</i>									
<i>NEURAL</i>	0.008	0.052	0.087	0.009	0.049	0.085	0.011	0.054	0.105
<i>TAYLOR</i>	0.012	0.042	0.083	0.014	0.053	0.097	0.009	0.053	0.104
<i>TSE</i>	0.027	0.084	0.134	0.023	0.080	0.139	0.086	0.144	0.201
<i>STCC</i>	0.011	0.054	0.107	0.012	0.042	0.094	0.011	0.049	0.104
<i>CCC</i>									
<i>NEURAL</i>	0.018	0.077	0.134	0.017	0.061	0.118	0.008	0.056	0.104
<i>TAYLOR</i>	0.021	0.083	0.141	0.017	0.070	0.133	0.010	0.046	0.105
<i>TSE</i>	0.018	0.060	0.117	0.015	0.062	0.107	0.017	0.052	0.105
<i>STCC</i>	0.015	0.059	0.112	0.013	0.059	0.106	0.008	0.047	0.098
<i>GARCH</i>									
<i>NEURAL</i>	0.014	0.053	0.107	0.015	0.057	0.113	0.011	0.055	0.106
<i>TAYLOR</i>	0.013	0.053	0.105	0.015	0.056	0.103	0.013	0.058	0.111
<i>TSE</i>	0.093	0.193	0.264	0.089	0.189	0.265	0.089	0.192	0.264
<i>STCC</i>	0.006	0.045	0.101	0.011	0.046	0.095	0.008	0.054	0.105
<i>GJR</i>									
<i>NEURAL</i>	0.013	0.061	0.119	0.008	0.052	0.099	0.010	0.061	0.109
<i>TAYLOR</i>	0.015	0.065	0.123	0.010	0.043	0.103	0.014	0.052	0.108
<i>TSE</i>	0.033	0.099	0.153	0.030	0.088	0.148	0.028	0.091	0.151
<i>STCC</i>	0.008	0.043	0.082	0.007	0.050	0.104	0.008	0.049	0.098

Note: The number of replications is  $S = 2000$ . For *NEURAL* the number of hidden units is  $p = 20$  and the number of lags of residuals is  $q = 3$ , and for *TAYLOR* the order of the Taylor expansion is  $k = 3$  and the number of lags of residuals is  $q = 2$ . The data-generating parameters are:

	$\zeta_1$	$\zeta_2$	$\alpha_{11}$	$\alpha_{12}$	$\alpha_{21}$	$\alpha_{22}$	$\beta_{11}$	$\beta_{12}$	$\beta_{21}$	$\beta_{22}$	$\gamma_{11}$	$\gamma_{22}$	$\rho_{12}$
ECCC	0.02	0.02	0.01	0.02	0.02	0.01	0.85	0.02	0.02	0.85	–	–	0.60
CCC	0.02	0.02	0.01	0.00	0.00	0.01	0.98	0.00	0.00	0.98	–	–	0.30
GARCH	0.40	0.20	0.15	0.00	0.00	0.20	0.80	0.00	0.00	0.70	–	–	0.00
GJR	0.02	0.01	0.05	0.00	0.00	0.04	0.91	0.00	0.00	0.93	0.02	0.01	0.00

The *BEKK*(1,1,1) model (ENGLE, R. F., and F. KRONER [1995]) is defined by:

$$H_t = Z'Z + A'\epsilon_{t-1}\epsilon'_{t-1}A + B'H_{t-1}B \quad (12)$$

where  $Z$ ,  $A$  and  $B$  are  $N \times N$  matrices and  $Z$  is upper triangular. With  $N = 2$ , we have:

$$Z = \begin{pmatrix} \zeta_{11} & \zeta_{12} \\ 0 & \zeta_{22} \end{pmatrix}, A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \text{ and } B = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}. \quad (13)$$

Note that the conditional covariance matrices  $H_t$  are positive definite by construction. This model will be denoted by *BEKK*.

An asymmetric scalar *BEKK* model is also tested (DING, Z., and R. F. ENGLE [2001]). This model is a special case of the *BEKK* model (13) and will be denoted by *ASBEKK*:

$$H_t = Z'Z + \alpha\epsilon_{t-1}\epsilon'_{t-1} + \gamma\nu_{t-1}\nu'_{t-1} + \beta H_{t-1}, \quad (14)$$

where  $\alpha, \gamma$  and  $\beta$  are scalars,  $v_t = I(\epsilon_t < 0) \odot \epsilon_t$  and  $\odot$  is the Hadamard product.

The third model is ENGLE, R. F. [2002] *DCC – GARCH* model (denoted by *DCC*):  $H_t$  is defined by  $H_t = S_t P_t S_t$  with  $P_t$  the  $N \times N$  conditional correlation matrix and  $S_t$  is the  $N \times N$  matrix given by  $S_t = \text{diag}(\sqrt{h_{11t}}, \dots, \sqrt{h_{NNt}})$ , with

$$P_t = \text{diag}(q_{11,t}^{-1/2}, \dots, q_{NN,t}^{-1/2}) Q_t \text{diag}(q_{11,t}^{-1/2}, \dots, q_{NN,t}^{-1/2}) \tag{15}$$

where  $Q_t = (q_{ij,t})$  is a symmetric positive definite  $N \times N$  matrix given by

$$Q_t = (1 - \alpha - \beta) \bar{Q} + \alpha z_{t-1} z'_{t-1} + \beta Q_{t-1}; \tag{16}$$

$\bar{Q}$  is a  $N \times N$  constant and positive definite matrix (see AIELLI, G. P. [2013] for more details on  $\bar{Q}$ ),  $z_t = S_t^{-1} \epsilon_t$  and  $\alpha$  and  $\beta$  are nonnegative scalar parameters satisfying  $\alpha + \beta < 1$ . Moreover,  $h_{iit}$  follows a simple *GARCH(1,1)* model for  $i = 1, \dots, N$ .

The last model is the *EDCC – GARCH* model (denoted by *EDCC*) which is the extended specification of the *DCC – GARCH* model, allowing for volatility spillovers; i.e. the individual conditional variance equations depend on the past squared returns and variances of all the series (see equation (11)). In the case  $N = 2$ , we have for  $i, j = 1, 2$ :

$$h_{iit} = \zeta_{ii} + \alpha_{ii} \epsilon_{i,t-1}^2 + \alpha_{ij} \epsilon_{j,t-1}^2 + \beta_{ii} h_{iit-1} + \beta_{ij} h_{jj,t-1}. \tag{17}$$

We take different DGPs for three of these models (*BEKK<sub>1</sub>* and *BEKK<sub>2</sub>*, *DCC<sub>1</sub>* and *DCC<sub>2</sub>*, *EDCC<sub>1</sub>* and *EDCC<sub>2</sub>*), characterized by different data-generating parameters and different variabilities of conditional correlation coefficients. The performance of the *STCC* test depends on the choice of the transition variable, but following SILVENNOINEN, A., and T. TERÄSVIRTA [2015], we define it as a linear combination of lags of squared returns:  $s_t = (0.2, 0.2, 0.2, 0.2, 0.2) \times (\bar{\epsilon}_{t-1}^2, \bar{\epsilon}_{t-2}^2, \bar{\epsilon}_{t-3}^2, \bar{\epsilon}_{t-4}^2, \bar{\epsilon}_{t-5}^2)'$  where  $\bar{\epsilon}_t^2$  is the mean of  $\epsilon_t$  over  $N$  of its squared elementwise.

TABLE II shows the parameter values used to generate the different models. Following TSE, Y. K. [2000], in the second part of this TABLE, we give the variability of the conditional correlation coefficients by calculating the range (maximum – minimum) of these coefficients in each simulated sample. The variability is high with *DCC<sub>2</sub>* and *EDCC<sub>1</sub>*, moderate with *BEKK<sub>2</sub>* and *DCC<sub>1</sub>* and low with *BEKK<sub>1</sub>*, *ASBEKK* and *EDCC<sub>2</sub>*. TABLE III summarizes the empirical powers of the tests by showing the rejection probabilities of the null hypothesis of constant conditional correlations. For all the tests, we can observe that these rejection probabilities increase with the number of observations  $T$ .

For the three *BEKK – GARCH* specifications, *BEKK<sub>1</sub>*, *BEKK<sub>2</sub>* and *ASBEKK*, the tests reject the null hypothesis of constancy of conditional correlation. Nevertheless, even if most of them present good properties with a rejection probability sometimes equal to 1, the power properties of the *TSE* test are the worst, especially for  $T = 1000$ . The *DCC – GARCH* models, *DCC<sub>1</sub>* and *DCC<sub>2</sub>*, are characterized by parameters implying very different ranges of variability of the conditional correlation coefficients. The intervals are very large in the case of the *DCC<sub>2</sub>* model. The *TAYLOR* test presents the best performances in both DGP models, followed by *NEURAL* for *DCC<sub>1</sub>* and *TSE* for *DCC<sub>2</sub>*. On the other hand, the tests performing worse than the others are *TSE* and *STCC* in the *DCC<sub>1</sub>* model and *NEURAL* and *STCC* in the *DCC<sub>2</sub>* model. Concerning the extended *DCC – GARCH* models, *EDCC<sub>1</sub>* and *EDCC<sub>2</sub>*, the tests showing the best performances are *TAYLOR* and *TSE* in the *EDCC<sub>1</sub>*

TABLE II. – MGARCH data-generating parameters for small sample powers

	$BEKK_1$	$BEKK_2$	$DCC_1$	$DCC_2$	$EDCC_1$	$EDCC_2$	$ASBEKK$
$\zeta_{11}$	0.20	0.20	0.01	0.01	0.01	0.01	0.7
$\zeta_{12}$	0.01	0.04	–	–	–	–	0.3
$\zeta_{22}$	0.20	0.20	0.02	0.02	0.02	0.02	0.9
$\alpha_{11}$	0.10	0.30	0.001	0.001	0.01	0.02	–
$\alpha_{12}$	0.10	0.10	0.00	0.00	0.01	0.05	–
$\alpha_{21}$	0.10	0.10	0.00	0.00	0.01	0.05	–
$\alpha_{22}$	0.10	0.30	0.002	0.002	0.02	0.02	–
$\beta_{11}$	0.70	0.30	0.98	0.98	0.80	0.75	–
$\beta_{12}$	0.10	0.20	0.00	0.00	0.15	0.10	–
$\beta_{21}$	0.10	0.20	0.00	0.00	0.15	0.10	–
$\beta_{22}$	0.70	0.30	0.90	0.90	0.75	0.55	–
$\alpha$	–	–	0.05	0.15	0.15	0.01	0.02
$\beta$	–	–	0.94	0.80	0.84	0.94	0.90
$\gamma$	–	–	–	–	–	–	0.05
$\rho_{12}$	–	–	0.80	0.30	0.80	0.30	–
$corr_{\min_{1000}}$	0.33	0.18	0.26	–0.66	–0.65	0.21	0.49
$corr_{\max_{1000}}$	0.53	0.77	0.94	0.86	0.99	0.38	0.75
$corr_{\min_{1500}}$	0.33	0.17	0.19	–0.69	–0.75	0.21	0.49
$corr_{\max_{1500}}$	0.54	0.78	0.94	0.87	0.99	0.39	0.76
$corr_{\min_{2500}}$	0.33	0.15	0.11	–0.71	–0.81	0.20	0.49
$corr_{\max_{2500}}$	0.56	0.80	0.95	0.88	0.99	0.39	0.77

case and *NEURAL* in the  $EDCC_2$  case. In the first case, the *NEURAL*, *TAYLOR* and *TSE* tests outperform *STCC*. In the second DGP model, the performances are disappointing for each test.

In terms of variability, *NEURAL* has better power properties when the variability of the conditional correlation coefficients is low ( $EDCC_2$ ), and in the latter case, *NEURAL* is the only test that can detect the nonconstancy of conditional correlations. However, *TAYLOR* is better in the case where the variability of the conditional correlation coefficients is high ( $DCC_2$  and  $EDCC_1$ ). Therefore, we can conclude that even if one or the other of the new tests is not the best for all the DGPs, it is useful to apply them jointly because at least one of them detects the time-varying conditional correlations.

All the multivariate GARCH-type models have thus been tested satisfactorily using the ANN-based test and the Taylor expansion-based test compared to *TSE* and *STCC* tests. Simulation experiments with more than two endogenous variables give similar conclusions (see p. 27 and TABLE VI of PÉGUIN-FEISSOLLE, A., and B. SANHAJI [2015]). This finding suggests that whatever the GARCH-type model, both the tests developed in this paper have good size and power properties to identify time-varying conditional correlations.

TABLE III. – Small sample powers of the different constant conditional correlation tests

<i>T</i>	1,000			1,500			2,500		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
<i>BEKK</i> <sub>1</sub>									
<i>NEURAL</i>	0.998	0.998	0.998	1.000	1.000	1.000	1.000	1.000	1.000
<i>TAYLOR</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>TSE</i>	0.832	0.986	0.995	0.989	0.998	0.999	1.000	1.000	1.000
<i>STCC</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>BEKK</i> <sub>2</sub>									
<i>NEURAL</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>TAYLOR</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>TSE</i>	0.983	0.987	0.990	0.995	1.000	1.000	1.000	1.000	1.000
<i>STCC</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>ASBEKK</i>									
<i>NEURAL</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>TAYLOR</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>TSE</i>	0.406	0.578	0.666	0.5335	0.720	0.800	0.739	0.876	0.919
<i>STCC</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>DCC</i> <sub>1</sub>									
<i>NEURAL</i>	0.392	0.536	0.612	0.410	0.546	0.619	0.433	0.585	0.659
<i>TAYLOR</i>	0.394	0.534	0.616	0.446	0.602	0.694	0.525	0.648	0.722
<i>TSE</i>	0.060	0.149	0.241	0.100	0.234	0.344	0.191	0.375	0.487
<i>STCC</i>	0.197	0.329	0.416	0.213	0.350	0.442	0.241	0.383	0.466
<i>DCC</i> <sub>2</sub>									
<i>NEURAL</i>	0.144	0.265	0.359	0.163	0.303	0.397	0.243	0.409	0.501
<i>TAYLOR</i>	0.970	0.988	0.992	0.999	0.999	0.999	1.000	1.000	1.000
<i>TSE</i>	0.829	0.943	0.966	0.967	0.991	0.997	0.997	0.999	1.000
<i>STCC</i>	0.237	0.369	0.454	0.232	0.388	0.463	0.249	0.397	0.456
<i>EDCC</i> <sub>1</sub>									
<i>NEURAL</i>	0.693	0.784	0.830	0.711	0.797	0.835	0.756	0.832	0.870
<i>TAYLOR</i>	0.923	0.956	0.969	0.969	0.981	0.986	0.994	0.996	0.996
<i>TSE</i>	0.799	0.865	0.895	0.874	0.909	0.986	0.911	0.917	0.929
<i>STCC</i>	0.454	0.566	0.631	0.509	0.604	0.654	0.615	0.708	0.747
<i>EDCC</i> <sub>2</sub>									
<i>NEURAL</i>	0.410	0.512	0.573	0.418	0.527	0.593	0.503	0.605	0.676
<i>TAYLOR</i>	0.023	0.088	0.163	0.031	0.104	0.178	0.127	0.188	0.247
<i>TSE</i>	0.022	0.081	0.138	0.025	0.094	0.158	0.131	0.199	0.262
<i>STCC</i>	0.064	0.156	0.238	0.069	0.165	0.243	0.154	0.250	0.335

Note: The number of replications is  $S = 2000$ . For *NEURAL* the number of hidden units is  $p = 20$  and the number of lags of residuals is  $q = 3$ , and for *TAYLOR* the order of the Taylor expansion is  $k = 3$  and the number of lags of residuals is  $q = 2$ .

## IV. Robustness to Nonnormality

In order to investigate the robustness of the different tests to nonnormality, we simulate some models with a nonnormal conditional distribution: the multivariate versions of *GARCH-t* model of BOLLERSLEV, T. [1987] and *Beta-t-EGARCH* model of HARVEY, A., and T. CHAKRAVARTY [2008], HARVEY, A. C. [2013] and HARVEY, A., and G. SUCARRAT [2014] (which is an unrestricted version of the Generalised Autoregressive Score (GAS) model of CREAL, D., S. J. KOOPMANS, and A. LUCAS [2013]). The *Beta-t-EGARCH* models are not traditional *GARCH*-type models in the sense that they do not nest *GARCH* models, but it is interesting to compare the performance of the different tests in such cases. In the Monte Carlo simulations, the number of replications is  $S = 2000$ , the sample sizes are  $T = 1000, 1500$  and  $2500$  and  $N = 2$  endogenous variables; we consider nominal sizes of 1%, 5% and 10%.

### IV.1. Size Simulations

We first study the empirical sizes of the different tests when the following two GDP with zero correlations are considered. The first DGP corresponds to two independent univariate *t-GARCH* models which are characterized by  $\alpha_{12} = \alpha_{21} = 0$ ,  $\beta_{12} = \beta_{21} = 0$  and  $\rho_{12} = 0$  in (11), and thus  $h_{12t} = 0$ . The second DGP consists of two independent univariate *Beta-t-EGARCH* models with and without a leverage effect. We use the same notations as HARVEY, A., and G. SUCARRAT [2014], who developed a more general framework than HARVEY, A., and T. CHAKRAVARTY [2008]. Each *Beta-t-EGARCH* model without a leverage effect is defined as:

$$y_{it} = \mu_i + \varepsilon_{it} \exp(\lambda_{i,t|t-1}), i = 1, 2, t = 1, \dots, T, \quad (18)$$

where  $\varepsilon_{it}$  has a  $t_{v_i}$ -distribution and is serially independent, and  $v_i$ , the number of degrees of freedom, is positive. Let us define the conditional score

$$u_{it} = \frac{(v_i + 1)(y_{it} - \mu_i)^2}{v_i \exp(2\lambda_{i,t|t-1}) + (y_{it} - \mu_i)^2} - 1, -1 \leq u_{it} \leq v_i, v_i > 0. \quad (19)$$

We consider the first order model given by

$$\lambda_{i,t|t-1} = \delta_i + \phi_i \lambda_{i,t-1|t-2} + \kappa_i u_{i,t-1} \quad (20)$$

with the stationarity condition  $|\phi_i| < 1$ . The first order *Beta-t-EGARCH* with a leverage effect is defined in HARVEY, A., and G. SUCARRAT [2014] as follows

$$\lambda_{i,t|t-1} = \delta_i + \phi_i \lambda_{i,t-1|t-2} + \kappa_i u_{i,t-1} + \kappa_i^* \text{sign}(-(y_{i,t-1} - \mu_i))(u_{i,t-1} + 1), \quad (21)$$

(see HARVEY, A., and T. CHAKRAVARTY [2008] and HARVEY, A., and G. SUCARRAT [2014] for a complete definition of these models and their properties).

The data-generating parameters are given in TABLE IV. TABLE V shows the small sample sizes of *t-GARCH* and *Beta-t-EGARCH*. When the DGP models are the *t-GARCH* and *Beta-t-EGARCH*, with or without leverage effects, the results highlight the relative

TABLE IV. – MGARCH data-generating parameters for small sample sizes and powers under nonnormality

	$t - GARCH$	$Beta - t - EGARCH$	$DCC - t - GARCH$	$DCC - Beta - t - EGARCH$
$\alpha$	–	–	0.05	0.15
$\beta$	–	–	0.94	0.80
$\rho_{12}$	0.00	0.00	0.70	0.60
$\nu$	10	10	10	10
$\zeta_1$	0.40	–	0.30	–
$\zeta_2$	0.20	–	0.20	–
$\alpha_{11}$	0.15	–	0.10	–
$\alpha_{12}$	0.00	–	0.00	–
$\alpha_{21}$	0.00	–	0.00	–
$\alpha_{22}$	0.20	–	0.20	–
$\beta_{11}$	0.80	–	0.70	–
$\beta_{12}$	0.00	–	0.00	–
$\beta_{21}$	0.00	–	0.00	–
$\beta_{22}$	0.70	–	0.60	–
$\mu_1$	–	0.00	–	0.00
$\mu_2$	–	0.00	–	0.00
$\delta_1$	–	0.007	–	0.02
$\delta_2$	–	0.05	–	0.04
$\phi_1$	–	0.99	–	0.95
$\phi_2$	–	0.90	–	0.90
$\kappa_1$	–	0.05	–	0.05
$\kappa_2$	–	0.07	–	0.07
$\kappa_1^*$	–	0.02* or 0.00	–	0.04* or 0.00
$\kappa_2^*$	–	0.04* or 0.00	–	0.06* or 0.00

Note: \* With leverage effect in the  $Beta - t - EGARCH$  and  $DCC - Beta - t - EGARCH$ .

robustness to nonnormality of the *TAYLOR*, *NEURAL* and *STCC* tests compared to the *TSE* test, which tends to over-reject the null hypothesis of constant conditional correlations. In the  $Beta - t - EGARCH$ , the conditional variances do not follow *GARCH* specification even if the conditional correlations are constant. Nevertheless, the test results are very similar in both DGPs.

#### IV.2. Power Simulations

To study the empirical powers of the different tests, we use the same models as in SECTION IV.1 except that the conditional correlations are no longer constant. The conditional expectation model can be written, for  $t = 1, \dots, T$ , and  $i = 1, \dots, N$ ,  $y_{it} = \sqrt{h_{iit}} \varepsilon_{it}$  where for the *GARCH*(1,1) models,  $h_{iit}$  is given by (2) and for the  $Beta - t - EGARCH$  models,

$$h_{iit} = [\exp(\lambda_{i,t|t-1})]^2 \quad (22)$$



TABLE V. – Small sample sizes of constant conditional correlation tests under nonnormality

$T$	1,000			1,500			2,500		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
<i>t</i> – GARCH									
NEURAL	0.014	0.058	0.116	0.016	0.059	0.102	0.016	0.056	0.102
TAYLOR	0.014	0.064	0.115	0.013	0.061	0.118	0.022	0.072	0.118
TSE	0.056	0.112	0.168	0.056	0.125	0.191	0.075	0.153	0.219
STCC	0.012	0.055	0.097	0.011	0.043	0.091	0.012	0.051	0.101
<i>Beta</i> – <i>t</i> – EGARCH									
NEURAL	0.024	0.057	0.107	0.017	0.064	0.117	0.014	0.053	0.100
TAYLOR	0.026	0.070	0.116	0.018	0.071	0.125	0.014	0.056	0.100
TSE	0.088	0.154	0.215	0.072	0.156	0.220	0.077	0.152	0.220
STCC	0.015	0.053	0.101	0.017	0.061	0.110	0.011	0.049	0.107
<i>Beta</i> – <i>t</i> – EGARCH*									
NEURAL	0.018	0.052	0.105	0.017	0.057	0.114	0.015	0.052	0.111
TAYLOR	0.024	0.063	0.119	0.015	0.063	0.117	0.015	0.055	0.107
TSE	0.064	0.133	0.189	0.061	0.138	0.211	0.065	0.141	0.195
STCC	0.012	0.052	0.103	0.012	0.058	0.102	0.011	0.048	0.099

Note: The number of replications is  $S = 2000$ . For *NEURAL* the number of hidden units is  $p = 20$  and the number of lags of residuals is  $q = 3$ , and for *TAYLOR* the order of the Taylor expansion is  $k = 3$  and the number of lags of residuals is  $q = 2$ . *Beta* – *t* – EGARCH and *Beta* – *t* – EGARCH\* are the models respectively without and with leverage effect.

as in equation (18);  $\lambda_{i,t|t-1}$  is determined as a first order model defined by (20) or (21) respectively without or with leverage effect. Moreover,  $\varepsilon_t$  has a conditional multivariate  $t_v$ -distribution with positive degrees of freedom  $v$ :  $\varepsilon_t | \Omega_{t-1} \sim t_v(0, H_t)$ .

We use the same notations as in the *DCC* – *GARCH* model:  $H_t = S_t P_t S_t$ , the  $N \times N$  conditional correlation matrix  $P_t$  is written as in (15) and  $S_t = \text{diag}(\hat{\sigma}_{11,t}, \dots, \hat{\sigma}_{NN,t})$ . To specify the different diagonal elements of  $S_t$ , we refer to HARVEY, A., and T. CHAKRAVARTY [2008]:  $\sigma_{ii,t}^* = \sqrt{h_{iit}}$  is the time-varying scale parameter, which is not necessarily equal to the conditional standard deviation. To ensure the existence of a unit variance we resample the variance of the residuals  $\varepsilon_{i,t}^* \sim \text{IID}(0, \sigma_{\varepsilon^*}^2)$  such that  $\hat{\sigma}_{ii,t} = \hat{\sigma}_{\varepsilon^*} \sigma_{ii,t}^*$ , where  $\hat{\sigma}_{\varepsilon^*}^2 = \sum (\hat{\varepsilon}_{i,t}^* - \bar{\varepsilon}_i^*)^2 (T - 1)^{-1}$ .  $Q_t = (q_{ij,t})$  is a symmetric positive definite  $N \times N$  matrix given by (16) with  $\bar{Q}$  a  $N \times N$  constant and positive definite matrix.

The data-generating parameters are given in TABLE IV. TABLE VI shows the empirical powers of the different tests. In all cases, the two new tests show the best empirical powers, even for  $T = 1000$  or  $1500$ , and *TSE* again is the least powerful.

As expected, the rejection probabilities of the null hypothesis increase with the number of observations  $T$ . To conclude, the simulation results in the case of time-varying conditional correlation and nonnormality assumptions show that *TAYLOR* and *NEURAL* are the best performing tests out of the four. The poor performance of the *TSE* test may be explained by the fact that the alternative hypothesis of this test is rather restrictive compared to that of the *TAYLOR* and *NEURAL* tests.

TABLE VI. – Small sample powers of constant conditional correlation tests under nonnormality

$T$	1,000			1,500			2,500		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
<i>DCC – t – GARCH</i>									
<i>NEURAL</i>	0.283	0.450	0.536	0.374	0.526	0.619	0.490	0.643	0.727
<i>TAYLOR</i>	0.511	0.668	0.739	0.636	0.772	0.834	0.788	0.885	0.922
<i>TSE</i>	0.113	0.223	0.329	0.183	0.338	0.436	0.245	0.432	0.530
<i>STCC</i>	0.246	0.380	0.461	0.301	0.427	0.493	0.339	0.455	0.529
<i>DCC – Beta – t – EGARCH</i>									
<i>NEURAL</i>	0.601	0.736	0.806	0.68	0.807	0.857	0.779	0.862	0.901
<i>TAYLOR</i>	0.792	0.886	0.928	0.878	0.942	0.961	0.965	0.986	0.994
<i>TSE</i>	0.107	0.256	0.376	0.188	0.349	0.471	0.284	0.463	0.567
<i>STCC</i>	0.459	0.548	0.595	0.594	0.663	0.704	0.591	0.669	0.720
<i>DCC – Beta – t – EGARCH*</i>									
<i>NEURAL</i>	0.676	0.801	0.851	0.762	0.855	0.897	0.854	0.917	0.943
<i>TAYLOR</i>	0.856	0.931	0.959	0.917	0.967	0.978	0.982	0.994	0.997
<i>TSE</i>	0.156	0.251	0.342	0.193	0.292	0.386	0.224	0.338	0.430
<i>STCC</i>	0.569	0.643	0.690	0.658	0.729	0.759	0.753	0.813	0.842

Note: The number of replications is  $S = 2000$ . For *NEURAL* the number of hidden units is  $p = 20$  and the number of lags of residuals is  $q = 3$ , and for *TAYLOR* the order of the Taylor expansion is  $k = 3$  and the number of lags of residuals is  $q = 2$ . *DCC – Beta – t – EGARCH* and *DCC – Beta – t – EGARCH\** are the models respectively without and with leverage effect.

## V. Empirical Application: Correlations between Three Asset Returns

We consider the asset returns of three “blue-chip” US daily stocks included in the Dow Jones Industrial Average index: JPMorgan Chase & Co. (JPM), The Coca-Cola Company (KO) and Exxon Mobil Corporation (XOM), from January 03, 2000 to October 12, 2012 ( $T = 3217$  daily observations). The series correspond to closing prices adjusted for dividends and splits and are obtained from Yahoo Finance. We compute the daily returns as  $100 \times \log(p_t / p_{t-1})$ , where  $p_t$  represents the daily closing price at time  $t$ ,  $t = 1, \dots, T$ . To avoid the singularity issues of the Hessian matrix for multivariate optimization and in order to have a comparable metric between the four tests, we winsorized the few peaks in the returns at  $-10$ . Panel A of TABLE VII presents the summary statistics of the winsorized asset returns; medians, means, standard deviations, skewness and kurtosis. The returns exhibit positive excess kurtosis.

We compute all the tests on the asset returns as bivariate and trivariate combinations in order to determine whether these series are conditionally correlated over time or not. We use the same values for  $p$ ,  $q$  and  $k$  as in the simulation experiments. Following SILVENNOINEN, A., and T. TERÄSVIRTA [2015], we use three different transition variables for the *STCC* test: the lagged absolute S&P 500 index returns averaged over ten days ( $STCC_1$ ), the lagged S&P 500 index average return over twenty days ( $STCC_2$ ) and the lagged S&P 500 index average return over two days ( $STCC_3$ ). Therefore, not rejecting the null of constant correlations using the *STCC* test indicates that there is not enough evidence of time-varying correlations given the chosen transition variable, but the conclusion may be different with another transition variable.

Panel B of TABLE VII shows the results of the tests. In the case of the relationship between JPM and KO, *NEURAL* and *TAYLOR* reject the null hypothesis of constant conditional

TABLE VII. – Summary statistics and constant conditional correlation tests of the daily asset returns

**Panel A: Summary Statistics**

	Mean	Med.	Max.	Min.	S.D.	Skew.	Kurt.
<i>JPM</i>	0.01	-0.02	22.39	-41.13	2.75	0.91	8.08
<i>KO</i>	0.00	0.02	13.00	-69.56	1.46	0.01	8.12
<i>XOM</i>	0.02	0.06	15.86	-67.07	1.67	0.18	8.25

**Panel B: Constant Conditional Correlation Tests**

	KO-JPM		JPM-XOM		XOM-KO		JPM-KO-XOM	
	Stat.	p-v	Stat.	p-v	Stat.	p-v	Stat.	p-v
<i>NEURAL</i>	21.8	0.001	11.8	0.067	10.9	0.015	53.8	0.000
<i>TAYLOR</i>	39.7	0.005	73.7	0.000	63.8	0.000	126	0.000
<i>TSE</i>	3.48	0.062	4.36	0.037	0.00	0.985	16.8	0.000
<i>STCC<sub>1</sub></i>	0.01	0.925	0.46	0.498	3.88	0.049	4.95	0.175
<i>STCC<sub>2</sub></i>	0.57	0.448	2.19	0.139	0.05	0.829	4.47	0.215
<i>STCC<sub>3</sub></i>	2.08	0.149	2.91	0.015	4.66	0.031	10.8	0.013

Note: Stat. and p-v represent respectively the value of the test statistic and the associated *p*-value. For *NEURAL* the number of hidden units is  $p = 20$  and the number of lags of residuals is  $q = 3$ , and for *TAYLOR* the order of the Taylor expansion is  $k = 3$  and the number of lags of residuals is  $q = 2$ . The transition variables used in *STCC* are the lagged absolute S&P 500 index returns averaged over ten days (*STCC<sub>1</sub>*), the lagged S&P 500 index average return over twenty days (*STCC<sub>2</sub>*) and over two days (*STCC<sub>3</sub>*).

correlations at the 1% level, whereas *TSE* rejects it only at the 10% level. *STCC*, whatever the transition variable (*STCC<sub>1</sub>*, *STCC<sub>2</sub>* or *STCC<sub>3</sub>*) does not reject the null hypothesis of constancy of conditional correlations. For JPM and XOM returns, *STCC<sub>1</sub>* and *STCC<sub>2</sub>* do not reject the null hypothesis while the other statistics reject the null at different levels (1% for *TAYLOR*, 5% for *TSE* and *STCC<sub>3</sub>* and 10% for *NEURAL*). When XOM and KO returns are considered, *TAYLOR* rejects the null at the 1% level, whereas *NEURAL*, *STCC<sub>1</sub>* and *STCC<sub>3</sub>* reject it at the 5% level. *TSE* and *STCC<sub>2</sub>* do not reject constancy. For the trivariate case (JPM-KO-XOM), all the tests reject the null hypothesis at the 1% or 5% level, except *STCC<sub>1</sub>* and *STCC<sub>2</sub>* which do not reject constancy.

Therefore, if we consider only the *STCC* test in the first case, *STCC<sub>1</sub>* and *STCC<sub>2</sub>* in the second and last cases, *STCC<sub>2</sub>* and *TSE* in the third case (the conclusions of the *STCC* test being obviously constrained by the chosen transition variables), we do not reject the constancy of conditional correlations but do conclude that the conditional correlations are time-varying with *NEURAL* and *TAYLOR*.

## VI. Concluding Remarks

In this paper, we propose two tests for the constancy of conditional correlations in MGARCH models. The main practical findings in this paper are that our new tests perform well in small-sample simulations, even in the case of nonnormality. They show better performances than the *STCC* and *TSE* tests in most cases, probably because their original form is very general. Indeed, the first test is based on artificial neural networks that are universal approximators, and the second comes from the linearization of an unknown relationship determining each conditional correlation.

Both tests should be applied jointly because they have different finite-sample properties, as shown by the small sample simulations and the financial application. When we consider their power, the *NEURAL* test has better properties when the variability of the conditional correlation coefficients is low, whereas *TAYLOR* has higher power when the variability is high. We recommend that a model with time-varying conditional correlations should be used if any one of the two tests rejects the null hypothesis of constant conditional correlations.

#### ACKNOWLEDGEMENT

The authors would like to thank Luc Bauwens, René Garcia, Christophe Hurlin, Roselyne Joyeux, Sébastien Laurent, Genaro Sucarrat and Timo Teräsvirta for stimulating discussions and remarks. Any remaining errors are ours.

*Correspondence:*

Anne Péguin-Feissolle

GREQAM, Château Lafarge

50 Chemin du Château Lafarge, 13290 Les Milles, France

Tel: +33(0)4 42 93 59 60, Fax: +33(0)4 42 93 09 68

E-mail: anne.peguin@univ-amu.fr

Bilel Sanhaji

Université Paris 8, LED

2 rue de la Liberté, 93526 Saint-Denis Cedex, France

E-mail: bilel.sanhaji@univ-paris8.fr

## Appendix

### A. Derivation of the Statistics of the Artificial Neural Network Test and the Taylor Expansion-Based Test

In order to compute the *NEURAL* and the *TAYLOR* statistics, we use broadly ABADIR, K. M., and J. R. MAGNUS [2005], ANDERSON, T. W. [2003], LÜTKEPOHL, H. [1996] and the Appendix of SILVENNOINEN, A., and T. TERÄSVIRTA [2005] for the derivations based on the rules of matrix algebra. Let us call  $r^*$  the number of parameters that are equal to zero under the null hypothesis, i.e.  $r^* = p$  for the *NEURAL* statistic and  $r^* = N^*$  for the *TAYLOR* statistic.

Each test statistic can be written (we simplify by  $T$ ):

$$STATISTIC = \left( \sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \theta} \right)' \mathfrak{S}(\theta)^{-1} \left( \sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \theta} \right). \quad (\text{A.1})$$

The vector  $\theta$  of all the parameters of the model is given by:  $\theta = (\omega', \rho', \delta')'$ ;  $\omega = (\omega'_1, \dots, \omega'_N)$  is composed by the  $3 \times 1$  parameter vector for the conditional variances  $\omega_i = (\zeta_i, \alpha_i, \beta_i)'$  for  $i = 1, \dots, N$ ,  $\rho$  is the vector of the conditional correlations  $\rho_{ij}$  with  $1 \leq i < j \leq N$  and  $\delta$  is the

vector of parameters that are equal to zero under the null hypothesis. Therefore,  $\sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \theta}$  is a  $\left(3N + (1 + r^*) \frac{N(N-1)}{2}\right) \times 1$  vector as follows

$$\sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \theta} = \begin{pmatrix} V_{\omega} \\ V_{\rho} \\ V_{\delta} \end{pmatrix} \quad (\text{A.2})$$

and  $\mathfrak{I}(\theta)$  is a  $\left(3N + (1 + r^*) \frac{N(N-1)}{2}\right) \times \left(3N + (1 + r^*) \frac{N(N-1)}{2}\right)$  matrix given by

$$\mathfrak{I}(\theta) = \begin{pmatrix} M_{\omega\omega} & M'_{\rho\omega} & M'_{\delta\omega} \\ M_{\rho\omega} & M_{\rho\rho} & M'_{\delta\rho} \\ M_{\delta\omega} & M_{\delta\rho} & M_{\delta\delta} \end{pmatrix}. \quad (\text{A.3})$$

So we can write:

$$\text{STATISTIC} = \begin{pmatrix} V_{\omega} \\ V_{\rho} \\ V_{\delta} \end{pmatrix}' \begin{pmatrix} M_{\omega\omega} & M'_{\rho\omega} & M'_{\delta\omega} \\ M_{\rho\omega} & M_{\rho\rho} & M'_{\delta\rho} \\ M_{\delta\omega} & M_{\delta\rho} & M_{\delta\delta} \end{pmatrix}^{-1} \begin{pmatrix} V_{\omega} \\ V_{\rho} \\ V_{\delta} \end{pmatrix} \quad (\text{A.4})$$

A.I. *Different Sub-Vectors of*  $\sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \theta}$

$$V_{\omega} = \left( \sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \omega'_1} \quad \dots \quad \sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \omega'_N} \right)', \quad V_{\rho} = \left( \sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \text{vecl}(P)} \right), \quad (\text{A.5})$$

and

$$V_{\delta} = \text{vecl} \left[ \left( \sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \delta_{ij}} \right)_{i,j} \right], \quad (\text{A.6})$$

where  $V_{\omega}$ ,  $V_{\rho}$  and  $V_{\delta}$  are respectively  $3N \times 1$ ,  $\frac{N(N-1)}{2} \times 1$  and  $r^* \frac{N(N-1)}{2} \times 1$  vectors. *vecl* is an operator stacking the columns of the strict lower diagonal, i.e. excluding the diagonal elements of the matrix. Moreover, we have, for  $i = 1, \dots, N$  and  $1 \leq i < j \leq N$ :

$$\begin{cases} \frac{\partial l_t(\theta)}{\partial \omega_i} = \frac{1}{2} \frac{1}{h_{iit}} \frac{\partial h_{iit}}{\partial \omega_i} [-1 + z_{it}' P_t^{-1} z_t] \\ \frac{\partial l_t(\theta)}{\partial \text{vecl}(P)} = -\frac{1}{2} U' [\text{vec}(P_t^{-1}) - (P_t^{-1} \otimes P_t^{-1})(z_t \otimes z_t)] \\ \frac{\partial l_t(\theta)}{\partial \delta_{ij}} = \frac{\partial l_t(\theta)}{\partial \delta_{ji}} = -\frac{1}{2} W_{ijt} [\text{vec}(P_t^{-1}) - (P_t^{-1} \otimes P_t^{-1})(z_t \otimes z_t)] \end{cases} \quad (\text{A.7})$$

where  $\frac{\partial l_t(\theta)}{\partial \omega_i}$ ,  $\frac{\partial l_t(\theta)}{\partial \text{vec}(P)}$  and  $\frac{\partial l_t(\theta)}{\partial \delta_{ij}}$  are respectively  $3 \times 1$ ,  $\frac{N(N-1)}{2} \times 1$  and  $r^* \times 1$  vectors.

Moreover,  $z_t = S_t^{-1} \epsilon_t$  and the matrix  $U$  is an  $N^2 \times \frac{N(N-1)}{2}$  matrix of zeros and ones whose columns are defined as

$$\left( \text{vec}(1_i 1'_j + 1_j 1'_i) \right)_{\substack{i=1, \dots, N-1 \\ j=i+1, \dots, N}} \tag{A.8}$$

and the columns appear in the same order from left to right as the indices in  $\text{vec}(P)$ . The matrix  $W_{ijt}$  is an  $r^* \times N^2$  matrix given by

$$W_{ijt} = \text{diag}(g_{ijt}) \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \otimes (\text{vec}(1_i 1'_j + 1_j 1'_i))' \right\}; \tag{A.9}$$

$g_{ijt}$  is a  $r^* \times 1$  vector defined in the *NEURAL* test as

$$g_{ijt} = \begin{pmatrix} g_{ijt,1} \\ \vdots \\ g_{ijt,r^*} \end{pmatrix} = \begin{pmatrix} (1 + \exp\{-w'_{ijt} \gamma_{ij1}\})^{-1} \\ \vdots \\ (1 + \exp\{-w'_{ijt} \gamma_{ijr^*}\})^{-1} \end{pmatrix} \tag{A.10}$$

and, in the *TAYLOR* test, it is replaced by  $d_{ijr}$ , composed by all terms involving functions of elements of lagged values of  $\epsilon_{i,t}$  and  $\epsilon_{j,t}$ . We have:

$$r^* = 2 \left[ \sum_{r=1}^k \binom{2q+r-1}{r} - \sum_{r=1}^k \binom{q+r-1}{r} \right]. \tag{A.11}$$

### A.II. Different Elements of the Matrix $\mathfrak{I}(\theta)$

—  $M_{\omega\omega}$  is the following  $3N \times 3N$  matrix:

$$M_{\omega\omega} = \sum_{t=1}^T \begin{pmatrix} M_{\omega\omega,11} & M'_{\omega\omega,12} & \cdots & M'_{\omega\omega,1N} \\ M_{\omega\omega,12} & M_{\omega\omega,22} & \cdots & M'_{\omega\omega,2N} \\ \vdots & \vdots & \ddots & \vdots \\ M_{\omega\omega,1N} & M_{\omega\omega,2N} & \cdots & M_{\omega\omega,NN} \end{pmatrix} \tag{A.12}$$

where, for  $i = 1, \dots, N$  and  $1 \leq i < j \leq N$ ,  $M_{\omega\omega,ii}$  and  $M_{\omega\omega,ij}$  are  $3 \times 3$  matrices defined as

$$\begin{cases} M_{\omega\omega,ii} = E_{t-1} \left[ \frac{\partial l_t(\theta)}{\partial \omega_i} \frac{\partial l_t(\theta)}{\partial \omega'_i} \right] = \frac{1}{4} \frac{1}{h_{iit}^2} \frac{\partial h_{iit}}{\partial \omega_i} \left[ 1'_i P_t^{-1} 1_i + 1 \right] \frac{\partial h_{iit}}{\partial \omega_i} \\ M_{\omega\omega,ij} = E_{t-1} \left[ \frac{\partial l_t(\theta)}{\partial \omega_i} \frac{\partial l_t(\theta)}{\partial \omega'_j} \right] = \frac{1}{4} \frac{1}{h_{iit} h_{jjt}} \frac{\partial h_{iit}}{\partial \omega_i} \left[ \rho_{ijt} 1'_i P_t^{-1} 1_j + 1 \right] \frac{\partial h_{jjt}}{\partial \omega_j} \end{cases} \tag{A.13}$$

—  $M_{\rho\rho}$  is the following  $\frac{N(N-1)}{2} \times \frac{N(N-1)}{2}$  matrix:

$$M_{\rho\rho} = \sum_{t=1}^T \frac{1}{4} U' [(P_t^{-1} \otimes P_t^{-1}) + (P_t^{-1} \otimes I)K(P_t^{-1} \otimes I)]U \quad (\text{A.14})$$

—  $M_{\rho\omega}$  is a  $\frac{N(N-1)}{2} \times 3N$  matrix defined as

$$M_{\rho\omega} = \sum_{t=1}^T (M'_{\rho\omega,1} \quad M'_{\rho\omega,2} \quad \dots \quad M'_{\rho\omega,N}) \quad (\text{A.15})$$

where each  $3 \times \frac{N(N-1)}{2}$  matrix  $M_{\rho\omega,i}$ , for  $i = 1, \dots, N$ , is:

$$\begin{aligned} M_{\rho\omega,i} &= E_{t-1} \left[ \frac{\partial l_t(\theta)}{\partial \omega_i} \frac{\partial l_t(\theta)}{\partial \text{vecl}(P)'} \right] \\ &= \frac{1}{4} \frac{1}{h_{iii}} \frac{\partial h_{iii}}{\partial \omega_i} [l'_i(l'_i \otimes P_t^{-1}) + l'_i(l'_i \otimes I)K(P_t^{-1} \otimes I)]U \end{aligned} \quad (\text{A.16})$$

—  $M_{\delta\rho}$  is a  $r^* \frac{N(N-1)}{2} \times \frac{N(N-1)}{2}$  matrix given by

$$M_{\delta\rho} = \sum_{t=1}^T \text{vecl}[(M_{\delta\rho,ij})_{i,j}] \quad (\text{A.17})$$

where each  $r^* \times \frac{N(N-1)}{2}$  matrix  $M_{\delta\rho,ij}$  is

$$\begin{aligned} M_{\delta\rho,ij} &= M_{\delta\rho,ji} = E_{t-1} \left[ \frac{\partial l_t(\theta)}{\partial \delta_{ij}} \frac{\partial l_t(\theta)}{\partial \text{vecl}(P)'} \right] \\ &= \frac{1}{4} W_{ijt} [(P_t^{-1} \otimes P_t^{-1}) + (P_t^{-1} \otimes I)K(P_t^{-1} \otimes I)]U \end{aligned} \quad (\text{A.18})$$

—  $M_{\delta\delta}$  is the following  $r^* \frac{N(N-1)}{2} \times r^* \frac{N(N-1)}{2}$  matrix:

$$M_{\delta\delta} = \sum_{t=1}^T [M_{\delta\delta}(i, j, n, m)]_{i,j,n,m} \quad (\text{A.19})$$

with  $M_{\delta\delta}(i, j, n, m)$  a  $r^* \times r^*$  matrix given by:

$$\begin{aligned} M_{\delta\delta}(i, j, n, m) &= E_{t-1} \left[ \frac{\partial l_t(\theta)}{\partial \delta_{ij}} \frac{\partial l_t(\theta)}{\partial \delta'_{nm}} \right] \\ &= \frac{1}{4} W_{ijt} [(P_t^{-1} \otimes P_t^{-1}) + (P_t^{-1} \otimes I)K(P_t^{-1} \otimes I)]W'_{nmt}; \end{aligned} \quad (\text{A.20})$$

for each couple  $(i, j)$  taken in the same order as for  $V_{\delta}$ , take the couple  $(n, m)$  in the same order.

—  $M_{\delta\omega}$  is a  $r^* \frac{N(N-1)}{2} \times 3N$  matrix given by:

$$M_{\delta\omega} = \sum_{t=1}^T [M_{\delta\omega}(i, j, n)]_{i,j,n} \tag{A.21}$$

with  $M_{\delta\omega}(i, j, n)$  a  $r^* \times 3$  matrix defined by:

$$\begin{aligned} M_{\delta\omega}(i, j, n) &= E_{t-1} \left[ \frac{\partial l_t(\theta)}{\partial \delta_{ij}} \frac{\partial l_t(\theta)}{\partial \omega'_n} \right] \\ &= -\frac{1}{4h_{mnt}} W_{ijt} \left[ \text{vec}(P_t^{-1}) - \left\{ (I \otimes 1_n) + (P_t^{-1} \otimes I)K(I \otimes P_t 1_n) \right. \right. \\ &\quad \left. \left. + \text{vec}(P_t^{-1})\text{vec}(1'_n P_t) \right\} P_t^{-1} 1_n \right] \frac{\partial h_{mnt}}{\partial \omega'_n}; \end{aligned} \tag{A.22}$$

for each couple  $(i, j)$  taken in the same order as for  $V_{\delta}$ , take all the  $n = 1, 2, \dots, N$ .

For the computation of all these vectors and matrices, we have:

$$\frac{\partial h_{iit}}{\partial \omega_i} = v_{i,t-1} + \beta_i \frac{\partial h_{i,t-1}}{\partial \omega_i}. \tag{A.23}$$

with  $v_{it} = (1, \ \varepsilon_{it}^2, \ h_{iit})'$  for  $i = 1, \dots, N$  (we compute recursively  $\frac{\partial h_{iit}}{\partial \omega_i}$ ).

Following SILVENNOINEN, A., and T. TERÄSVIRTA [2015], we have for a model with general correlation matrix  $P_t$

$$\begin{cases} E_{t-1}[z_t z'_t \otimes z_t z'_t] = (P_t \otimes P_t) + (I \otimes P_t)K(I \otimes P_t) + \text{vec}P_t(\text{vec}P_t)' \\ E_{t-1}(z_{it} z'_{jt} \otimes z_t z'_t) = (1'_i \otimes I)E_{t-1}(z_t z'_t \otimes z_t z'_t)(1_j \otimes I) \\ E_{t-1}[z_{it} z'_t \otimes z_t z'_t] = (1'_i \otimes I)E_{t-1}[z_t z'_t \otimes z_t z'_t] \end{cases} \tag{A.24}$$

with  $K$  a  $N^2 \times N^2$  matrix defined as

$$K = \begin{bmatrix} 1_1 1'_1 & \dots & 1_i 1'_1 & \dots & 1_N 1'_1 \\ \vdots & & \vdots & & \vdots \\ 1_1 1'_j & \dots & 1_i 1'_j & \dots & 1_N 1'_j \\ \vdots & & \vdots & & \vdots \\ 1_1 1'_N & \dots & 1_i 1'_N & \dots & 1_N 1'_N \end{bmatrix} \tag{A.25}$$

Following (A.4), and using the rules of the partitioned matrices and the null hypothesis, the test statistic can be written

$$STATISTIC = V'_{\delta} \left[ M_{\delta\delta} - (M_{\delta\omega} \ M_{\delta\rho}) \begin{pmatrix} M_{\omega\omega} & M'_{\rho\omega} \\ M_{\rho\omega} & M_{\rho\rho} \end{pmatrix}^{-1} \times \begin{pmatrix} M'_{\delta\omega} \\ M'_{\delta\rho} \end{pmatrix} \right]^{-1} V_{\delta}. \tag{A.26}$$



## References

- ABADIR, K. M., and J. R. MAGNUS (2005): *Matrix Algebra*. Cambridge University Press. [94]
- AIELLI, G. P. (2013): “Dynamic Conditional Correlation: On Properties and Estimation”, *Journal of Business and Economic Statistics*, **31**, 282-299. [86]
- ANDERSON, T. W. (2003): *An Introduction to Multivariate Statistical Analysis*. New York: Wiley, 3 edn. [94]
- BAUWENS, L., S. LAURENT, and J. V. K. ROMBOUTS (2006): “Multivariate GARCH Models: A Survey”, *Journal of Applied Econometrics*, **21**, 79-109. [77]
- BERA, A. K., and S. KIM (2002): “Testing Constancy of Correlation and Other Specifications of the BGARCH Model with an Application to International Equity Returns”, *Journal of Empirical Finance*, **9**, 171-195. [77]
- BERBEN, R. P., and W. J. JANSEN (2005): “Comovement in International Equity Markets: A Sectoral View”, *Journal of International Money and Finance*, **24**, 832-857. [77, 79]
- BOLLERSLEV, T. (1987): “A Conditionally Heteroskedastic Time Series Model for Security Prices and Rates of Return Data”, *Review of Economics and Statistics*, **59**, 542-547. [89]
- CARROLL, S. M., and B. W. DICKINSON (1989): “Construction of Neural Nets Using the Radon Transform”, *Proceedings of the International Joint Conference on Neural Networks*, Washington, D.C., IEEE Press I, 607-611. [79]
- CASTLE, J. L., and D. F. HENDRY (2010): “A Low-Dimension Portmanteau Test for Non-Linearity”, *Journal of Econometrics*, **150**, 231-245. [81]
- CAULET, R., and A. PÉGUIN-FEISSOLLE (2000): “Un test d'hétéroscédasticité conditionnelle inspiré de la modélisation en termes de réseaux neuronaux artificiels”, *Annales d'Économie et de Statistique*, **59**, 177-197. [78]
- CREAL, D., S. J. KOOPMANS, and A. LUCAS (2013): “Generalized Autoregressive Score Models with Applications”, *Journal of Applied Econometrics*, **28**, 777-795. [89]
- CYBENKO, G. (1989): “Approximation by Superpositions of a Sigmoid Function”, *Mathematics of Control, Signals and Systems*, **2**, 303-314. [79]
- DING, Z., and R. F. ENGLE (2001): “Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing”, *Academia Economic Papers*, **29**, 157-184. [85]
- ENGLE, R. F. (2002): “Dynamic Conditional Correlation – A Simple Class of Multivariate GARCH Models”, *Journal of Business and Economic Statistics*, **20**, 339-350. [86]
- ENGLE, R. F., and F. KRONER (1995): “Multivariate Simultaneous Generalized ARCH”, *Econometric Theory*, **11**, 122-150. [85]
- HARVEY, A. C. (2013): *Dynamic Models for Volatility and Heavy Tails*. Cambridge University Press. [89]

- HARVEY, A., and T. CHAKRAVARTY (2008): “Beta-t-(E)GARCH”. Cambridge Working Papers in Economics, CWPE 0840. [89, 91]
- HARVEY, A., and G. SUCARRAT (2014): “EGARCH Models with Fat Tails, Skewness and Leverage”, *Computational Statistics and Data Analysis*, **76**, 320-338. [89]
- HARVEY, A., and S. THIELE (2014): “Testing for Time Varying Correlation in a Dynamic Conditional Score Framework”. [78]
- HORNIK, K. (1991): “Approximation Capabilities of Multilayer Feedforward Networks”, *Neural Networks*, **4**, 251-257. [79]
- HORNIK, K., M. STINCHCOMBE, and H. WHITE (1989): “Multi-Layer Feedforward Networks Are Universal Approximators”, *Neural Networks*, **2**, 359-366. [79]
- HORNIK, K., M. STINCHCOMBE, and H. WHITE (1990): “Universal Approximation of an Unknown Mapping and Its Derivatives Using Multi-Layer Feedforward Networks”, *Neural Networks*, **3**, 551-560. [79]
- JEANTHEAU, T. (1998): “Strong Consistency of Estimators for Multivariate ARCH Models”, *Econometric Theory*, **14**, 70-86. [84]
- KAMSTRA, M. (1993): “A Neural Network Test for Heteroscedasticity”. Working Paper, Simon Fraser University, Burnaby, Canada. [78]
- LEBRETON, M., and A. PÉGUIN-FEISSOLLE (2007): “Robust Tests for Heteroscedasticity in a General Framework”, *Annales d'Économie et de Statistique*, **85**, 159-187. [78, 80, 81]
- LEE, T. H., H. WHITE, and C. W. J. GRANGER (1993): “Testing for Neglecting Nonlinearity in Time Series Models”, *Journal of Econometrics*, **56**, 269-290. [78, 80, 83]
- LÜTKEPOHL, H. (1996): *Handbook of Matrices*. Chichester: John Wiley & Sons. [94]
- PÉGUIN-FEISSOLLE, A. (1999): A Comparison of the Power of Some Tests for Conditional Heteroscedasticity”, *Economics Letters*, **63**, 5-17. [78, 79]
- PÉGUIN-FEISSOLLE, A., and B. SANHAJI (2015): “Testing the Constancy of Conditional Correlations in Multivariate GARCH-type Models”. Working Paper AMSE, WP 2015 Nr 16, Aix-Marseille University, France. [80, 83, 87]
- PÉGUIN-FEISSOLLE, A., B. STRIKHOLM, and T. TERÄSVIRTA (2013): “Testing the Granger Noncausality Hypothesis in Stationary Nonlinear Models of Unknown Functional Form”, *Communications in Statistics – Simulation and Computation*, **42**, 1063-1087. [78, 81, 82]
- PÉGUIN-FEISSOLLE, A., and T. TERÄSVIRTA (1999): “A General Framework for Testing the Granger Noncausality Hypothesis”. SSE/EFI Working Paper Series in Economics and Finance No. 343, Stockholm School of Economics. [78, 80, 81]
- SILVENNOINEN, A., and T. TERÄSVIRTA (2005): “Multivariate Autoregressive Conditional Heteroskedasticity with Smooth Transitions in Conditional Correlations”. SSE/EFI Working

Paper Series in Economics and Finance No. 577, Stockholm School of Economics, Sweden. [94]

SILVENNOINEN, A., and T. TERÄSVIRTA (2009a): "Multivariate GARCH Models". In *Handbook of Financial Time Series*, Ed. by T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch. New York: Springer, 201-229. [77]

SILVENNOINEN, A., and T. TERÄSVIRTA (2009b): "Modeling Multivariate Autoregressive Conditional Heteroskedasticity with the Double Smooth Transition Conditional Correlation GARCH Model", *Journal of Financial Econometrics*, **7**, 373-411. [78, 79]

SILVENNOINEN, A., and T. TERÄSVIRTA (2015): "Modelling Conditional Correlations of Asset Returns: A Smooth Transition Approach", *Econometric Reviews*, **34**, 174-197. [78, 79, 83, 84, 86, 92, 98]

STINCHCOMBE, M., and H. WHITE (1989): Universal Approximation Using Feedforward Networks with Non-sigmoid Hidden Layer Activation Functions. In *Proceedings of the international joint conference on neural networks*, Washington, D.C., IEEE Press, I, 613-618. [79]

TERÄSVIRTA, T., C. F. LIN, and C. W. J. GRANGER (1993): "Power of the Neural Network Linearity Test", *Journal of Time Series Analysis*, **14**, 209-220. [78]

TSE, Y. K. (2000): "A Test for Constant Correlations in a Multivariate GARCH Model", *Journal of Econometrics*, **98**, 107-127. [77, 78, 79, 83, 86]