2018/09

DP

Luc Bauwens and Edoardo Otranto

Nonlinearities and regimes in conditional correlations with different dynamics

CORE
CENTER FOR
OPERATIONS RESEARCH
AND ECONOMETRICS

CORE
Voie du Roman Pays 34, L1.03.01
Tel (32 10) 47 43 04
Fax (32 10) 47 43 01
Email: immaq-library@uclouvain.be
https://uclouvain.be/en/research-institutes/immaq/core/discussion-papers.html

CORE DISCUSSION PAPER
2018/09

Nonlinearities and Regimes in Conditional Correlations with Different Dynamics

Luc Bauwens
Université catholique de Louvain, CORE,
B-1348 Louvain-La-Neuve, Belgium
E-mail: luc.bauwens@uclouvain.be
Edoardo Otranto[1]
Dipartimento di Economia, Universitá di Messina, Italy
E-mail: eotranto@unime.it

March 19, 2018

Abstract

New parameterizations of the dynamic conditional correlation (DCC) model and of the regime-switching dynamic correlation (RSDC) model are introduced, such that these models provide a specific dynamics for each correlation. They imply a nonlinear autoregressive form of dependence on lagged correlations and are based on properties of the Hadamard exponential matrix. The new models are applied to a data set of twenty stock market indices, comparing them to the classical DCC and RSDC models. The empirical results show that the new models improve their classical versions in terms of several criteria.

Keywords: dynamic conditional correlations, regime-switching dynamic correlations, Hadamard exponential matrix.

JEL Classification: C32, C58.

[1]Corresponding author

1 Introduction

The increasing globalization of financial markets and the enduring quest of asset and risk management techniques have developed a large interest in multivariate volatility modeling of asset returns, resulting in a long sequence of proposals to represent the interdependence of financial markets. The seminal paper of Engle (2002) has triggered the development of models based on a two-step specification and estimation procedure, where the volatility of each single asset return is estimated in the first step within a univariate model, and the correlation model is estimated in the second step. Only the second step genuinely consists of a multivariate model linking a set of financial returns. The multivariate aspect, coupled with the related difficulty of estimation, justifies that many efforts were devoted to propose models featuring a small number of parameters (to save the feasibility of estimation) but able to capture the main stylized facts of the correlations. This evolution has enriched the models belonging to the GARCH family, working under the conditioning on the past information: the Constant Conditional Correlation (CCC) model of Bollerslev (1990) has been followed by the Dynamic Conditional Correlation (DCC) models of Engle (2002) and the Time-Varying Correlation (TVC) of Tse and Tsui (2002), which provide time-varying conditional correlations between assets, with a GARCH-type dynamics. Concurrently, the smooth transition (STCC) model of Silvennoinen and Teräsvirta (2015) and the Regime Switching Dynamic Correlation (RSDC) model of Pelletier (2006) were designed to capture the presence of smooth and abrupt changes in the correlation dynamics, respectively. All the previously mentioned models have been extended by Bauwens and Otranto (2016) to include the dependence of the correlations on the market volatility, resulting in the wide class of Volatility Dependent Conditional Correlation (VDCC) models.

A solid body of empirical evidence confirms the presence of regimes in the conditional correlations of financial markets, see e.g. Billio and Caporin (2005), Lee and Yoder (2007), Silvennoinen and Teräsvirta (2012), and Bauwens and Otranto (2016). In fact, it is linked to co-movements in volatility, which is frequently subjected to abrupt changes, as shown by Gallo and Otranto (2015 and 2018). The RSDC model of Pelletier (2006) is a simple solution to include regime changes in the correlation dynamics, but with some strong constraints on the parameters. Pelletier distinguishes between the case of at most three assets and the case of a larger number: only in the former it is possible to estimate different parameters that characterize the conditional correlation of each pair of assets, whereas in the latter it is necessary to assume common parameters. This is a strong constraint that typically limits the performance of the RSDC model. In this respect, the examples shown by Bauwens and Otranto (2016) are instructive: when they use three assets and estimate the most widespread conditional correlation models (CCC, DCC, STCC, RSDC) and their VDCC extensions without constraints on the correlation parameters, the RSDC family clearly outperforms all the others; when they use thirty stocks imposing a common dynamics to all the correlations, the models with a GARCH-type dynamics (DCC) have a better in-sample and out-of-sample forecasting performance.

To avoid imposing the constraints of a common dynamics on conditional correlations in large dimensions, new versions of the DCC model and of the RSDC model are introduced. These models provide a specific dynamics to each correlation. They use parameterizations implying a nonlinear autoregressive form of dependence on lagged correlations and are based on properties of the Hadamard exponential matrix. The parameterizations proposed for these new models ensure the positive definiteness of the conditional correlation matrix and in their simplest versions they require the estimation of one parameter more than the corresponding scalar forms of DCC (Engle, 2002; Aielli, 2013) and the RSDC (Pelletier, 2006). More flexible versions of the models are also available, with the introduction of a larger number of parameters, for which a general-to-specific procedure is proposed to identify a more parsimonious model. The new models, called the NonLinear AutoRegressive Correlation (NLARC) and the Flexible RSDC (FRSDC) models, are applied to a data set of twenty stock market indices, comparing them to Aielli's version of the DCC model (Aielli, 2013) and the RSDC model of Pelletier (2006). The empirical results show that the new models improve their classical versions in terms of statistical fit criteria and in-sample forecasting performance, while differences are small in out-of-sample forecast comparisons.

The paper is structured in five sections. Section 2 sets the modeling framework, reminding the DCC and the RSDC models and underlying their constraints. Section 3 introduces the NLARC and FRSDC models and their properties, describing also the possible parameterizations of the Hadamard exponential of the conditional correlation matrix. Section 4 illustrates the new models on real data, and compares them to the RSDC and DCC models both in in-sample and out-of-sample terms. Section 5 concludes the paper.

2 The Modeling Framework

Let us consider a set of time series of n asset returns at time t collected in the vector $\boldsymbol{y}_t$, and available for $t = 1, 2, \ldots, T$. Denoting by $\boldsymbol{\Psi}_t$ the information set containing all the values of the returns until time t, $\boldsymbol{y}_t|\boldsymbol{\Psi}_{t-1}$ is assumed to follow a multivariate Normal distribution with zero mean and covariance matrix $\boldsymbol{H}_t = \boldsymbol{S}_t\boldsymbol{R}_t\boldsymbol{S}_t$, where $\boldsymbol{S}_t$ is the diagonal matrix containing the conditional standard deviations of the returns and $\boldsymbol{R}_t = (\rho_{ij,t})$ is the positive definite (PD) matrix containing the conditional correlations between the returns.

The assumption that the conditional variance of each return depends on the past returns of the considered asset but not on those of the other assets is adopted, so that each element of $\boldsymbol{S}_t$ can be specified as a univariate GARCH model and each of these models can be estimated independently of the other. The first step estimates of the matrices $\boldsymbol{S}_t$ provide the 'devolatilized' (or 'degarched') residuals $\boldsymbol{u}_t = \boldsymbol{S}_t^{-1}\boldsymbol{y}_t$.

After the first step, a dynamic multivariate model for the $\boldsymbol{R}_t$ matrix can be specified and estimated in a second step, conditioning on the results of the first step. The second step estimation becomes prohibitive when n is large enough. The models proposed and widely used in the literature are specified in such a way that the correlation matrix depends on the past residuals $\boldsymbol{u}_t$ through a small number of parameters, to make the estimation feasible for large n.

The most widespread model is the DCC model of Engle (2002). Using the 'corrected

DCC‘ (cDCC) of Aielli (2013), it is given by the following equations:

$$\begin{aligned}\boldsymbol{R}_t &= \tilde{\boldsymbol{Q}}_t^{-1}\boldsymbol{Q}_t\tilde{\boldsymbol{Q}}_t^{-1},\\ \boldsymbol{Q}_t &= (\boldsymbol{J}_n - \boldsymbol{A} - \boldsymbol{B}) \odot \boldsymbol{C} + \boldsymbol{A} \odot \tilde{\boldsymbol{Q}}_{t-1}\boldsymbol{u}_{t-1}\boldsymbol{u}_{t-1}^{'}\tilde{\boldsymbol{Q}}_{t-1} + \boldsymbol{B} \odot \boldsymbol{Q}_{t-1},\\ \tilde{\boldsymbol{Q}}_t &= \text{diag}(\sqrt{q_{11,t}}\,\sqrt{q_{22,t}}, \ldots, \sqrt{q_{nn,t}})\end{aligned} \tag{2.1}$$

where $q_{ii,t}$ ($i = 1, \ldots, n$) are the diagonal elements of $\boldsymbol{Q}_t$, $\odot$ indicates the Hadamard (element-by-element) product, $\boldsymbol{J}_n$ is a square matrix with all entries equal to unity, and $\boldsymbol{C}$, $\boldsymbol{A}$ and $\boldsymbol{B}$ are square and symmetric parameter matrices of order n. These matrices, and also $\boldsymbol{J}_n - \boldsymbol{A} - \boldsymbol{B}$, must be positive semi-definite (PSD), and at least one of them must be PD, to ensure that $\boldsymbol{Q}_t$ be PD. In addition, the diagonal elements of $\boldsymbol{C}$ are set equal to unity for identification (see Aielli, 2013). This model has potentially $[3n(n+1)/2] - n$ parameters, involving nonlinear constraints (due to the positivity constraints), revealing the estimation difficulty for large n. When n is not small, $\boldsymbol{A}$ and $\boldsymbol{B}$ are restricted to depend on a small number of parameters. Engle (2002) adopts the scalar restrictions $\boldsymbol{A} = a\boldsymbol{J}_n$ and $\boldsymbol{B} = b\boldsymbol{J}_n$, where a and b are non-negative scalars constrained by $a + b < 1$, which is very convenient for estimation but may be considered to be too restrictive for large n.

More flexible, still practically feasible for estimation, alternative parameterizations are proposed by Billio et al. (2006) where each matrix $\boldsymbol{M}$ of parameters ($\boldsymbol{M} = \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$) is restricted to be a rank-one matrix defined as the outer product $\boldsymbol{m}\boldsymbol{m}'$, where $\boldsymbol{m}$ is a $n \times 1$ vector. Bauwens et al. (2016) extend this idea to rank-two matrices and extend the scalar model by constraining the elements of the vector $\boldsymbol{m}$ to lie on a polynomial of low degree. Another approach is to group the assets in a small number of clusters following the same dynamics (the same parameters in $\boldsymbol{m}$). Otranto (2010) proposes a clustering algorithm to detect these groups. Similar ideas are used for the nonlinear models exposed in Section 3.

An alternative to the GARCH-type dynamics of DCC models is to let the correlations remain constant during different regimes (sequences of periods of random lengths) but let their levels change between the different regimes. The RSDC model of Pelletier (2006) with two regimes is of this type, but it imposes the relative change of correlation to be the same for all correlations. More in detail, the parameterization of the RSDC model (named RSDC-1λ) proposed by Pelletier (2006) is:

$$\begin{aligned}&\boldsymbol{R}_t = \boldsymbol{R}_{s_t},\ s_t \in \{h, l\},\\ &\boldsymbol{R}_h = \bar{\boldsymbol{R}}, \quad \boldsymbol{R}_l = \bar{\boldsymbol{R}}\lambda_l + \boldsymbol{I}_n(1 - \lambda_l),\\ &\lambda_l \in [0, 1],\end{aligned} \tag{2.2}$$

where $\bar{\boldsymbol{R}} = (\bar{r}_{ij})$ is the sample correlation matrix of the residuals $\boldsymbol{u}_t$ ($t = 1, 2, \ldots, T$), and s_t is the unobserved regime indicator driven by a two-state Markov chain with transition probabilities $p_{gk} = Pr(s_t = k|s_{t-1} = g),\ g$ and $k \in \{h, l\}$, where h is the label of the high correlation regime and l the label of the low correlation regime. $\boldsymbol{R}_t$ is a PD correlation matrix, being equal to the sample correlation under the regime of high correlation, and a convex combination of two PD matrices under the low correlation regime.

In this model the relative variation between each element of $\boldsymbol{R}_l$ and the corresponding element of $\boldsymbol{R}_h$ is the same for all pairs of assets (being equal to $(1 - \lambda_l)/\lambda_l$), and likewise when the regime changes from h to l (being equal to $\lambda_l - 1$). This feature imposes a strong constraint on the model.

An alternative parameterization (RSDC-2λ) was proposed by Bauwens and Otranto (2016):

$$\begin{aligned} &\boldsymbol{R}_h = \bar{\boldsymbol{R}}\lambda_h + \boldsymbol{I}_n(1-\lambda_h), \quad \boldsymbol{R}_l = \bar{\boldsymbol{R}}\lambda_l + \boldsymbol{I}_n(1-\lambda_l), \\ &\lambda_l \in [0,1], \; \lambda_h \in [1, 1/\bar{r}_{max}), \end{aligned} \tag{2.3}$$

where $\bar{r}_{max}(>0)$ is the maximum correlation coefficient in $\bar{\boldsymbol{R}}$. This model is more flexible than RSDC-1λ because it allows each high correlation to exceed the corresponding sample correlation. To provide a PD $\boldsymbol{R}_h$ matrix, it requires the constraint that the smallest eigenvalue of $\bar{\boldsymbol{R}}$ be larger than $(\lambda_h - 1)/\lambda_h$, this being easily verified considering that the eigenvalues of $\boldsymbol{R}_h$ are equal to $(1-\lambda_h)$ plus the eigenvalues of $\bar{\boldsymbol{R}}$ multiplied by λ_h. This restriction is not very strong, in particular when $\bar{r}_{max}$ is close to unity. Anyway, this model also implies that the relative variation of each correlation coefficient, switching from state l to state h, is the same for all correlations (being equal to $(\lambda_h - \lambda_l)/\lambda_l$), and likewise, switching from state h to state l (being equal to $(\lambda_l - \lambda_h)/\lambda_h$).

In other words, the presence of regimes in the correlations between asset returns is a verified stylized fact, but the constraint of identical dynamics for all the correlations is a strong hypothesis that justifies our search for more flexible models.

3 Nonlinearities and Different Dynamics in Correlations

New models are proposed, where each correlation coefficient can have a specific dynamics or relative variation when changing regime. Nevertheless, these models can remain very parsimonious in parameters. The proposed models include a cDCC model with a nonlinear dependence of the conditional correlations on the past values of the correlations, and two parameterizations in the regime switching framework, corresponding to extensions of the models (2.2) and (2.3). These models involve, in their simplest specifications, one parameter more than the scalar cDCC in the first case, and the same number of parameters in the regime-switching models. Their formulation uses the entry-wise exponential operator of a matrix, which has the interesting property to preserve the positive definiteness of a positive (semi-)definite matrix. The first kind of model is named the NonLinear AutoRegressive Correlation (NLARC) model and the second the Flexible RSDC (FRSDC) model.

3.1 The Nonlinear Autoregressive Correlation Model

A feature of the cDCC model in its general version (2.1) is that each correlation process has its specific dynamic parameters (a_{ij} and b_{ij}), so that the dynamics of different correlations can vary considerably. This interesting property implies that the model becomes difficult, if not impossible, to estimate for large n since the number of parameters is $O(n^2)$. The scalar version of the model, where $\boldsymbol{A}$ is replaced by the scalar a and $\boldsymbol{B}$ by the scalar b, implies on the contrary that all correlations have the same dynamic properties and that the model is easy to estimate for large n (when $\boldsymbol{C}$ is replaced by a preliminary estimator). Although intermediate specifications (mentioned in Section 2) exist, a new one is proposed. It consists formally of the same equations as (2.1), with a time-varying, newly parameterized, matrix $\boldsymbol{A}$, and the scalar version of $\boldsymbol{B}$. In this version, the second equation of (2.1) is:

$$\boldsymbol{Q}_t = (1-a-b)\boldsymbol{C} + a\boldsymbol{A}_t \odot \tilde{\boldsymbol{Q}}_{t-1}\boldsymbol{u}_{t-1}\boldsymbol{u}_{t-1}^{'}\tilde{\boldsymbol{Q}}_{t-1} + b\boldsymbol{Q}_{t-1}, \tag{3.1}$$

where

$$\boldsymbol{A}_t = \exp^{\odot}\left[\phi_A\left(\boldsymbol{R}_{t-1} - \boldsymbol{J}_n\right)\right]. \tag{3.2}$$

For any matrix $\boldsymbol{M} = (m_{ij})$, $\exp^{\odot}(\boldsymbol{M}) = [\exp(m_{ij})]$ is the corresponding Hadamard exponential matrix, so $\exp^{\odot}$ is the entry-wise exponential operator. In (3.2), the three scalar parameters of the model are restricted by $0 \leq a < 1$, $0 \leq b < 1$, $b = 0$ if $a = 0$, $a + b < 1$, and $\phi_A \geq 0$. The entry-wise exponential transformation used to obtain $\boldsymbol{A}_t$ provides a symmetric PD matrix (see Proposition 1 below) with diagonal elements equal to 1 and non-negative off- diagonal elements smaller than 1. In (2.1), the matrix $\boldsymbol{B}$ could also be specified to be time–varying, but empirically this turned out not to be useful; this is consistent with the findings of Bauwens and Otranto (2016) and Clements et al. (2018), applying the DCC–TVV model proposed by Bauwens and Otranto (2016).

More explicitly, the dynamic equation for a diagonal element of (3.1) with parameterization (3.2) is $q_{ii,t} = (1-a-b) + a\, u_{i,t-1}^2 q_{ii,t-1} + bq_{ii,t-1}$ ($\boldsymbol{C}$ is unit-diagonal) . As shown by Aielli (2013), this implies that $E(q_{ii,t}) = 1$. For a covariance element $q_{ij,t}$ ($i \neq j$) the dynamic equation is:

$$q_{ij,t} = (1-a-b)c_{ij} + a\exp[\phi_A(\rho_{ij,t-1} - 1)]\; u_{i,t-1}u_{j,t-1}\sqrt{q_{ii,t-1}q_{jj,t-1}} + bq_{ij,t-1}. \tag{3.3}$$

This shows that two separate autoregressive dependences are considered: one of them is the linear dependence on the lagged covariance $q_{ij,t-1}$, with the same parameter b for all the covariances, as in the scalar cDCC model; the other one is embedded within the matrix $\boldsymbol{A}_t$ and adds a nonlinear dependence on the lagged conditional correlation, the latter being itself a linear function of $q_{ij,t-1}$. Computing the derivative of $q_{ij,t}$ with respect to $q_{ij,t-1}$ gives $a\phi_A \exp[\phi_A(\rho_{ij,t-1} - 1)]u_{i,t-1}u_{j,t-1} + b$, where the first term is the nonlinear and time-varying effect. This effect is illustrated graphically in the empirical study presented in Section 4 (see Figure 3).

The impact of the lagged covariance shock $u_{i,t-1}u_{j,t-1}\sqrt{q_{ii,t-1}q_{jj,t-1}}$ on the next covariance $q_{ij,t}$ is given by $a_{ij,t} = a\exp[\phi_A(\rho_{ij,t-1} - 1)] \in [0, 1)$. Thus it is both time-varying and asset-pair specific through the lagged correlation $\rho_{ij,t-1}$. Figure 1 illustrates the estimated time series $a_{ij,t}$ for three asset pairs of the empirical study. The three lines have different patterns even though they come from a single estimated NLARC model (see Section 4.1). The corresponding impact for the cDCC model is constant and lower. In the NLARC model, a given positive (negative) lagged covariance shock increases (decreases) more the next covariance when the lagged correlation is strong than when it is weak. Figure 2 shows several impact curves $a_{ij,t} = 0.1\exp[\phi_A(\rho_{ij,t-1} - 1)]$ as functions of $\rho_{ij,t-1}$ for values of ϕ_A ranging from 0 to ∞. The impact function is flat (at $a = 0.1$, which is a scale coefficient and does not affect the general shape of the curves) for $\phi_A = 0$ and a step function for ϕ_A approaching infinity; for positive values of ϕ_A, it is convex, and increasing ϕ_A increases the convexity.

Moreover, $a_{ij,t} + b$ is strictly less than 1 for each pair (i, j) such that $i \neq j$ (because $\rho_{ij,t-1} < 1$ and $\phi_A \geq 0$). Thus the proposed parameterization satisfies *at each* t one of the sufficient conditions for stationarity ($|a_{ij} + b_{ij}| < 1$) of the general cDCC model (2.1).

The proposed parameterization (3.2) for the cDCC model (3.1) satisfies the required property that the resulting $\boldsymbol{Q}_t$ matrix be PD for all t. This result is based on the following proposition.

Proposition 1: *If $\boldsymbol{D}$ is a PSD correlation matrix, then $\boldsymbol{F} = \exp^{\odot}\left[\delta\left(\boldsymbol{D} - \boldsymbol{J}_n\right)\right]$, is PD if $\delta \geq 0$.*

This proposition is based on the property according to which the entry-wise exponential function preserves the positive (semi-)definiteness of a matrix. Rewriting $\boldsymbol{F}$ as $\exp^{\odot}\left(\delta \boldsymbol{D}\right) / \exp(\delta)$, since δ is non-negative, the argument of the entry-wise exponential is again PSD and the result is divided by a positive constant, so that $\boldsymbol{F}$ is PSD. Moreover, another property of the entry-wise exponential of a PSD matrix establishes that it is PD if no two rows of the matrix are identical (see Theorem 7.5.9 (c) in Horn and Johnson, 2013); hence if $\boldsymbol{D}$ is a PD correlation matrix and its off-diagonal elements are strictly less than 1, then all its rows are different, so that $\boldsymbol{F}$ is PD.

In (3.2), the matrix $\boldsymbol{R}_{t-1}$ is a PSD, and in practice PD, correlation matrix, so that Proposition 1 can be applied to $\boldsymbol{A}_t$.

The parameterization (3.2) of $\boldsymbol{A}_t$ is named the *scalar parameterization*; it is very parsimonious, involving just one parameter. More flexibility can be reached by associating different parameters to each element of the correlation matrix $\boldsymbol{R}_{t-1}$. A natural extension of (3.2), is given by:

$$\boldsymbol{A}_t = \exp^{\odot}\left[\boldsymbol{\Phi}_A \odot \left(\boldsymbol{R}_{t-1} - \boldsymbol{J}_n\right)\right], \tag{3.4}$$

where $\boldsymbol{\Phi}_A$ is a square symmetric matrix with strictly non-negative entries; on the diagonal, any fixed constant can be chosen because the elements on the diagonal of $\boldsymbol{A}_t$ are equal to 1 by construction. A nice property of the parameterization (3.4) is given by the following proposition.

Proposition 2: *If $\boldsymbol{D}$ is a PSD correlation matrix and $\boldsymbol{\Delta}$ is a symmetric PSD matrix with all strictly positive entries and just one positive eigenvalue, then $\boldsymbol{F} = exp^{\odot}\left[\boldsymbol{\Delta} \odot \left(\boldsymbol{D} - \boldsymbol{J}_n\right)\right]$ is PD.*

The matrix $\boldsymbol{F}$ is actually the Hadamard product of two entry-wise exponential functions:

$$\boldsymbol{F} = \exp^{\odot}\left[\boldsymbol{\Delta} \odot \boldsymbol{D}\right] \odot \exp^{\odot}\left[\left(\boldsymbol{\Delta}\right)^{\odot(-1)}\right], \tag{3.5}$$

where $^{\odot(-1)}$ indicates the Hadamard inverse, so that the element (i, j) of$\boldsymbol{\Delta}^{\odot(-1)}$ is equal to $1/\delta_{ij}$ when δ_{ij} is the element (i, j) of $\boldsymbol{\Delta}$. $\boldsymbol{\Delta}$ is PSD and, from Schur's theorem, its Hadamard product with $\boldsymbol{D}$ is PSD since $\boldsymbol{D}$ is PSD. As a consequence $\exp^{\odot}\left[\boldsymbol{\Delta} \odot \boldsymbol{D}\right]$ is PD because it is the entry-wise exponential of a PSD matrix with distinct rows. Bapat (1988) has proven that the Hadamard inverse of a symmetric matrix with all positive entries and just one positive eigenvalue is PSD. This implies that, under the not restrictive hypothesis of Proposition 2, $\left(\boldsymbol{\Delta}\right)^{\odot(-1)}$ is PSD, so that its entry-wise exponential transformation is PSD. So $\boldsymbol{F}$ is the Hadamard product of a PD matrix and a PSD matrix having non-zero diagonal entries. By Lemma 2.2 of Reams (1999), this type of Hadamard product provides a PD matrix.

Proposition 2 implies that, if $\boldsymbol{\Phi}_A$ has just one positive eigenvalue, then $\boldsymbol{A}_t$ is PD. Setting $\boldsymbol{\Phi}_A = \boldsymbol{\phi}_A \boldsymbol{\phi}'_A$, where $\boldsymbol{\phi}_A$ is a vector of n strictly positive elements, the resulting matrix satisfies the conditions of Proposition 2. The parameterization (3.4) with the matrix defined in the previous sentence is called the *rank–1 parameterization*.

3.2 The Flexible Regime Switching Dynamic Correlation Model

The extension of the model (2.2) of Pelletier (2006) to the FRSDC case (FRSDC-1λ) is simple and given by:

$$\begin{aligned}
&\boldsymbol{R}_t = \boldsymbol{R}_t^{(s_t)},\ s_t \in \{h, l\}, \\
&\boldsymbol{R}_t^{(h)} = \bar{\boldsymbol{R}}, \quad \boldsymbol{R}_t^{(l)} = \bar{\boldsymbol{R}} \odot \boldsymbol{\Lambda}_t^{(l)}, \\
&\boldsymbol{\Lambda}_t^{(l)} = \exp^{\odot}\left[\phi^{(l)}\left(\boldsymbol{R}_{t-1} - \boldsymbol{J}_n\right)\right] \\
&p_{gk} = Pr(s_t = k | s_{t-1} = g),\ g \text{ and } k \in \{h, l\},\ p_{kk} = 1 - p_{gk} \text{ if } g \neq k
\end{aligned} \tag{3.6}$$

where $\bar{\boldsymbol{R}}$ is the sample correlation matrix of the residuals $\boldsymbol{u}_t$ ($t = 1, 2, \ldots, T$), assumed to be PD, $\phi^{(l)}$ is a scalar parameter restricted to be strictly positive, and $\boldsymbol{\Lambda}_t^{(l)}$ is a time-varying symmetric PD matrix (by Proposition 1) with all its diagonal elements equal to 1 and other elements $\lambda_{ij,t}^{(l)} = \exp\left[\phi^{(l)}(\rho_{ij,t-1} - 1)\right]$ in $[0, 1]$. The resulting $\boldsymbol{R}_t^{(l)}$ matrix is a correlation matrix since it is obtained as the Hadamard product of two PD matrices with the characteristics of correlation matrices (ones on the diagonal and less or equal to 1 in absolute value otherwise). The last formula of (3.6) implies that each element $\rho_{ij,t}$ of the correlation matrix $\boldsymbol{R}_t$ follows a two-state Markov Switching model with a time invariant transition probability matrix that is common to all correlations.

The first three equations of the model (3.6), written for each element (i, j) of the correlation matrices, are:

$$\begin{aligned}
&\rho_{ij,t} = \rho_{s_t,t},\ s_t \in \{h, l\}, \\
&\rho_{ij,t}^{(h)} = \bar{r}_{ij}, \quad \rho_{ij,t}^{(l)} = \bar{r}_{ij}\lambda_{ij,t}^{(l)} = \bar{r}_{ij} \exp\left[\phi^{(l)}(\rho_{ij,t-1} - 1)\right].
\end{aligned} \tag{3.7}$$

This model implies constant correlations under the high correlation regime, but dynamic correlations under the low correlation regime; the classification high/low is justified by the inequality $\rho_{ij,t}^{(l)} \leq \rho_{ij,t}^{(h)}$ at each time t. The dynamics of the low correlation has a nonlinear autoregressive structure of order 1. The main novelty with respect to the specification (2.2) is that, even if all the correlations are in the same regime at each time, the relative variations differ at each time and for each pair of assets, being equal to $(1 - \lambda_{ij,t}^{(l)})/\lambda_{ij,t}^{(l)}$ when the regime switches from l to h, and equal to $\lambda_{ij,t}^{(l)} - 1$ when the regime changes from h to l.

A more flexible specification (FRSDC-2λ) provides a dynamic structure also for the high correlation matrix. It is given by:

$$\begin{aligned}
&\boldsymbol{R}_t = \boldsymbol{R}_t^{(s_t)},\ s_t \in \{h, l\}, \\
&\boldsymbol{R}_t^{(h)} = \boldsymbol{W}_t - (1 - \theta)\boldsymbol{I}_n, \quad \boldsymbol{R}_t^{(l)} = \bar{\boldsymbol{R}} \odot \boldsymbol{\Lambda}_t^{(l)}, \\
&\boldsymbol{W}_t = \bar{\boldsymbol{R}} + \boldsymbol{R}^* \odot \boldsymbol{\Lambda}_t^{(h)}, \\
&\boldsymbol{R}^* = S\left(\boldsymbol{J_n} - \bar{\boldsymbol{R}}\right), \\
&\boldsymbol{\Lambda}_t^{(s_t)} = \exp^{\odot}\left[\phi^{(s_t)}\left(\boldsymbol{R}_{t-1} - \boldsymbol{J}_n\right)\right], \\
&p_{gk} = Pr(s_t = k | s_{t-1} = g),\ g \text{ and } k \in \{h, l\},\ p_{kk} = 1 - p_{gk} \text{ if } g \neq k
\end{aligned} \tag{3.8}$$

where $S(\cdot)$ is a shrinking function that transforms a symmetric matrix into a PD matrix. A simple specification of $S(\boldsymbol{A})$ is a convex combination of $\boldsymbol{A}$ and the identity matrix $\boldsymbol{I}_n$:

$$S\left(\boldsymbol{A}\right) = \theta\boldsymbol{A} + (1 - \theta)\boldsymbol{I}_n, \tag{3.9}$$

where θ is chosen to be the highest value in $(0, 1]$ such that all eigenvalues of $S(\boldsymbol{A})$ are positive (see Devlin et al., 1975; other techniques are illustrated in Rousseeuw and Molenberghs, 1993). Subtracting $(1-\theta)\boldsymbol{I}_n$ from $\boldsymbol{W}_t$ in the second formula of (3.8) is necessary to render the diagonal elements of $\boldsymbol{R}_t^{(h)}$ equal to 1.

The matrix $\boldsymbol{R}_t^{(l)}$ is PD by construction (as in the FRSDC-1λ model), whereas $\boldsymbol{R}_t^{(h)}$ is PD under the constraint expressed in the following proposition.

Proposition 3: *$\boldsymbol{R}_t^{(h)}$ is PD if the smallest eigenvalue of $\boldsymbol{W}_t$ is strictly greater than $1-\theta$.*

Let $\boldsymbol{V}_t$ be the orthonormal matrix of eigenvectors of $\boldsymbol{W}_t$ and $\boldsymbol{L}_t$ the diagonal matrix of associated eigenvalues (all positive and different). By definition, $\boldsymbol{W}_t = \boldsymbol{V}_t\boldsymbol{L}_t\boldsymbol{V}_t'$ and $\boldsymbol{V}_t'\boldsymbol{V}_t = \boldsymbol{V}_t\boldsymbol{V}_t' = \boldsymbol{I}_n$. As a consequence:

$$\begin{aligned} \boldsymbol{V}_t'\boldsymbol{R}_t^{(h)}\boldsymbol{V}_t &= \boldsymbol{V}_t'\boldsymbol{W}_t\boldsymbol{V}_t - (1-\theta)\boldsymbol{V}_t'\boldsymbol{I}_n\boldsymbol{V}_t \\ &= \boldsymbol{V}_t'\boldsymbol{V}_t\boldsymbol{L}_t\boldsymbol{V}_t\boldsymbol{V}_t' - (1-\theta)\boldsymbol{V}_t'\boldsymbol{V}_t = \boldsymbol{L}_t - (1-\theta)\boldsymbol{I}_n. \end{aligned} \tag{3.10}$$

The last matrix of (3.10) is a diagonal matrix. It is obviously PD if the smallest eigenvalue of $\boldsymbol{W}_t$ is strictly larger than $1-\theta$. Under this condition, $\boldsymbol{V}_t'\boldsymbol{R}_t^{(h)}\boldsymbol{V}_t$ is thus PD, and $\boldsymbol{R}_t^{(h)}$ also (since $\boldsymbol{V}_t$ has full rank), so the proposition is proven.

From (3.8) each element $\rho_{ij,t}^{(l)}$ is smaller than or equal to the corresponding element $\bar{r}_{ij}$ of the sample correlation matrix. Each off-diagonal element of the high correlation matrix $\boldsymbol{R}_t^{(h)}$ is larger than the corresponding element of the sample correlation matrix, since it is equal to $\bar{r}_{ij} + \theta(1-\bar{r}_{ij})\exp\left(\phi^{(h)}(\rho_{ij,t-1}-1)\right)$ where the second term is positive. Like the low correlations, the high correlations follow a nonlinear autoregressive process of order 1, instead of being constant like in (3.6).

The scalar parameterizations of $\boldsymbol{\Lambda}_t^{(s_t)}$ in (3.6) and (3.8) can be extended to a more general one, which uses rank-one matrices (a *rank–1 parameterization*), as exposed in the next sub-section.

3.3 Groupwise Formulation

The scalar models defined in (3.1)-(3.2), (3.6) and (3.8) provide a practical way to limit the number of parameters, while ensuring the positive definiteness of the corresponding conditional correlation matrices. On the other hand they introduce constraints that may be considered strong, especially when n is very large. A more flexible solution is the rank–1 formulation, which provides also PD correlation matrices but with a larger yet manageable number (of $O(n)$) of parameters. A middle ground can be achieved by a *groupwise formulation*, which is a rank–1 parameterization where groups of correlations with similar dynamics are formed. Of course the detection of groups is a difficult task. A possible solution consists in adopting model-based clustering algorithms (such as Otranto, 2010, for DCC models), where a parameter-dependent distance measure is used to identify similar correlation dynamics. Extending the Otranto (2010) algorithm to the FRSDC model does not bring a successful outcome. Alternatively some a priori information about assets could be used to create groups of assets characterized by a similar correlation structure (as in Billio et al., 2006, for DCC models), but of course this approach is subjective.

If the number of assets is not huge (say less than thirty), a heuristic general-to-specific search procedure to obtain a parsimonious version of the rank–1 models is can be applied. For the FRSDC-2λ case, let us consider a rank–1 version of $\mathbf{\Phi}^{(s_t)}$:

$$\mathbf{\Phi}^{(s_t)} = \boldsymbol{\phi}^{(s_t)}\boldsymbol{\phi}^{(s_t)\prime} \tag{3.11}$$

where $\boldsymbol{\phi}^{(s_t)}$ is a $n \times 1$ vector of strictly positive elements $\phi_i^{(s_t)}$ $(i = 1, \ldots, n)$. By Proposition 2, this formulation provides a PD matrix $\mathbf{\Lambda}_t^{(s_t)}$. Then the search procedure to reduce the number of parameters from $2n$ to $2k$ with $k \leq n$ involves the following steps:

1. Estimate the FRSDC-2λ model with the parameterization (3.11).
2. Compute the p-value of the joint null hypothesis $H_0 : \phi_i^{(h)} = \phi_j^{(h)}, \quad \phi_i^{(l)} = \phi_j^{(l)}$ for each pair i, j $(i \neq j)$. This test can be done using a Wald statistic having asymptotically a $\chi^2(2)$ distribution under H_0.
3. Select all the pairs for which H_0 is not rejected with highest p-values (for example more than 0.9) and set $\phi_i^{(h)} = \phi_j^{(h)}$ and $\phi_i^{(l)} = \phi_j^{(l)}$ for these pairs. Of course it is necessary to check if the different results are consistent: for example if the null hypotheses $\phi_1^{(s_t)} = \phi_2^{(s_t)}$ and $\phi_2^{(s_t)} = \phi_3^{(s_t)}$ are accepted with a large p-value, the restriction $\phi_1^{(s_t)} = \phi_2^{(s_t)} = \phi_3^{(s_t)}$ is adopted only if the p-value of the statistic for $\phi_1^{(s_t)} = \phi_3^{(s_t)}$ is also large.
4. Estimate the model imposing the constraints obtained in step 3.
5. If the constrained model of the previous step is preferred to the initial model in terms of a loss function (for example BIC or AIC), compute the p-values of the null hypothesis (as defined in step 2) for all the pairs of parameters in $\boldsymbol{\phi}^{(s_t)}$. Select the new pairs of parameters with highest p-value for which the null hypothesis is not rejected. Obviously, if no pair is selected, stop the procedure.
6. Constrain the selected pairs of parameters of the previous step to be equal and estimate the new constrained model.
7. Repeat steps 5 and 6 until the loss function no longer decreases or all the null hypotheses are rejected.

This procedure partitions each vector $\boldsymbol{\phi}^{(h)}$ and $\boldsymbol{\phi}^{(h)}$ into k groups of parameters; the grouping is the same in both vectors, the elements of a group of $\boldsymbol{\phi}^{(h)}$ are the same, as are those of a group in $\boldsymbol{\phi}^{(l)}$, and the parameters of the same group in $\boldsymbol{\phi}^{(h)}$ and $\boldsymbol{\phi}^{(l)}$ differ.

For the NLARC model, the procedure is the same, but substituting $\mathbf{\Phi}_A$ for $\mathbf{\Phi}^{(l)}$. For the FRSDC-1λ model, the null hypothesis considered in step 2 involves only the vector $\boldsymbol{\phi}^{(l)}$.

4 Empirical Findings

It is instructive to apply the previous models to a real data set, following some steps to come to the best specifications of the NLARC and FRDSC models. Twenty daily series of stock indices have been downloaded from the Oxford-Man Institute Realized Library version 0.2 (Heber et al., 2009): S&P 500 (abbreviated to SP), FTSE 100 (FTSE), Nikkei 225 (NIK), DAX (DAX), Russell 2000 (RUS), All Ordinaries (AO), Dow Jones Industrial Average (DJ), Nasdaq 100 (NAS), CAC 40 (CAC), Hang Seng (HS), KOSPI Composite Index (KOS), AEX Index (AEX), Swiss Market Index (SMI), IBEX 35 (IBEX), S&P CNX Nifty (CNX), IPC Mexico (IPC), Bovespa Index (BOV), S&P/TSX Composite Index (TSX), Euro STOXX 50 (EU), FTSE MIB (MIB). The time span starts the 8th of July 2002 and ends the 27th of April 2017. This provides 2555 daily observations for each series, keeping only the dates where all the indices are recorded. The series of the degarched returns $\boldsymbol{u}_t = \boldsymbol{S}_t^{-1}\boldsymbol{y}_t$ $(t = 1, 2, \ldots, 2554)$ have been obtained after estimating the conditional variances of the log-returns by univariate GARCH-GJR(1,1) models (Glosten et al.,1993).

The estimated models are: scalar cDCC (see Section 2), NLARC (see Section 3.1), the RSDC-1λ (see eq. (2.2)), RSDC-2λ (eq. (2.3)), FRSDC-1λ (eq.(3.6)) and FRSDC-2λ (eq. (3.8)). The NLARC and FRSDC-2λ models have been estimated both in the scalar, the rank–1 and the groupwise versions.

4.1 Estimation Results for cDCC and NLARC

The estimation results for the scalar cDCC model and NLARC models are shown in Table 1. These results were obtained by maximizing the second step log-likelihood function with respect to the remaining parameters, after substituting the sample correlation $\bar{\boldsymbol{R}}$ for the constant $\boldsymbol{C}$. Estimating $\boldsymbol{C}$ at the second step involves 190 additional parameters and is unfeasible. Nonscalar versions of the cDCC model have also been estimated, using the rank–1parameterization proposed in Billio et al. (2006). First, the procedure of Otranto (2010) was applied to detect series with similar cDCC dynamics to reduce the number of parameters, but this identifies a single group of twenty series. Next a cDCC model with a rank–1 parameterization of the matrices $\boldsymbol{C}$ and $\boldsymbol{A}$, and another with also a rank–1 $\boldsymbol{B}$ matrix, were estimated. In both cases, the estimated elements of the matrices $\boldsymbol{C}$ and $\boldsymbol{A}$ are all close to zero and in the second case those of the matrix $\boldsymbol{B}$ are comprised in a small range (between 0.93 and 1). Moreover the BIC of the most general model is equal to 1.888, which is higher than for the scalar model (1.866). Given this evidence, the scalar cDCC model is adopted as the benchmark for this study. The estimates show a small $a = 0.006$ coefficient (with asymptotic standard error 0.001) and a large $b = 0.977$; they are comparable to the values found in other empirical studies.

The scalar NLARC shows a clear increase in the log-likelihood and a decrease both in AIC and BIC with respect to the scalar cDCC model. Since the NLARC model simplifies to the cDCC if $\phi_A = 0$, this null hypothesis can be tested with a likelihood ratio (LR) test, but since the restriction $\phi_A = 0$ is on the boundary of the parameter space, the LR statistic does not follow the standard $\chi^2(1)$ asymptotic distribution. However, the value of the statistic (60.04) seems sufficiently large to provide some evidence in favor of the NLARC against the cDCC.

Figure 3 illustrates the nonlinear time-varying part of the first derivative of $q_{ij,t}$ with respect to $q_{ij,t-1}$ in the estimated scalar NLARC model, for i corresponding to SP and j to the other 19 stock indices. The time-varying effect, given by
$0.008(1.157)\exp[1.157(\rho_{ij,t-1} - 1)]u_{i,t-1}u_{j,t-1}$, is the effect added by the NLARC specification with respect to the cDCC one, in which the effect is equal to zero. The time-invariant effect is given by $b = 0.975$ (0.977 in the cDCC). The time-varying one is most often positive, though it is occasionally negative, which happens when $u_{i,t-1}u_{j,t-1} < 0$, and it can vary substantially. The graphs of the time-varying effects show some similar features for some groups of correlations, in particular, the strong positive effects on 28/2/2007 (beginning of the Chinese stock bubble) and 16/8/2011 (beginning of the European contagion) seem common to most of the indices (exceptions are IBEX, EU, HS, DJ). Other relevant shocks are:
-10/10/2008: positive effect after the collapse of several US banks (DJ, HS, EU, IBEX);
-27/6/2016, some days after the Brexit referendum: all the indices (except IPC) are affected, some with a negative effect (CNX, MIB,AO, CAC, DAX, SMI, KOS, TSX);
-10/11/2016, after the US election: negative effect for DJ, HS, IBEX, EU.

For the rank–1 version of the NLARC model, the number of groups detected by the algorithm, described in subsection 3.3, is just one, in the sense that all the coefficients of the vector $\boldsymbol{\phi}_A$ are not significantly different. In practice the groupwise specification is identical to the scalar one, a very parsimonious specification. This is confirmed by a comparison of the AIC and BIC of the scalar and the rank–1 specifications: they are equal to 1.45 and 1.46 respectively for the scalar NLARC (see Table 1), whereas they are 1.45 and 1.50 respectively for the rank–1 NLARC. The Wald statistic of the hypothesis that the twenty different parameters in $\boldsymbol{\phi}_A$ are equal (resulting in 19 constraints) has a p-value of 0.02, not rejecting the null at the 1% nominal size (this test is standard).

4.2 Estimation Results for RSDC and FRSDC

The estimation results of the RSDC and FRSDC models are shown in Table 2. The 2-λ models improve the 1-λ in terms of AIC and BIC. Moreover, it is possible to test, using the LR statistic, the (F)RSDC-1λ model against the corresponding 2-λ, by defining the null hypothesis as $\lambda_h = 1$ for the RSDC 2-λ model and $\phi^{(h)} = 0$ for the scalar FRSDC 2-λ model. In the first case, the test is standard and the null is rejected with a p-value very close to zero. In the second case the null is on the frontier but the statistic suggests that it can be rejected. For this reason only the 2-λ version of the rank–1 and groupwise parameterization of FRSDC are reported. The estimates of the parameters of the FRSDC 2-λ with rank–1 parameterization are shown in Figure 4; several estimated parameters seem similar and some are not significantly different from zero. The grouping algorithm determines three groups:

Group 1: SP, RUS, DJ, NAS, IPC, TSX;

Group 2: FTSE, DAX, CAC, AEX, SMI, IBEX, EU, MIB;

Group 3: NIK, AO, HS, KOS, CNX, BOV

It is interesting to notice that the three groups correspond to natural geographical partitions: the first group collects the North American indices, the second group the European indices, the third one all the rest (Asian, Australian and Brazilian markets).

The comparison in terms of AIC shows a very similar fitting of the groupwise and the rank–1 parameterizations, whereas the BIC clearly favors the more parsimonious groupwise representation.

In terms of AIC and BIC, the correlation models belonging to the Markow Switching family of Table 2 outperform the cDCC and NLARC models of Table 1, and all the 2-λ models outperform the 1-λ models.

The probability to stay in the high correlation regime and, as a consequence, the persistence of that regime, is higher in the RSDC models than in the FRSDC models, the latter involving as a consequence more frequent changes from high to low correlations. The inverse occurs for changes from low to high for the 2–λ case.

4.3 In-sample Forecasting

A complementary comparison among non nested models can be made by evaluating the in-sample and out-of-sample forecasting performance of the models, using a statistical criterion and an economic one. The former is the Model Confidence Set (MCS) of Hansen et al. (2003), with the purpose to detect a set of models having the best forecasting performance. This is done in the way proposed by Becker et al. (2015), by adopting the Quasi-Likelihood loss function with the semi-quadratic statistics. The important finding of this experiment is that the set of models with the best in-sample forecasting performance is given by the FRSDC-2λ models, which outperform the RSDC models and the FRSDC-1λ (upper part of Table 3). The approach excludes first the models belonging to the cDCC family; it is interesting to notice that the models with 1-λ parameterization are excluded before the RSDC 2-λ.

The economic criterion is based on theoretical portfolio performances, following the approach of Engle and Colacito (2006). The purpose is to compare the sample volatility of portfolios with weights depending on the alternative correlation matrices, having the same variances (obtained in step 1 of the estimation procedure), and fixing the expected returns equal to $1/\sqrt{n}$ (so the corresponding vector has length equal to one). In practice the difference between the portfolios depends only on the correlation matrix, and the model with the smallest portfolio volatility is the best model in economic terms. In the second part of Table 3 the result is the opposite with respect to the MCS experiment; the best model is the scalar NLARC and the cDCC family outperforms the RSDC family which provides an increase of 1 to 3$\%$ in the portfolio variance. The bottom part of Table 3 reports the results of the Diebold-Mariano (DM) test proposed by Engle and Colacito (2006) to verify the equality of each pair of models, comparing the differences between the squared returns within each pair of portfolios. Using the 5$\%$ significance level, the scalar NLARC model, which provides the minimum variance portfolio, and the cDCC model outperform only the FRSDC–2λ models. The groupwise FRSDC–2λ model does not provide a significantly different performance with respect to the scalar version.

In summary, the in-sample forecast comparisons indicate that the two alternative families have different performances if evaluated with statistical or economic criteria, favoring the RSDC family in the former, the cDCC family in the latter.

4.4 Out–of–sample Forecasting

The out–of–sample performance of the models is evaluated by generating one-step ahead forecasts of the correlation matrices for the last 441 days of the sample. Thus the estimation sample ends on the 30th of December 2014 and the estimated parameters of each model are frozen in the forecast sample until the end (27th of April 2017). The forecasts are used to compute the MCS and MVP as in the in-sample comparisons.

The results in Table 4 show that the MCS includes all the FRSDC and RSDC models (with p-values over 0.27), whereas the NLARC and cDCC models are excluded by this procedure.

In terms of MVP evaluation, an opposite order of the best performing models is observed: NLARC (closely followed by cDCC) provides the smallest portfolio minimum variance, the FRSDC–2λ models (scalar and groupwise) provide the largest values (5% and 8% higher), and the other models provide variances between 2 and 3% higher than NLARC. According to the DM tests, NLARC performs significantly better than the other models at the 2.5% significance level, and than FRSDC–1λ at 10%.

To check the importance of fixing the estimates at the end of the estimation sample, the parameter estimates were updated every fifth day (updating every day was too costly) during the forecast period before computing the next five forecasts. Table 5 contains the results. The MCS includes all models, meaning that their forecast performances are not significantly different at usual nominal sizes. In terms of MVP, the models also do not show significantly different performances.

5 Concluding Remarks

The proposed new versions of the DCC and RSDC models enrich the existing versions by allowing the correlation dynamics to be different for different pairs of assets, without implying a big increase in the parameter dimension. Indeed, in the scalar versions, the NLARC model has one parameter more than cDCC, and the flexible RSDC models have the same number of parameters as the corresponding less flexible versions. This is achieved by parameterizing some matrices as the Hadamard exponential transformations of a lagged correlation matrix. The Hadamard exponential of a matrix has some mathematical properties that might prove useful in other modeling contexts. The empirical results presented in Section 4, though contingent on the data set used, illustrate that the new parameterizations may be useful in practice.

Acknowledgments

This research was started during the Semester ‘Stochastic Dynamical Models in Mathematical Finance, Econometrics, and Actuarial Sciences’ (January-June 2017) held at the ‘Centre Interunivesitaire Bernouilli’ (CIB) of the Ecole Polytechnique Fédérale de Lausanne. Both authors are grateful to the CIB for funding their stay in Lausanne and providing an excellent research infrastructure.

References

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

[2] Bapat, R.B. (1988). Multinomial probabilities, permanents and a conjecture of Karlin and Rinott. *Proceedings of the American Mathematical Society* 102, 467–472.

[3] Bauwens, L., Grigoryeva, L., Ortega, J.–P. (2016). Estimation and empirical performance of non–scalar dynamic conditional correlation models, *Computational Statistics and Data Analysis*, 100, 17–36.

[4] Bauwens, L., Otranto, E. (2016). Modeling the dependence of conditional correlations on market volatility, *Journal of Business & Economic Statistics*, 34, 254–268.

[5] Becker, R., Clements, A., Doolan, M., Hurn, S. (2015). Selecting volatility forecasting models for portfolio allocation purposes, *International Journal of Forecasting*, 31, 849–861.

[6] Billio, M., Caporin, M. (2005). Multivariate Markov switching dynamic conditional correlation GARCH representations for contagion analysis, *Statistical Methods & Applications* 14, 145–161.

[7] Billio, M., Caporin, M., Gobbo, M. (2006). Flexible dynamic conditional correlation multivariate GARCH models for asset allocation, *Applied Financial Economics Letters* 2, 123–130.

[8] Bollerslev, T. (1990), Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach. *Review of Economics and Statistics* 72, 498–505.

[9] Clements, A., Scott, A., Silvennoinen, A. (2018). Volatility dependent correlations: further evidence of when, where and how. *Mimeo*.

[10] Devlin, S.J., Gnanadesikan, G., Kettenring, J.R. (1975). Robust estimation and outlier detection with correlation coefficients. *Biometrika* 62, 531–545.

[11] Ding, Z., Engle, R.F. (2001). Large scale conditional covariance matrix modeling, estimation and testing, *Academia Economic Papers* 29, 157–184.

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

[13] Engle, R., Colacito, R. (2006). Testing and valuing dynamic correlation for asset allocation, *Journal of Business and Economic Statistics*, 24, 238–253.

[14] Gallo, G.M., Otranto, E. (2015), Forecasting realized volatility with changing average volatility levels, *International Journal of Forecasting*, 31, 620–634.

[15] Gallo, G.M., Otranto, E. (2018), Combining sharp and smooth transitions in volatility dynamics: a fuzzy regime approach, *Journal of the Royal Statistical Society, Series C, Applied Statistics*, 67, 549–573.

[16] Glosten, L., Jagannathan, R., Runkle, D. (1993). On the relation between expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* 48, 1779–1801.

[17] Hansen, P., Lunde, A., Nason, J. (2003). Choosing the best volatility models: the Model Confidence Set approach, *Oxford Bulletin of Economic and Statistics*, 65, 839–861.

[18] Heber, G., Lunde, A., Shephard, N., Sheppard, K. (2009). Omi' s realised library, version 0.1. *Technical report*, Oxford-Man Institute, University of Oxford.

[19] Horn, R.A., Johnson, C.R. (2013). *Matrix Analysis*, 2nd Edition, Cambridge University Press.

[20] Lee, H. T., Yoder, J. K. (2007). Optimal hedging with a regime–switching time–varying correlation GARCH model, *The Journal of Futures Markets* 27, 495–516.

[21] Otranto, E. (2010). Identifying financial time series with similar dynamic conditional correlation. *Computational Statistics and Data Analysis* 54, 1–15.

[22] Pelletier, D. (2006), Regime switching for dynamic correlations, *Journal of Econometrics*, 131, 445–473.

[23] Reams, R. (1999). Hadamard inverses, square roots and products of almost semidefinite matrices. *Linear Algebra and its Applications* 288, 35–43.

[24] Rousseeuw, P., Molenberghs, G. (1993). Transformation of non positive semidefinite correlation matrices. *Communications in Statistics: Theory and Methods* 22, 965–984.

[25] Silvennoinen, A. and Teräsvirta, T. (2015), Modelling conditional correlations of asset returns: a smooth transition approach. *Econometric Reviews* 34, 174–197.

[26] Tse, Y.K., Tsui, K.C. (2002). A multivariate generalized autoregressive conditional heteroskedasticity model with time-varying correlations. *Journal of Business and Economic Statistics* 20, 351–362.

Tables and Figures

Table 1: QML estimation results for scalar DCC and NLARC models
(robust standard errors in parentheses)

Model	a	b	ϕ_A	Log–Lik	AIC	BIC
cDCC	0.006	0.977		-1876.1	1.472	1.476
	(0.001)	(0.004)				
NLARC	0.008	0.975	0.157	-1851.9	1.454	1.461
	(0.001)	(0.004)	(0.068)			

Table 2: QML estimation results for RSDC and FRSDC models
(robust standard errors in parentheses)

	RSDC			FRSDC			
	1–λ	2–λ		1–λ scalar	2–λ scalar	2–λ rank–1	2–λ group
p_{hh}	0.980	0.974	p_{hh}	0.967	0.895	0.905	0.898
	(0.004)	(0.006)		(0.006)	(0.013)	(0.013)	(0.013)
p_{ll}	0.340	0.323	p_{ll}	0.305	0.453	0.397	0.406
	(0.066)	(0.061)		(0.098)	(0.094)	(0.060)	(0.060)
λ_h		1.004	$\phi_1^{(h)}$		1.461	1.517 (SP)	0.928
		(0.001)			(0.195)	(0.243)	(0.065)
λ_l	0.552	0.591	$\phi_2^{(h)}$				1.259
	(0.095)	(0.097)					(0.178)
			$\phi_3^{(h)}$				0.317
							(0.128)
			$\phi_1^{(l)}$	1.907	1.058	0.969 (RUS)	1.464
				(0.522)	(0.300)	(0.124)	(0.182)
			$\phi_2^{(l)}$				0.586
							(0.079)
			$\phi_3^{(l)}$				5.251
							(1.301)
Log–Lik	-1744.0	-1719.2		-1734.5	-1577.9	-1453.0	-1494.8
AIC	1.369	1.350		1.362	1.240	1.172	1.178
BIC	1.376	1.360		1.369	1.249	1.268	1.196

The coefficients $\phi^{(h)}$ and $\phi^{(l)}$, in the case of the FRSDC models with rank–1, are the medians of the estimates (in parentheses the corresponding indices), considering, as median, the 10-th index in increasing order. For the FRSDC 2–λ models, the shrinking coefficient θ is equal to 0.21.

Table 3: In–sample forecast evaluation

Model Confidence Set						
NLARC	cDCC	RSDC 1–λ	FRSDC 1–λ	RSDC 2–λ	FRSDC 2–λ g	FRSDC 2–λ
0.000	0.000	0.000	0.000	0.000	0.691	1.000
Minimum Variance Portfolio						
FRSDC 2–λ g	FRSDC 2–λ	FRSDC 1–λ	RSDC 2–λ	RSDC 1–λ	cDCC	NLARC
102.76	102.34	101.60	101.27	101.15	100.30	100.00
Diebold–Mariano p-values for MVP						
	RSDC 2–λ	FRSDC 1–λ	FRSDC 2–λ	FRSDC 2–λ g	cDCC	NLARC
RSDC 1–λ	-0.074	-0.143	-0.008	-0.030	0.149	0.087
RSDC 2–λ		-0.245	-0.013	-0.039	0.102	0.065
FRSDC 1–λ			-0.039	-0.077	0.087	0.055
FRSDC 2–λ				-0.329	0.010	0.007
FRSDC 2–λ g					0.007	0.005
cDCC						0.117

For the MCS criterion: the models are in the order (from left to right) in which they are removed from the MCS and the corresponding p-value is indicated below the model name. The letter g indicates the groupwise version (all other are scalar). For the MVP: the models are shown (from left to right) in decreasing MVP order, setting to 100 the smallest MVP, so a number like $(100 + x)$ means that the corresponding model provides, on average, a $x\%$ higher MVP than the model having the lowest MVP. The bottom panel shows the p-values of the Diebold and Mariano statistics to compare the model in row with the model in column. A negative sign means that the model in row is better than the model in column.

Table 4: Out–of–sample evaluation of 441 forecasts (fixed estimates)

Model Confidence Set						
NLARC	cDCC	FRSDC 2–λ g	RSDC 1–λ	RSDC 2–λ	FRSDC 1–λ	FRSDC 2–λ
0.000	0.005	0.273	0.839	0.703	0.873	1.000
Minimum Variance Portfolio						
FRSDC 2–λ g	FRSDC 2–λ	FRSDC 1–λ	RSDC 1–λ	RSDC 2–λ	cDCC	NLARC
108.48	105.05	102.93	102.40	102.22	100.19	100.00
Diebold–Mariano p-values for MVP						
	RSDC 2–λ	FRSDC 1–λ	FRSDC 2–λ	FRSDC 2–λ g	cDCC	NLARC
RSDC 1–λ	0.229	-0.376	-0.013	-0.000	0.031	0.021
RSDC 2–λ		-0.340	-0.009	-0.000	0.029	0.020
FRSDC 1–λ			-0.020	-0.001	0.090	0.073
FRSDC 2–λ				-0.001	0.002	0.002
FRSDC 2–λ g					0.000	0.000
cDCC						0.024

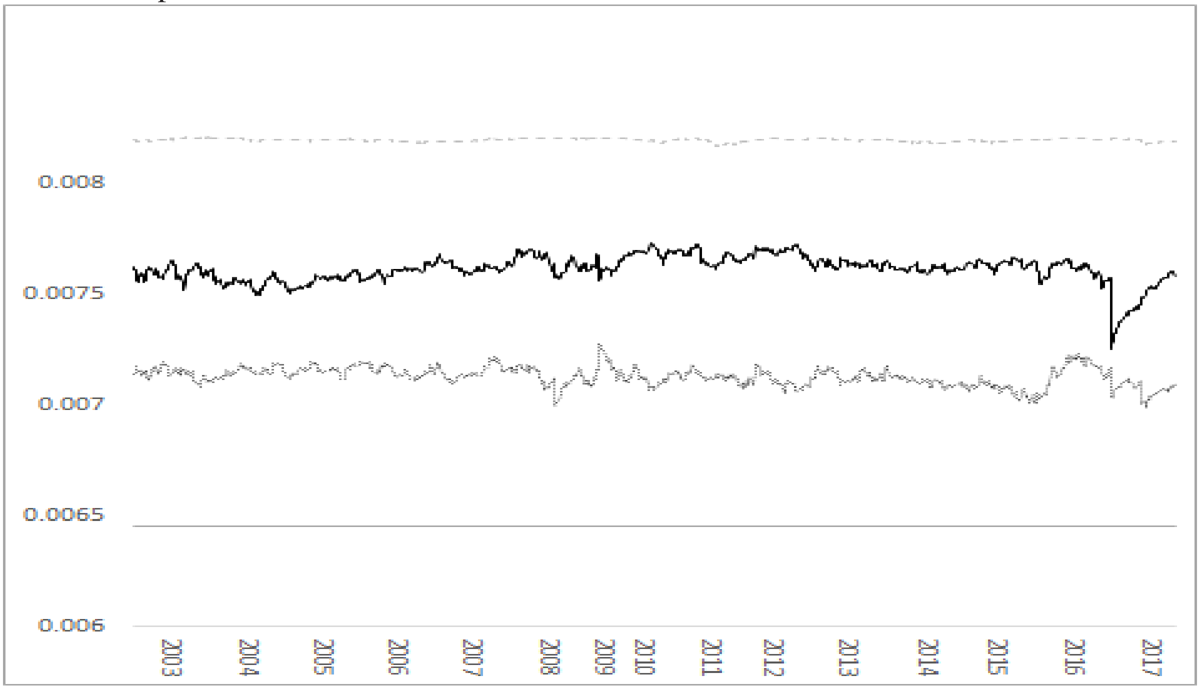
The models are estimated using the sample from 8th July 2002 to 30th December 2014; then the estimated models are fixed and the successive 441 correlations are forecasted. The letter g indicates the groupwise version (all other are scalar). For a description of the table contents, see the note below Table 3.

Table 5: Out–of–sample evaluation of 441 forecasts (updated estimates)

Model Confidence Set						
FRSDC 2–λ	RSDC 2–λ	RSDC 1–λ	FRSDC 1–λ	FRSDC 2–λ g	cDCC	NLARC
0.060	0.139	0.108	0.073	0.036	0.087	1.000
Minimum Variance Portfolio						
FRSDC 2–λ g	RSDC 2–λ	RSDC 1–λ	FRSDC 1–λ	FRSDC 2–λ	cDCC	NLARC
104.37	102.32	102.31	101.92	101.20	101.05	100.00
Diebold–Mariano p-values for MVP						
	RSDC 2–λ	FRSDC 1–λ	FRSDC 2–λ	FRSDC 2–λ g	cDCC	NLARC
RSDC 1–λ	-0.450	0.172	0.176	-0.013	0.303	0.086
RSDC 2–λ		0.152	0.191	-0.026	0.285	0.067
FRSDC 1–λ			0.292	-0.017	0.369	0.114
FRSDC 2–λ				-0.001	-0.498	0.295
FRSDC 2–λ g					0.128	0.030
cDCC						0.129

The first models are estimated using the sample from 8th July 2002 to 30th December 2014; then the estimates are updated each 5 observations and the last 441 correlations are forecasted. The letter g indicates the groupwise version (all other are scalar). For a description of the table contents, see the note below Table 3.

Figure 1: Estimated time series of $a_{ij,t} = a\exp[\phi_A(\rho_{ij,t-1} - 1)]$ in the NLARC model (see eq. (3.3)) for $i = SP$ and $j = FTSE$ (solid line), for $i = SP$ and $j = CNX$ (dotted line) and for $i = SP$ and $j = BOV$ (dashed line). The horizontal line corresponds to the estimated a parameter of the cDCC model.



.

.

Figure 2: Impact of the lagged covariance shock on the covariance $q_{ij,t}$ in the NLARC model, see eq. (3.3). An impact curve is defined by $a_{ij,t} = a \exp[\phi_A(\rho_{ij,t-1} - 1)]$ as a function of $\rho_{ij,t-1}$. Impact curves are shown for $a = 0.1$ and values of $\phi_A = 0, 0.5, 1, ..5$. The horizontal line at 0.1 corresponds to $\phi_A = 0$, the first curve below it to $\phi_A = 0.5$, the next one to $\phi_A = 0.5$, etc. The step function is obtained when ϕ_A approaches infinity.

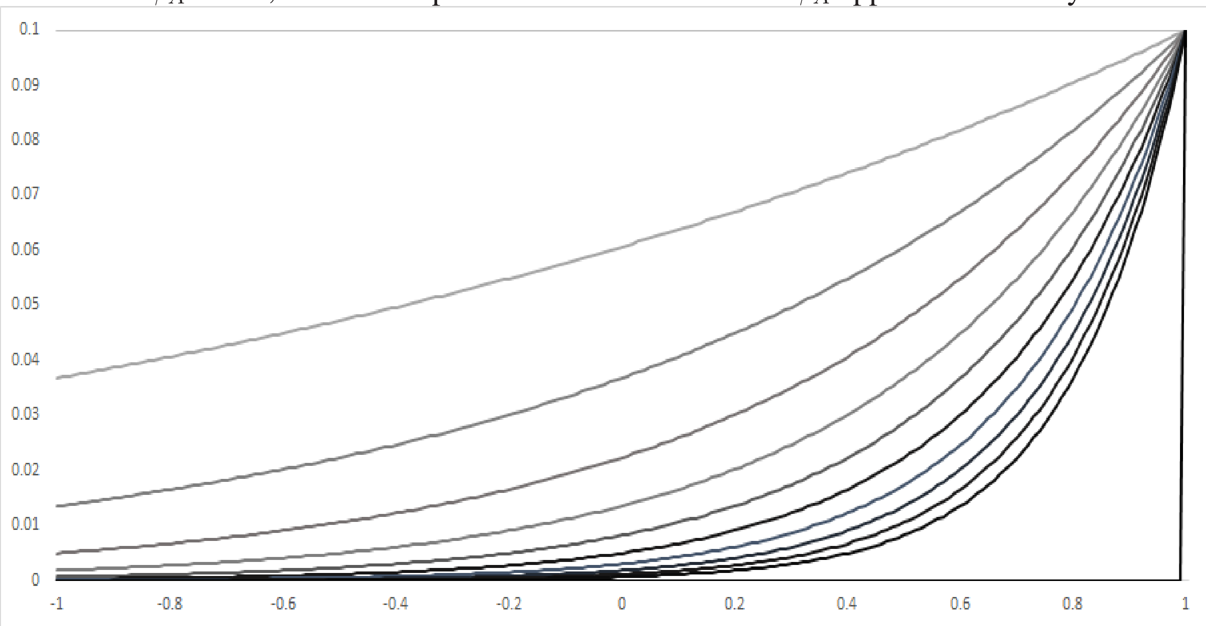


Figure 3: Time-varying part of the derivative of $q_{ij,t}$ with respect to $q_{ij,t-1}$ in the NLARC model for the conditional correlations between SP and the other 19 indices. The last graph represents the average of the 19 other.

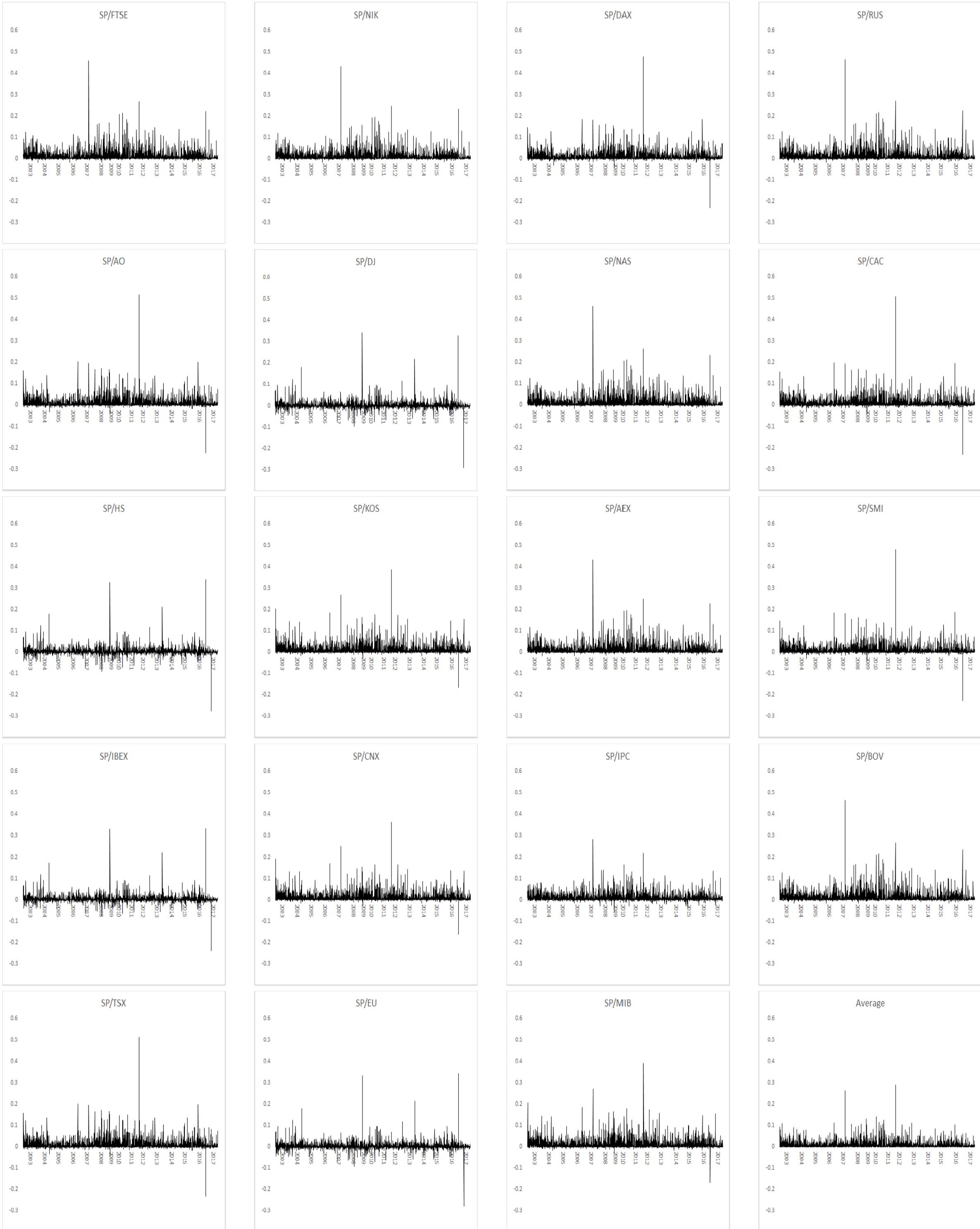


Figure 4: Estimated vectors $\phi^{(l)}$ (continuous line, left axis) and $\phi^{(h)}$ (dotted line, right axis) of the FRSDC–2λ model with rank–1 parameterization. The x axis refers to the corresponding financial indices.