DEPARTMENT OF ECONOMICS, MANAGEMENT AND STATISTICS

UNIVERSITY OF MILAN – BICOCCA

DEMS WORKING PAPER SERIES

Semiparametric Estimation of Multivariate GARCH Models

Claudio Morana

No. 317 – December 2015

Semiparametric Estimation of Multivariate GARCH Models

Claudio Morana[1,2,3]

[1]Department of Economics, Management and Statistics, University of Milan-Bicocca, Milan, ITALY
Email: claudio.morana@unimib.it

[2]Center for Research on Pensions and Welfare Policies, Collegio Carlo Alberto, Moncalieri, ITALY
[3]Rimini Center for Economic Analysis, Rimini, ITALY

December 2015

Abstract

The paper introduces a new simple semiparametric estimator of the conditional variance co-variance and correlation matrix (SP-DCC). While sharing a similar sequential approach to existing dynamic conditional correlation (DCC) methods, SP-DCC has the advantage of not requiring the direct parameterization of the conditional covariance or correlation processes, therefore also avoiding any assumption on their long-run target. In the proposed framework, conditional variances are estimated by univariate GARCH models, for actual and suitably transformed series, in the first step; the latter are then nonlinearly combined in the second step, according to basic properties of the covariance and correlation operator, to yield non-parametric estimates of the various conditional covariances and correlations. Moreover, in contrast to available DCC methods, SP-DCC allows for straightforward estimation also for the non-symultaneous case, i.e., for the estimation of conditional cross-covariances and correlations, displaced at any time horizon of interest. A simple ex-post procedure, to ensure well behaved conditional covariance and correlation matrices, grounded on nonlinear shrinkage, is finally proposed. Due to its sequential implementation and scant computational burden, SP-DCC is very simple to apply and suitable for the modeling of vast sets of conditionally heteroskedastic time series.

Keywords

Multivariate GARCH model, dynamic conditional correlation, semiparametric estimation

1. Introduction

Since the seminal contribution of Bollerslev et al. (1988), the literature on multivariate Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models has rapidly developed (see Bauwens et al., 2006 and Silvennoinen and Terasvirta, 2009, for surveys). To date, three generations of models can be counted. First generation models, likewise the VEC model of Bollerslev et al. (1988) and the BEKK model of Engle and Kroner (1995), are straightforward extensions of the univariate GARCH model. They allow for very general conditional variance covariance matrix dynamics, yet at the cost of a very profligate parameterization, which limits their use to small sets of time series. This drawback has been overcome by second generation models, yet at the cost of imposing either parameter restrictions on the BEKK model, as for the case of the scalar BEKK model and the exponentially weighted moving average model introduced by J.P.Morgan (1996), or on the conditional correlation matrix, assumed time-invariant in the constant conditional correlation CCC model of Bollerslev (1990). Alternatively, restrictions have been imposed through factor structures, likewise Engle and Gonzalez-Rivera (1991) and the orthogonal models of Alexander (2002), van der Weide (2002), Vrontos et al. (2003) and Fan et al. (2008). On the other hand, a different approach has been pursued by the most recent third generation of multivariate GARCH models, i.e., the *dynamic conditional correlation* models, grounded on a two-step estimation procedure, involving the estimation of univariate GARCH models for the conditional variances in the first step and then the estimation of the conditional covariances in the second step. Although inefficient, the latter sequential procedure is consistent and asymptotically normal. Moreover, by dramatically reducing the numerical optimization burden, it can be implemented also in the case of vast sets of time series. Seminal in this respect are the Dynamic Conditional Correlation models (DCC) of Engle (2002), Engle and Sheppard (2006) and Tse and Tsui (2002). Further extensions are Cappiello et al. (2006), the Dynamic Conditional Equi-Correlation (DECO) model of Engle and Kelly (2008), Engle (2009) and Pakel et al. (2014).

Dynamic conditional correlation models, in order to ensure positive definiteness of the conditional variance-covariance matrix, posit the correlation matrix to be a transformation of a latent matrix, which is a function of past devolatilized innovations. In particular, while the CCC model of Bollerslev (1990) assumes time invariant, but pairwise specific correlations, the DECO model of Engle and Kelly (2008) makes the opposite assumption, positing time varying correlations, but equal across series. Both CCC and DECO therefore rely on assumptions on conditional correlation dynamics which are unlikely to be supported by the data. On the other hand, in the alternative formulation of Tse and Tsui (2002), the correlation matrix is modeled directly and as a function of past correlations of devolatilized innovations. As a common drawback, all of the available dynamic conditional correlation models rely on the choice, neither unique nor obvious, of a long run target for the conditional variance covariance or correlation matrix.

In the light of the above issues, the paper then contributes to the literature on dynamic conditional correlation models, by introducing a new simple semiparametric estimator of the conditional variance covariance and correlation matrix (SP-DCC). While sharing a similar sequential approach to DCC and DECO, SP-DCC has the advantage of not requiring the direct parameterization of the conditional covariance or correlation processes, therefore also avoiding any assumption on their long-run target. In the proposed framework, conditional variances are estimated by univariate GARCH models, for actual and suitable transformed series, in the first step; the latter are then nonlinearly combined in the second step, according to basic properties of the covariance and correlation operator, to yield nonparametric estimates of the corresponding conditional covariances and correlations. In contrast to available DCC methods, SP-DCC allows for straightforward estimation also for the non-symultaneous case, i.e., for the estimation of conditional cross-covariances and correlations displaced at any time horizon of interest. A simple ex-post procedure, to ensure well behaved conditional covariance and correlation matrices, grounded on nonlinear shrinkage, is finally proposed. Due to its sequential implementation and scant computational burden, SP-DCC is very simple to apply and suitable for the modeling of vast sets of conditionally heteroskedastic time series. We point to Morana (2015) for an empirical application of the proposed approach.

2. Semiparametric estimation of dynamic conditional correlations

Consider a discrete time, real-valued vector stochastic process $\{y_t\}$ of dimension $N \times 1$

$$y_t = \mu_t(\theta) + \varepsilon_t \tag{1}$$

where $\mu_t(\theta)$ is the conditional mean vector $E(y_t|I_{t-1})$, θ is a vector of parameter, I_{t-1} is the sigma field, and

$$\varepsilon_t = H_t^{1/2}(\theta)z_t \tag{2}$$

where $H_t^{1/2}(\theta)$ is a positive definite matrix of dimension $N \times N$.

The random vector z_t is of dimension $N \times 1$ and assumed to be $i.i.d.$ with first two moments

$$E(z_t) = 0 \tag{3}$$
$$Var(z_t) = I_N \tag{4}$$

where I_N is the identity matrix of dimension N.

It is straightforward to show that $H_t(\theta)$ is the conditional covariance matrix; in fact

$$
\begin{aligned}
Var(y_t|I_{t-1}) &\equiv Var_{t-1}(y_t) = Var_{t-1}(\varepsilon_t) \\
&= H_t^{1/2}(\theta)Var_{t-1}(z_t)H_t^{1/2}(\theta)' \\
&= H_t(\theta).
\end{aligned}
$$

In general both μ_t and H_t depend on the parameter vector θ. While, the conditional mean vector does not depend on the conditional variance parameter, apart from the GARCH-in-mean case, the conditional variance matrix depends on the conditional mean parameters through the residuals. In what follow, for simplicity, we leave out θ from notation and neglect the conditional mean vector, which might be modelled in various ways, i.e. by means of univariate or multivariate ARMA models.

2.1. The conditional variance process

We assume the elements along the main diagonal of H_t follow a GARCH(1,1) process

$$h_{ii,t} = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{i,t-1} \qquad i = 1, ..., N \tag{5}$$

subject to the usual restrictions required to ensure that the generic ith conditional variance process $Var_{t-1}(y_{i,t}) \equiv h_{ii,t}$ is positive almost surely at any point in time. For instance, sufficient (not necessary) conditions are $\omega_i > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$, with stationarity condition $\alpha_i + \beta_i < 1$.[1]

An extended specification is in principle also viable, i.e.

$$h_{ii,t} = \omega_i + \alpha_{i,1}\varepsilon_{1,t-1}^2 + ... + \alpha_{i,N}\varepsilon_{N,t-1}^2 + \beta_{i,1}h_{1,t-1} + ... + \beta_{i,N}h_{N,t-1} \qquad i = 1, ..., N$$

yet actually feasible only for small N.

2.2. The conditional covariance process

Consider the identity

$$Cov(A, B) \equiv \frac{1}{4}[Var(A + B) - Var(A - B)] \tag{6}$$

given that $Var(A \pm B) = Var(A) + Var(B) \pm 2Cov(A, B)$.

---

[1] The GARCH(1,1) model is chosen for simplicity; the approach is very flexible and can accommodate any model of the GARCH family.

The off-diagonal elements of H_t can then be defined accordingly, i.e.,

$$Cov_{t-1}(y_{i,t}, y_{j,t}) \equiv h_{ij,t} = \frac{1}{4} \left[Var_{t-1}(y_{i,t} + y_{j,t}) - Var_{t-1}(y_{i,t} - y_{j,t}) \right] \qquad i,j = 1, ..., N \qquad i \neq j. \quad (7)$$

By defining the new variables $y_{ij,t}^+ \equiv y_{i,t} + y_{j,t}$ and $y_{ij,t}^- \equiv y_{i,t} - y_{j,t}$, and assuming a GARCH(1,1) specification for their conditional variance processes $Var_{t-1}(y_{ij,t}^+) \equiv h_{ij,t}^+$ and $Var_{t-1}(y_{ij,t}^-) \equiv h_{ij,t}^-$

$$Var_{t-1}(y_{ij,t}^+) \equiv h_{ij,t}^+ = \omega_{i,j}^+ + \alpha_{i,j}^+ \varepsilon_{i,j,t-1}^{+2} + \beta_{i,j}^+ h_{i,t-1}^+ \qquad i,j = 1, ..., N \qquad i \neq j \quad (8)$$

$$Var_{t-1}(y_{ij,t}^-) \equiv h_{ij,t}^- = \omega_{i,j}^- + \alpha_{i,j}^- \varepsilon_{i,j,t-1}^{-2} + \beta_{i,j}^- h_{i,t-1}^- \qquad i,j = 1, ..., N \qquad i \neq j \quad (9)$$

subject to the usual restrictions required to ensure a well behaved conditional variance process, (7) becomes

$$h_{ij,t} = \frac{1}{4} \left[h_{ij,t}^+ - h_{ij,t}^- \right] \qquad i,j = 1, ..., N \qquad i \neq j. \quad (10)$$

Moreover, if residuals are obtained from linear transformations of the original variables[2], then $\varepsilon_{ij,t-1}^{+2} = (\varepsilon_{i,t-1} + \varepsilon_{j,t-1})^2$ and $\varepsilon_{ij,t-1}^{-2} = (\varepsilon_{i,t-1} - \varepsilon_{j,t-1})^2$; hence, (8) and (9) can be written as

$$h_{ij,t}^+ = \omega_{i,j}^+ + \alpha_{i,j}^+ (\varepsilon_{i,t-1} + \varepsilon_{j,t-1})^2 + \beta_{i,j}^+ h_{i,t-1}^+ \qquad i,j = 1, ..., N \qquad i \neq j \quad (11)$$

$$h_{ij,t}^- = \omega_{i,j}^- + \alpha_{i,j}^- (\varepsilon_{i,t-1} - \varepsilon_{j,t-1})^2 + \beta_{i,j}^- h_{i,t-1}^- \qquad i,j = 1, ..., N \qquad i \neq j. \quad (12)$$

By means of the proposed method conditional cross-covariances and correlations can also be computed, as

$$Cov_{t-1}(y_{i,t}, y_{j,t-k}) = \frac{1}{4} \left[Var_{t-1}(y_{i,t} + y_{j,t-k}) - Var_{t-1}(y_{i,t} - y_{j,t-k}) \right] \qquad i,j = 1, ..., N \qquad i \neq j, k \neq 0. \quad (13)$$

2.3. Estimation

Consistent, and asymptotically normal estimation is performed in two steps.

Firstly, the conditional variances $h_{i,t}$, $i = 1, ..., N$, i.e., the elements along the main diagonal of H_t, and $h_{ij,t}^+$, $h_{ij,t}^-$, $i,j = 1, ..., N$, $i \neq j$, are estimated equation by equation by means of QML; this yields $\hat{h}_{i,t}$, $i = 1, ..., N$, and $\hat{h}_{ij,t}^+$, $\hat{h}_{ij,t}^-$, $i,j = 1, ..., N$, $i \neq j$.

Then, in the second step the off-diagonal elements of H_t, $h_{ij,t}$, $i,j = 1, ..., N$, $i \neq j$, are estimated nonparametrically by computing

$$\hat{h}_{ij,t} = \frac{1}{4} \left[\hat{h}_{ij,t}^+ - \hat{h}_{ij,t}^- \right] \qquad i,j = 1, ..., N \qquad i \neq j. \quad (14)$$

By defining

$$\hat{D}_t = diag \left(\hat{h}_{1,t}^{1/2}, ..., \hat{h}_{N,t}^{1/2} \right)$$

the conditional correlation matrix R_t is then estimated as

$$\hat{R}_t = \hat{D}_t^{-1} \hat{H}_t \hat{D}_t^{-1}.$$

By definition the matrix H_t is positive definite; our estimation approach does not restrict $\hat{H}_t$ to be almost surely positive definite at any point in time. The latter property can however be checked ex-post

---

[2] This, for instance, would occurr when the conditonal mean vector is specified as a vector autoregressive (VAR) process, yet not in the presence of a VARMA structure. In the latter case residuals $\varepsilon_{i,j,t}^+$ and $\varepsilon_{i,j,t}^-$ should be computed from time series models specified for the new variables $y_{i,t}^+$ and $y_{i,t}^-$.

by computing the eigenvalues of $\hat{H}_t$, which by being a real, square and symmetric matrix, under positive definiteness are expected to be all positive. In practice this can be performed by means of Descartes' rule of alternating signs applied to the its characteristic polynomial[3], as well as by means of Sylvester's criterion[4], or by assessing the existence and uniqueness of its Cholesky decomposition.

However, the positive definiteness property might also be imposed ex-post, by means of shrinkage methods, as in Ledoit and Wolf (2004, 2012). In the latter case a compromise estimate of the conditional correlation matrix $\hat{R}_t^c$ is obtained by shrinking the estimated conditional correlation matrix $\hat{R}_t$ towards the identity matrix, i.e., by computing

$$\hat{R}_t^c = \lambda_t I_N + (1 - \lambda_t)\, \hat{R}_t$$

where λ_t is the shrinkage intensity at time period t. The compromise estimate of H_t, i.e., $\hat{H}_t^c$ can then be obtained as

$$\hat{H}_t^c = \hat{D}_t \hat{R}_t^c \hat{D}_t$$

which is positive definite by construction, as $\hat{R}_t^c$ is positive definite and the elements of $\hat{D}_t$ are well-defined.

2.4. Ex-post correction for well-behaved conditional covariances and correlations

Alternatively, the validity of the Cauchy-Schwarz inequality and the condition of positive definiteness can be imposed sequentially, at each point in time t, following the below procedure.

Firstly, the estimated conditional correlations in $\hat{R}_t$, $\hat{\rho}_{ij}$, $i, j = 1, ..., N$, $i \neq j$, are bounded to lie within the range $-1 \leq \hat{\rho}_{ij} \leq 1$, by applying the sign-preserving bounding transformation

$$\hat{\rho}_{ij}^* = \hat{\rho}_{ij}(1 + \hat{\rho}_{ij}^k)^{-1/k} \tag{15}$$

where $k > 0$ and even; the value of k can be selected optimally by solving

$$\min_k \sum_{t=1}^{T} \left\| \hat{R}_t - \hat{R}_t^* \right\|_F = \min_k \sum_{t=1}^{T} \sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} \left| \hat{\rho}_{ij,t} - \hat{\rho}_{ij,t}^* \right|^2} \tag{16}$$

i.e., by setting k in such a way that the sum of Frobenious norms over the temporal sample is minimized; this yields $\hat{R}_t^*$, the transformed correlation matrix, which satisfies, by construction, the Cauchy-Schwarz inequality.

Secondly, positive definiteness is enforced by computing the eigenvalue-eigenvector decomposition of the transformed conditional correlation matrix $\hat{R}_t^*$, yielding

$$\hat{R}_t^* = \hat{E}_t \hat{V}_t \hat{E}_t'$$

where $\hat{V}_t$ is the $N \times N$ diagonal matrix containing the N ordered eigenvalues along the main diagonal, and $\hat{E}_t$ is the $N \times N$ matrix containing the N associated orthogonal eigenvectors. In the case of violation of the positive definiteness condition one or more of the eigenvalues will be negative; an empirically viable strategy to impose positive definiteness ex-post consists of replacing the negative sample eigenvalues with positive values, computed for instance from their sample average value when positive or from the grand average across sample eigenvalues. The rationale guiding this practice is the well-known issue of downward biased

---

[3] By ordering the terms of the characteristic polynomial with real coefficients by descending variable exponent, the number of positive roots of the polynomial is then either equal to the number of sign differences between consecutive nonzero coefficients, or is less than it by an even number. Multiple roots of the same value are counted separately.

[4] According to Sylvester's criterion, a real, square symmetric positive definite matrix shows all positive leading principal minors, where the kth leading principal minor of a matrix M is the determinant of its upper-left k by k sub-matrix. In practice the M matrix is reduced to an upper triangular matrix by means of row operations, as in the first part of the Gaussian elimination method, preserving the sign of its determinant during pivoting process. Since the kth leading principal minor of a triangular matrix is the product of its diagonal elements up to row k, positive definitness can the be assessed by checking whether its diagonal elements are all positive. The latter condition is then checked each time a new row k of the triangular matrix is obtained.

estimation of the smallest eigenvalues (versus upward biased estimation of the largest eigenvalues). Rather than shrinking all the sample eigenvalues towards their grand average, as occurring by implementing Ledoit and Wolf (2004), only the negative eigenvalues are shrinked towards positive average values. The latter practice is consistent with nonlinear shrinkage of the covariance matrix (Ledoit and Wolf, 2012), allowing in principle for different shrinkage intensities to be applied to the various eigenvalues.

The shrank matrix of eigenvalues $\hat{V}_t^*$ would then be obtained, and therefore

$$\hat{R}_t^{**} = \hat{E}_t \hat{V}_t^* \hat{E}_t' \tag{17}$$

which, by construction, is well-behaved at each point in time. The implied conditional covariance process at time period t can then be obtained as

$$\hat{H}_t^{**} = \hat{D}_t \hat{R}_t^{**} \hat{D}_t$$

where

$$\hat{D}_t = diag\left(\hat{h}_{1,t}^{1/2}, ..., \hat{h}_{N,t}^{1/2}\right)$$

as before. The implied estimated variance covariance matrix $\hat{H}_t^{**}$ then obeys the Cauchy-Schwarz inequality and the positive definiteness condition, at each point in time, by construction.

2.5. Asymptotic properties

Under assumptions (1) through (5), estimation and inference for the parameters of the univariate GARCH(1,1) processes (5), (8) and (9) can be performed by means of QML. The Gaussian log likelihood function for the generic process $y_{i,t}$, assuming for simplicity $E(y_{i,t}|I_{t-1}) = 0$ and a GARCH(1,1) structure $h_{i,t} = \omega_i + \alpha_i y_{i,t-1}^2 + \beta h_{i,t-1}$ can then be written as

$$ln\{L(\boldsymbol{\vartheta}, y_{i,t}, ..., y_{i,T})\} = -0.5T\ln(2\pi) - 0.5\sum_{t=1}^{T}\{\ln(h_{i,t}) + y_{i,t}^2 h_{i,t}^{-1}\}$$

and numerically maximized with respect to the vector of parameters $\boldsymbol{\vartheta} = (\omega_i, \alpha_i, \beta_i)'$. Similarly for the other variables $y_{i,j,t}^+$ and $y_{i,j,t}^-$.

Under fairly general conditions, the asymptotic distribution of QML is

$$T^{1/2}\left(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0\right) \to N\{\mathbf{0}, \mathbf{A}(\boldsymbol{\vartheta}_0)^{-1}\mathbf{B}(\boldsymbol{\vartheta}_0)\mathbf{A}(\boldsymbol{\vartheta}_0)^{-1}\}$$

where $\boldsymbol{\vartheta}_0$ denotes the true value of the vector of parameters, and where $\mathbf{A}(\boldsymbol{\vartheta}_0)$ is the Hessian and $\mathbf{B}(\boldsymbol{\vartheta}_0)$ is the outer product gradient, both of which are evaluated at the true parameter values. This also establishes the consistent and asymptotically normal estimation of the conditional variance of y_t, $E\left(y_{i,t}^2|I_{t-1}\right)$, as well as of the transformed variables $y_{ij,t}^+$ and $y_{ij,t}^-$.

Consistent and asymptotically normal estimation of the off-diagonal elements of the conditional variance covariance matrix H_t then follows directly from the consistent and asymptotically normal estimation of the conditional variances of the transformed variables in (8) and (9). In fact, considering the generic off-diagonal i,j element of H_t, $i,j = 1, ..., N$, $i \neq j$, one has

$$\widehat{Cov}_{t-1}(y_{i,t}, y_{j,t}) = \frac{1}{4}\left[\widehat{Var}_{t-1}(y_{i,t} + y_{j,t}) - \widehat{Var}_{t-1}(y_{i,t} - y_{j,t})\right]$$

as the conditional covariance estimator $\widehat{Cov}_{t-1}(y_{i,t}, y_{j,t})$ is a linear combination of the (consitent and asymptotically Normal) conditional variance estimators for the transformed variables $\widehat{Var}_{t-1}(y_{i,t} + y_{j,t})$ and $\widehat{Var}_{t-1}(y_{i,t} - y_{j,t})$.

6

3. Conclusions

The paper introduces a new simple semiparametric estimator of the conditional variance covariance and correlation matrix (SP-DCC). While sharing a similar sequential approach to existing dynamic conditional correlation methods, SP-DCC has the advantage of not requiring the direct parameterization of the conditional covariance or correlation processes. In the first step, conditional variances are estimated by univariate GARCH models for actual and suitably transformed series. In the second step, the estimated conditional covariances are then nonlinearly combined, according to basic properties of the covariance and correlation operator, to yield nonparametric estimates of the various conditional covariances and correlations. At this step, SP-DCC also allows for the estimation of conditional cross-covariances and correlations, displaced at any time horizon. In the third step, well behaved conditional covariance and correlation matrices are obtained by means of nonlinear shrinkage. Due to its sequential implementation and scant computational burden, SP-DCC is very simple to apply and suitable for the modeling of vast sets of conditionally heteroskedastic time series.

References

[1] C.O. Alexander. Principal component models for generating large GARCH covariance matrices *Economic Notes*, 31, 337-359, 2002. doi: 10.1111/1468-0300.00089

[2] L. Bauwens, S. Laurent and L. Rombouts. Multivariate GARCH models: A survey *Journal of Applied Econometrics*, 21, 79-109, 2006. doi: 10.1002/jae.842

[3] T. Bollerslev, R.F. Engle and J. Wooldridge. A capital asset pricing model with time varying covariances *Journal of Political Economy*, 96, 116-131, 1988. doi: 10.1086/261527

[4] T. Bollerslev. Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH approach *Review of Economics and Statistics*, 72, 489-505, 1990. doi: 10.2307/2109358

[5] L. Cappiello, R.F. Engle and K. Sheppard. Asymmetric dynamics in the correlations of global equity and bond returns *Journal of Financial Econometrics*, 4, 537-572, 2006. doi: 10.1093/jjfinec/nbl005

[6] R.F. Engle and F. Kroner. Multivariate simultaneous generalized ARCH *Econometric Theory*, 11, 122-150, 1995. doi: 10.1017/S0266466600009063

[7] R.F. Engle and G. Gonzalez-Rivera. Semiparametric ARCH models *Journal of Business and Economic Statistics*, 9, 345-59, 1991. doi: 10.1080/07350015.1991.10509863

[8] R.F. Engle Dynamic conditional correlation: A simple class of multivariate GARCH models *Journal of Business and Economic Statistics*, 20, 339-350, 2002. doi: 10.1198/073500102288618487

[9] R.F. Engle. High dimension dynamic correlations. In *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, J. Castle and N. Shephard (eds.),Oxford University Press, 2009. ISBN: 9780199237197.

[10] R.F. Engle and B. Kelly. Dynamic equicorrelation *Journal of Business and Economic Statistics*, 30, 212-228, 2008. doi: 10.1080/07350015.2011.652048.

[11] J. Fan, M. Wang and Q. Yao. Modelling multivariate volatilities via conditionally uncorrelated components *Journal of Royal Statistical Society B*, 70, 679-702, 2008. doi: 10.1111/j.1467-9868.2008.00654.x

[12] J.P.Morgan. *Riskmetrics Technical Document*, 4th ed. J.P.Morgan, New York. doi: 10.1007/s10260-015-0332-9

[13] O. Ledoit and M. Wolf. Honey, I shrunk the sample covariance matrix *Journal of Portfolio Management*, 30, 10-119, 2004. doi: 10.3905/jpm.2004.110

[14] O. Ledoit and M. Wolf. Nonlinear shrinkage estimation of large-dimensional covariance matrices *The Annals of Statistics*, 40, 1024-1060, 2012. doi: 10.1214/12-AOS989

[15] C. Morana. The US\$/€ exchange rate: Structural modeling and forecasting during the recent financial crisis, Università di Milano-Bicocca, mimeo, 2015.

[16] C. Pakel, D. Shephard, K. Sheppard and R.F. Engle. Fitting vast dimensional time-varying covariance models, New York University, mimeo, 2014.

[17] A. Silvennoinen and T. Terasvirta Multivariate GARCH models. In *Handbook of Financial Time Series*, T. Andersen, R. Davis, J.F. Kreiss and T. Mikosch (eds), Springer, 2009. ISBN: 978-3-540-71297-8

[18] Y. Tse and A Tsui. A multivariate GARCH model with time-varying correlations *Journal of Business and Economic Statistics*, 20, 351-362, 2002. doi: 0.1198/073500102288618496.

[19] R. van der Weide. GO-GARCH: A multivariate generalized orthogonal GARCH model *Journal of Applied Econometrics*, 17, 549-564, 2002. doi: 10.1002/jae.688

[20] I. Vrontos, P. Dellaportas and V. Politis. A full-factor multivariate GARCH model *Econometrics Journal*, 6, 311-333, 2003. doi: 10.1111/1368-423X.t01-1-00111