UC3M Working Papers
Statistics and Econometrics
15-16
ISSN 2387-0303
July 2015

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-48

MGARCH MODELS: TRADEOFF BETWEEN FEASIBILITY AND FLEXIBILITY

Daniel de Almeida[a], Luiz Hotta[b], Esther Ruiz[a]

Abstract

The parameters of popular multivariate GARCH (MGARCH) models are restricted so that their estimation is feasible in large systems and covariance stationarity and positive definiteness of conditional covariance matrices are guaranteed. These restrictions limit the dynamics that the models can represent, assuming, for example, that volatilities evolve in an univariate fashion, not being related neither among them nor with the correlations. This paper updates previous surveys on parametric MGARCH models focusing on their limitations to represent the dynamics observed in real systems of financial returns. The conclusions are illustrated using simulated data and a five-dimensional system of exchange rate returns.

Keywords: BEKK, DCC, multivariate conditional heteroscedasticity, variance targeting, VECH

[a] Department of Statistics, Universidad Carlos III de Madrid.
[b] University of Campinas.

Acknowledgements: The first author was supported by grants of 0969/13-3 CAPES, Coordination of Improvement of Higher Education Personnel. The second author acknowledges financial support from CAPES, grant 10600/13-2, São Paulo Research Foundation (FAPESP), grant 2013/00506-1, and Laboratory EPIFISMA. Financial support from ECO2012-32401 project by the Spanish Government is gratefully acknowledged by the third author. Finally, we acknowledge Juan Pablo Ortega for the codes of the VECH estimation.

MGARCH models: Tradeoff between feasibility and flexibility

DANIEL DE ALMEIDA[1]

Department of Statistics, University Carlos III

LUIZ K. HOTTA[2]

University of Campinas

ESTHER RUIZ[3]

Department of Statistics, University Carlos III

July, 2015.

[1] Calle Madrid 126, 28903 Getafe (Madrid), Spain, telephone: +34 91 6249851, email: dani.d.almeida89@gmail.com

[2] Rua Sérgio Buarque de Holanda, 651, Barão Geraldo, Campinas, SP, Brazil, telephone: + 55 19 3521 6081, email: hotta@ime.unicamp.br

[3] Calle Madrid 126, 28903 Getafe (Madrid), Spain, telephone: +34 91 6249851, email: ortega@est-econ.uc3m.es

Abstract

The parameters of popular multivariate GARCH (MGARCH) models are restricted so that their estimation is feasible in large systems and covariance stationarity and positive definiteness of conditional covariance matrices are guaranteed. These restrictions limit the dynamics that the models can represent, assuming, for example, that volatilities evolve in an univariate fashion, not being related neither among them nor with the correlations. This paper updates previous surveys on parametric MGARCH models focusing on their limitations to represent the dynamics observed in real systems of financial returns. The conclusions are illustrated using simulated data and a five-dimensional system of exchange rate returns. (*JEL*: C32, C52, C58)

KEYWORKS: BEKK, DCC, multivariate conditional heteroscedasticity, variance targeting, VECH

Generalized autoregressive conditional heteroscedastic (GARCH) models, originally proposed by Engle (1982) and Bollerslev (1986), were a big advance in the statistical analysis of univariate financial returns and are widely fitted to describe and forecast their volatilities. Univariate GARCH models were soon extended to a multivariate framework by Bollerslev *et al.* (1988). Since then, multivariate GARCH (MGARCH) models have attracted a great deal of attention due to several applications that require estimates of conditional variances, covariances and correlations of multivariate time series. The largest number of implementations of MGARCH models appear in the context of systems of financial returns; see, for example, Bollerslev *et al.* (1988), Attanasio (1991), De Santis and Gerard (1997), Hansson and Hordahl (1998), Lien and Tse (2002), Engle and Colacito (2006), Andersen *et al.* (2007), McNeil *et al.* (2010), Santos *et al.* (2012) and Santos *et al.* (2013), for a few selected asset pricing, portfolio selection, risk management and future hedging applications. It is important to note that, depending on the particular application, the number of returns in the system can vary from being rather small to extremely large; see, for example, Bollerslev *et al.* (1988), Kavussanos and Visvikis (2004), Kawakatsu (2006) and Beirne *et al.* (2013) for small systems and Cappiello *et al.* (2006), Diebold and Yilmaz (2009), Santos *et al.* (2012), Santos *et al.* (2013) and Rombouts *et al.* (2014) for large systems. Besides financial applications, MGARCH models have also been fitted to systems of macroeconomic and commodity related variables. For example, Conrad and Karanasos (2010) explain the inflation-growth interaction fitting a generalized version of the constant conditional correlation (CCC) model, while Baillie and Myers (1991) estimate the optimal hedge ratios of

[0]The first author was supported by grants of 0969/13-3 CAPES, Coordination of Improvement of Higher Education Personnel. The second author acknowledges financial support from CAPES, grant 10600/13-2, São Paulo Research Foundation (FAPESP), grant 2013/00506-1, and Laboratory EPIFISMA. Financial support from ECO2012-32401 project by the Spanish Government is gratefully acknowledged by the third author. Finally, we acknowledge Juan Pablo Ortega for the codes of the VECH estimation.

commodity futures and Kavussanos and Visvikis (2004) analyze the interaction between spot and forward returns and volatilities in the shipping freight markets. MGARCH models have also been implemented in agricultural economics; see Gardebroek and Hernandez (2013) and Haixia and Shiping (2013) for some references. Finally, MGARCH models are also useful when modeling and forecasting non-economic time series. For example, Cripps and Dunsmuir (2003) and Jeon and Taylor (2012) fit bivariate vector autoregressive moving average (VARMA) models with MGARCH errors to model the wind speed and direction.

The original MGARCH models, which were direct generalizations from their univariate counterparts, were rather flexible allowing all volatilities and conditional covariances in the model to be related with each other. However, in practice, the empirical implementation of MGARCH models was limited due to two main limitations. First, due to the need of estimating a large number of parameters, their implementation was originally restricted to systems with very few series. Second, their parameters need to be restricted to guarantee covariance stationarity and positive definiteness of conditional covariance matrices. Consequently, many popular MGARCH models implemented to represent the dynamic evolution of volatilities, covariances and correlations of real systems are restricted in a such way that parameter estimation is feasible and it is easy to guarantee covariance stationarity and positiveness. The corresponding restrictions are often based on assuming that volatilities depend on their own past without interrelations neither among them nor with covariances. However, if these restrictions are not satisfied, the estimated conditional variances, covariances and correlations may suffer from strong biases; see Kroner and Ng (1998), Ledoit *et al.* (2003), Caporin and McAleer (2008), Rossi and Spazzini (2010) and Amado and Teräsvirta (2014) for consequences of misspecifying conditional variances, covariances and correlations. Engle (2009) argues that, although it seems important to allow for square and cross-products of one asset to help forecasting variances and covariances of other assets, in fact, there are few striking examples of these interrelations in the literature. On the contrary, several works conclude that allowing for interrelations among conditional variances and correlations may be important. First, empirical evidence on volatility feedbacks is plentiful; see Granger *et al.* (1986), Engle *et al.* (1990), Comte and Lieberman (2000), Hafner and Herwartz (2006, 2008b), Bai and Chen (2008), Diebold and Yilmaz (2009), Nakatani and Teräsvirta (2009), Conrad and Karanasos (2010), Beirne *et al.* (2013) and Aboura and Chevallier (2015) for studies related with volatility feedbacks. Therefore, it seems important to allow for volatility interactions when specifying MGARCH models. Second, it has often been found that volatilities and cross-correlations across assets move together over time and, consequently, they cannot be estimated separately. For example, Ramchand and Susmel (1998), Longin and Solnik (2001) and Okimoto (2008) find that cross-correlations between markets are higher during unstable periods when the markets are more volatile. Kasch and Caporin (2013) also describe correlations between conditional correlations and variances, while Bauwens and Otranto (2013) have a very complete description of the literature on volatility as a determinant of correlations. Finally, we did not find empirical evidence in the related literature about the influence of correlations on volatilities.

The choice of a particular MGARCH model can lead to substantially different conclusions when forecasting dynamic covariance matrices. However, in empirical applications, the particular specification fitted to the data is often chosen on an ad hoc basis; see the discussion by Caporin and McAleer (2012). In many cases, the easy of estimation is the primary factor affecting the selection of the model; see Kroner and Ng (1998) for an early reference. In this paper, we survey the main developments of parametric MGARCH models, updating previous surveys by Bauwens *et al.* (2006), Engle (2009) and Silvennoinen and Teräsvirta (2009a). The objective is to analyze the empirical implications of the restrictions imposed on MGARCH models to reduce the number of parameters and/or to guarantee covariance stationarity and/or positiveness. Comparisons between alternative MGARCH specifications have been previously carried out by Rossi and Spazzini (2010), who compare a smaller number of MGARCH models, and Caporin and McAleer (2014) who focus on the effect of the dimension of the system.

The rest of this paper is organized as follows. Section 1 describes several popular MGARCH specifications often implemented in empirical applications when estimating conditional variances, covariances and correlations. Section 2 compares their finite sample performance by carrying out Monte Carlo experiments to analyze the effect of restrictions on the estimation of conditional covariance matrices. An empirical application to a five-dimensional system of daily exchange rate returns of the Euro, British Pound, Swiss Franc, Australian Dollar and Japanese Yen against the US Dollar is carried out in Section 3. Finally, Section 4 concludes.

1 Multivariate GARCH models

This section describes several popular MGARCH models often implemented to represent the dynamic evolution of conditional variances, covariances and correlations of multivariate conditionally heteroscedastic time series.

1.1 The VECH model

Assuming zero conditional mean, the original multivariate GARCH model, proposed by Bollerslev *et al.* (1988) and denoted by VECH, is given by[1]

$$\mathbf{r}_t = \mathbf{H}_t^{1/2}\varepsilon_t, \tag{1}$$

$$\text{vech}(\mathbf{H}_t) = \mathbf{C} + \mathbf{A}\ \text{vech}(\mathbf{r}_{t-1}\mathbf{r}_{t-1}') + \mathbf{B}\ \text{vech}(\mathbf{H}_{t-1}), \tag{2}$$

where $\mathbf{r}_t$ is the $N \times 1$ vector of returns[2] observed at time t, for $t = 1, \ldots, T$, $\mathbf{H}_t$ is the $N \times N$ conditional covariance matrix of $\mathbf{r}_t$ and ε_t is a serially independent multivariate white noise process with covariance matrix $\mathbf{I}_N$, the identity matrix of order N. The operator vech$(\cdot)$ stacks the columns of the lower triangular part of a square matrix. Finally, $\mathbf{C}$ is an $N(N+1)/2$ vector of constants and $\mathbf{A}$ and $\mathbf{B}$ are square $N(N+1)/2$ parameter matrices. The initial condition for $\mathbf{H}_t$ is given by $\mathbf{H}_1 = \mathbf{\Sigma} = E(\mathbf{r}_t\mathbf{r}_t')$, the unconditional covariance matrix of returns. The VECH model is covariance stationary if the moduli of the eigenvalues of $\mathbf{A}+\mathbf{B}$ are less than one; see Engle and Kroner (1995). Furthermore, Hafner (2003) derives analytical expressions of the fourth order moments of returns and Hafner and Preminger (2009) establish sufficient conditions for geometric ergodicity. Although there are not known necessary conditions for the positivity of $\mathbf{H}_t$, Gourieroux (1997) provides sufficient conditions.

The VECH model in equations (1) and (2) is very flexible to represent symmetric responses of conditional variances and covariances to past square returns and cross-products of returns. The conditional variances depend on each other and on past conditional covariances. Similarly, the conditional covariances depend not only on past cross-products of returns but also on past conditional variances. However, the main limitation of the VECH model appears when trying to estimate their parameters in relatively large systems. The most popular estimator is Gaussian quasi-maximum likelihood (G-QML) based on maximizing the Gaussian log-likelihood function. Provided the conditional mean vector and the conditional covariance matrix are well specified, Bollerslev and Wooldridge (1992) show that the G-QML estimator is consistent even if the Data Generating Process (DGP) is not conditionally Gaussian. Hafner and Preminger (2009) establish consistency and asymptotic normality of the G-QML estimator under the existence of sixth-order moments; see also Hafner and Herwartz (2008a) who provide analytical expressions of the score and the Hessian. It is important to note that the dimension of the covariance matrix of the parameter estimator is at least of order N^2. Consequently, the asymptotic covariance matrix of the parameter estimator, computed as the average of the T outer-products of the score, will not be full rank for large N. This is a feature inherent to all MGARCH models and all their estimators; see Palandri (2009).

As argued by Bauwens and Laurent (2005) and Rossi and Spazzini (2010), among many others, the MGARCH models combined with Gaussian innovations could be inadequate once conditional financial returns exhibit fat tails and are often skewed. A natural alternative to the G-QML estimator is based on maximizing the Student-ν likelihood; see Fiorentini *et al.* (2003), Hafner and Herwartz (2006) and Bai and Chen (2008). The corresponding estimator will be denoted by S-QML. When compared with the Gaussian likelihood, the Student-ν likelihood has an extra scalar parameter, the degrees of freedom, ν, so that the complexity of the estimation process increases. Bauwens and Laurent (2005) further propose a multivariate skew-Student distribution and show empirically that it provides better, or at least not worse, out-of-sample Value at Risk (VaR) forecasts than a symmetric density. Other distributions used in MGARCH models are the fat-tailed multivariate Laplace of Rombouts *et al.* (2014) and the Multiple Degrees of Freedom t of Serban *et al.* (2007); see Rossi and Spazzini (2010) for a comparison of the performance of these distributions associated with different MGARCH specifications.

Regardless of the particular distribution assumed, the estimation procedure is computationally demanding and could be unfeasible when the dimension of the system is relatively large. Indeed, the log-likelihood function is nonlinear on the parameters and, in each iteration of its maximization algorithm, the matrix $\mathbf{H}_t$ needs to be inverted T times. Another difficulty of estimating the VECH model is that their parameters need to be subjected to nonlinear constraints to ensure the existence of covariance stationary solutions and the positive semidefinite character of the conditional covariance matrices. Chrétien and Ortega (2014) solve the estimation problem by incorporating these non-linear constraints in an efficient and natural way, using a Bregman-proximal trust-region method. They fit the VECH model for real systems of stock returns for dimensions up to eight and, with considerable computational effort, find a superior performance of the estimated VECH model in relation to other traditional parsimonious models.

The parameters of the VECH model can also be estimated using the two-step covariance targeting (VT) procedure proposed by Engle and Mezrich (1996); see Kristensen and Linton (2004) and Francq *et al.* (2011) for the properties of the VT estimator in the univariate case as well as for a comparison with the standard G-QML estimator and Caporin and McAleer (2012) for a general description. The extension of VT to the multivariate case is theoretically straightforward. The VT estimator can be implemented after rewriting equation (2) in terms of the unconditional covariance matrix as follows

$$\text{vech}(\mathbf{H}_t) = \text{vech}(\mathbf{\Sigma})(\mathbf{I}_{\frac{N(N+1)}{2}} - \mathbf{A} - \mathbf{B}) + \mathbf{A}\ \text{vech}(\mathbf{r}_{t-1}\mathbf{r}'_{t-1}) + \mathbf{B}\ \text{vech}(\mathbf{H}_{t-1}). \tag{3}$$

The estimation procedure is then divided into two steps. First, the unconditional covariance matrix, $\mathbf{\Sigma}$, is estimated by the sample unconditional covariance matrix of $\mathbf{r}_t$ and substituted in (3). Then, the remaining parameters are estimated by G-QML, conditional on the sample estimates of the unconditional covariances. Hence, in the second step, there are less free parameters to be estimated compared with G-QML. However, even for moderate systems, one still has to estimate a large number of parameters.

As an illustration, we generate a bivariate system of size $T = 1000$ by a VECH model with Gaussian errors and parameters $\mathbf{\Sigma} = 10^{-4}\times((2.217\ 0.887)'\ (0.887\ 1.763)')$, $\mathbf{A} = ((0.097\ 0.014\ 0.022)'\ (0.016\ 0.069\ 0.011)'\ (0.025\ 0.01\ 0.105)')$ and $\mathbf{B} = \text{diag}(0.8695, 0.857, 0.85)$. The parameters are chosen so as to represent the conditional variances and covariances usually estimated when dealing with real data and to guarantee covariance stationarity and positiveness of the covariance matrices; see, for example, Bai and Chen (2008) and Conrad and Karanasos (2010).[3]

The first two rows of Figure 1 plot the simulated conditional standard deviations for each of the two variables in the system, while the last two rows plot the conditional covariances and correlations, respectively. It is important to point out that, although we estimate the parameters by G-QML[4] and by VT without restricting them to ensure positivity and covariance stationarity, the estimated parameters satisfy the corresponding conditions. The estimated conditional standard deviations, covariances and correlations are also plotted in the first column of Figure 1, while the second column plots the simulated values (x-axis) versus the estimated values (y-axis) corresponding to the estimates obtained with the two procedures. In the second column plots, the errors can be measured as the Euclidean distance between the points and the identity line. Note that, as the fitted model is the true DGP, the errors plotted in Figure 1 can be attributed to parameter estimation. This figure illustrates that the errors have means close to zero and relatively small dispersion. Furthermore, there are not large differences when estimating the parameters using G-QML or VT, but both procedures tend to overestimate the conditional standard deviations when they are small.

Alternative models have been proposed in the literature to overcome the main problems in estimating the VECH model. The restricted models should be parsimonious enough not to hamper the estimation and interpretation of their parameters and simultaneously ensuring covariance stationarity and positive definiteness of the conditional covariance matrices. The later issue can be solved in two ways: either finding conditions under which the conditional covariance matrices implied by the model are positive definite and guarantee covariance stationarity, or defining a model whose conditional covariance matrices are positive definite and/or stationary by the model structure. The second approach is preferred, at least when dealing with positiveness, since as mentioned above, the conditions to guarantee positivity are only partially known and sometimes are non-linear on the parameters; see Gourieroux (2007). Moreover, it is important to point out that although the restrictions are usually imposed on the matrices of parameters that govern the dynamic evolution of the conditional covariances, $\mathbf{A}$ and $\mathbf{B}$, very recently, Caporin and Paruolo (2015) propose a model with restrictions on the unconditional covariance matrix. In any case, the restricted models may fail to represent the rich dynamics of real systems of financial returns.

The most popular restricted parametric MGARCH models can be divided basically into two categories: models for conditional covariance matrices and for conditional correlation matrices. The first category is related with the early parametric MGARCH models based on directly restricting the parameters in equation (2). The most popular ones are the diagonal VECH (DVECH) and the BEKK models. The second category includes models that represent the conditional variances and correlations rather than directly modeling the conditional covariance matrices. Examples of these models are the CCC and the dynamic conditional correlation (DCC) models. Next, we briefly describe each of these two families of models.

1.2 Restricted models for conditional covariance matrices

1.2.1 Diagonal VECH model

The diagonal VECH (DVECH) model, suggested by Bollerslev *et al.* (1988), assumes that the $\mathbf{A}$ and $\mathbf{B}$ matrices in equation (2) are diagonal. Consequently, each conditional variance and covariance in the system has a univariate GARCH-type specification without allowing for feedbacks among volatilities and between volatilities and covariances; see, for example, Bauwens *et al.* (2007) for a bivariate empirical application.

Sufficient conditions to guarantee the positivity of the covariance matrices can be derived by reparametrizing the DVECH model in terms of Hadamard products, as follows

$$\mathbf{H}_t = \mathbf{C}^* + \mathbf{A}^* \circ \mathbf{r}_{t-1}\mathbf{r}_{t-1}' + \mathbf{B}^* \circ \mathbf{H}_{t-1}, \tag{4}$$

where $\mathbf{C}^*$, $\mathbf{A}^*$ and $\mathbf{B}^*$ are $N \times N$ symmetric parameter matrices and $\circ$ denotes the Hadamard product. Ding and Engle (2001) show that the conditional covariance matrix, $\mathbf{H}_t$, is positive definite if $\mathbf{C}^*$ is positive definite and $\mathbf{A}^*$ and $\mathbf{B}^*$ are positive semi-definite. Later, Gourieroux (2007) derives necessary and sufficient conditions for the bivariate DVECH model and shows that the sufficient condition by Ding and Engle (2001) is also a necessary one. He mentions that the extension of the necessary positivity conditions to volatility matrices with dimensions larger than 3 is an open question. Finally, Ledoit *et al.* (2003) also derive a sufficient condition for positiveness. In particular, denoting by $\div$ the elementwise division, if $\mathbf{C}^* \div (1 - \mathbf{B}^*)$, $\mathbf{A}^*$ and $\mathbf{B}^*$ are positive semidefinite, then $\mathbf{H}_t$ is also positive semidefinite. Furthermore, they find the following necessary condition to ensure covariance stationarity: $\mathbf{A}^*$ and $\mathbf{B}^*$ matrices are positive semidefinite and $a_{ij}^* + b_{ij}^* < 1, \forall i$, where a_{ij}^* and b_{ij}^* are the elements of $\mathbf{A}^*$ and $\mathbf{B}^*$, respectively.

Although G-QML estimation of the parameters of the DVECH model is easier than in the complete VECH model, Ledoit *et al.* (2003) argue that it is not computationally feasible for systems of dimension $N > 5$. The DVECH model still has too many parameters that interact in a complex way and, as a consequence, it is difficult to obtain convergence using existing optimization algorithms. Moreover, the estimation of the DVECH model does not necessarily yield positive semidefinite conditional covariance matrices. To solve these issues, Ledoit *et al.* (2003) propose estimating the DVECH model using a flexible two-step procedure, hereafter denoted by LSW (after Ledoit-Santa Clara-Wolf). In the first step, the volatilities are estimated by fitting univariate GARCH models and the conditional covariances by fitting bivariate GARCH models. These estimates do not necessarily yield positive conditional covariance matrices. Consequently, in the second step, the estimated matrices, $\hat{\mathbf{C}}^*$, $\hat{\mathbf{A}}^*$ and $\hat{\mathbf{B}}^*$, are transformed in such a way that they guarantee positive semi-definite conditional covariance matrices with the transformation being the least disruptive. Finally, standard errors of the parameters could be obtained by bootstrapping.

Next, we illustrate the biases incurred when estimating conditional variances, covariances and correlations after fitting the DVECH model to the bivariate system generated by the DGP described in subsection 1.1. The pseudo-parameters of the DVECH model are estimated by the G-QML, VT and LSW procedures. The first column of Figure 2 plots the simulated and estimated conditional standard deviations, covariances and correlations, while in the second column are the scatter plots of the simulated values versus the estimated values. We can observe that, first, the estimates of the conditional standard deviations, covariances and correlations obtained when implementing the three alternative estimators considered are very similar. Second, as expected, the errors in the conditional standard deviations and covariances are larger than when the true VECH model is fitted. Finally, the conditional covariances are underestimated when they are large. Therefore, in the particular DGP considered in this illustration and for the particular time series generated, the restrictions imposed by the DVECH model have greater influence on the estimation of the conditional standard deviations and covariances than on conditional correlation estimates.

1.2.2 BEKK models

Engle and Kroner (1995) propose the BEKK (after Baba-Engle-Kraft-Kroner) model that guarantees positivity of the conditional covariance matrices. The BEKK(1,1,K) is given by

$$\mathbf{H}_t = \mathbf{C} + \sum_{k=1}^{K} \mathbf{A}_k' \mathbf{r}_{t-1} \mathbf{r}_{t-1}' \mathbf{A}_k + \sum_{k=1}^{K} \mathbf{B}_k' \ \mathbf{H}_{t-1} \mathbf{B}_k, \tag{5}$$

where $\mathbf{C}$, $\mathbf{A}_k$ and $\mathbf{B}_k$ are square $N \times N$ parameter matrices, with $\mathbf{C}$ being a positive definite symmetric matrix and K determines the generality of the process. The positivity of $\mathbf{H}_t$ is guaranteed by the parametrization of the model.[5] Denoting by $\otimes$ the Kronecker product of two matrices, the BEKK model is covariance stationary if and only if the eigenvalues of $\sum_{k=1}^{K} \mathbf{A}_k \otimes \mathbf{A}_k + \sum_{k=1}^{K} \mathbf{B}_k \otimes \mathbf{B}_k$ are less than one in modulus; see Engle and Kroner (1995). The conditions for strict stationarity and geometric ergodicity can be found in Boussama *et al.* (2011).

Engle and Kroner (1995) show that all BEKK models are representable as VECH models and that the BEKK parametrization eliminates very few if any interesting model allowed by the VECH representation. Indeed, Scherrer and Ribarits (2007) and Stelzer (2008) show that, when $N = 2$, both specifications are equivalent. Nonetheless, when $N > 3$, the VECH model allows for more flexibility than the BEKK model; see, for example, Stelzer (2008) who presents a three-dimensional VECH model with no BEKK representation. It is important to note that the equivalence between the BEKK and VECH models can be established when $K > 1$. However, the version of the BEKK model predominantly fitted in practice restricts $K = 1$;[6] see, for instance, Hafner and Herwartz (2006), Silvennoinen and Teräsvirta (2009a), Rossi and Spazzini (2010), Caporin and McAleer (2012), Laurent *et al.* (2012), Pedersen and Rahbek (2014) and Burda (2015). In this case, the VECH and BEKK models are not equivalent even if $N = 2$.

Comte and Lieberman (2003) and Hafner and Preminger (2009) (as a special case of the VECH model) prove consistency and asymptotic normality for the G-QML estimator under eighth and sixth finite moments of the observed variables, respectively. Avarucci *et al.* (2013) argue that finite fourth order moment restrictions for the G-QML estimator cannot be relaxed, even in the simple ARCH form of the BEKK model. However, the G-QML estimation is computationally demanding due to the need for several matrix inversions and, since the model is not linear on the parameters, it is difficult to obtain convergence. Alternatively, Boudt and Croux (2010) show the good robustness properties of an M-estimator with a fat-tail Student-ν loss function.

Pedersen and Rahbek (2014) study the asymptotic properties of the VT estimator of the following specification of the BEKK model

$$\mathbf{H}_t = \mathbf{\Sigma} - \mathbf{A}\mathbf{\Sigma}\mathbf{A}' - \mathbf{B}\mathbf{\Sigma}\mathbf{B}' + \mathbf{A}'\mathbf{r}_{t-1}\mathbf{r}'_{t-1}\mathbf{A} + \mathbf{B}'\ \mathbf{H}_{t-1}\mathbf{B}, \tag{6}$$

and establish its strong consistency under finite second-order moments and asymptotic normality under finite sixth-order moments. Note that these conditions are identical to those of the G-QML estimator; see Hafner and Preminger (2009). However, as pointed out by Caporin and McAleer (2012), when estimating the BEKK model by VT, it is extremely complicated imposing positive definiteness and covariance stationary restrictions. Recently, Burda (2015) deals with this shortcoming and suggests an approach based on Constrained Hamiltonian Monte Carlo that solves both the nonlinear constraints resulting from BEKK targeting and the complex form of the BEKK likelihood in relatively large dimensions. They conclude that the BEKK model estimated by VT presents an effective way of reducing dimensionality without compromising the model fit.

Alternatively, Noureldin *et al.* (2014) propose a rotated version of the BEKK model, denoted by rotated BEKK (RBEKK), based on fitting a covariance-targeting BEKK-type specification to the rotated returns, $\mathbf{e}_t = \mathbf{\Sigma}^{-1/2}\mathbf{r}_t$, as follows

$$\mathbf{G}_t = \mathbf{I}_N - \mathbf{A}\mathbf{A}' - \mathbf{B}\mathbf{B}' + \mathbf{A}'\mathbf{e}_{t-1}\mathbf{e}'_{t-1}\mathbf{A} + \mathbf{B}'\ \mathbf{G}_{t-1}\mathbf{B}, \tag{7}$$

where $\mathbf{G}_t$ is the conditional covariance matrix of $\mathbf{e}_t$ and $\mathbf{G}_1 = \mathbf{I}_N$. The RBEKK model can be written as a restricted BEKK model with parameters $\overline{\mathbf{A}} = \mathbf{\Sigma}^{1/2}\mathbf{A}\mathbf{\Sigma}^{-1/2}$, $\overline{\mathbf{B}} = \mathbf{\Sigma}^{1/2}\mathbf{B}\mathbf{\Sigma}^{-1/2}$ and $\overline{\mathbf{C}} = \mathbf{\Sigma}^{1/2}(\mathbf{I}_N - \mathbf{A}\mathbf{A}' - \mathbf{B}\mathbf{B}')\mathbf{\Sigma}^{1/2}$.

The RBEKK model can be easily estimated by VT, ensuring positiveness and covariance stationary. In the first step, $\mathbf{\Sigma}$ is estimated by the sample unconditional covariance matrix of $\mathbf{r}_t$, denoted by $\hat{\mathbf{\Sigma}}$. Using the spectral decomposition of this estimate, $\mathbf{e}_t$ is computed. The second step is based on estimating the parameters in (7) by G-QML, conditional on the estimates of $\mathbf{e}_t$ obtained in the first step.

As mentioned above, even if K is restricted to be 1, fully parametrized BEKK models are only feasible for small systems. Two restricted popular versions of the BEKK model that reduce the number of parameters are the diagonal BEKK (DBEKK) and the scalar BEKK (SBEKK) models, which are obtained by restricting the $\mathbf{A}_k$ and $\mathbf{B}_k$ matrices in (5) to be diagonal and scalar, respectively; see Ding and Engle (2001). The DBEKK model is representable as a DVECH model; see Bauwens *et al.* (2006). Consequently, the variances only depend on their own lags and past squared returns and the covariances only depend on their own lags and past cross products of returns. McAleer *et al.* (2008) show that the DBEKK and SBEKK models can be derived as multivariate extensions of the random coefficients autoregression of Tsay (1987). Noureldin *et al.* (2014) also study restricted versions of the RBEKK model denoted as diagonal RBEKK (D-RBEKK) and scalar RBEKK (S-RBEKK) models.

Consider again the same simulated system described in subsection 1.1. Figure 3 plots the estimated conditional standard deviations, covariances and correlations obtained after estimating the parameters of the BEKK model by G-QML and VT, while Figure 4 plots the same quantities after estimating the parameters of the DBEKK and SBEKK models by G-QML and of the D-RBEKK and S-RBEKK models by VT. For the sake of comparison, we also plot the simulated conditional standard deviations, covariances and correlations in both figures. Figure 3 shows that the BEKK model estimated by VT has slight larger errors than the BEKK model estimated by G-QML. However, the errors in the BEKK model are much larger than in the VECH model (see Figure 1) and, surprisingly, larger than in the DBEKK, SBEKK, D-RBEKK and S-RBEKK models plotted in Figure 4, which are restricted versions of the BEKK model. Note that Rossi and Spazzini (2010) also find the counterintuitive result that the SBEKK model has a better performance than the less restrictive DBEKK model. A possible explanation is that the non-linear restrictions in the parameters imposed by the BEKK model do not hold in practice for volatilities and conditional correlations. These restrictions could be even worse than considering that the matrices in (5) are diagonal or scalar. Figure 4 shows that the errors of the estimated conditional correlations are larger than when estimating the true DGP, besides the fact that, when the conditional correlation are small, they are underestimated. However, the errors obtained when estimating conditional volatilities and covariances are similar in magnitude to those obtained when the true DGP is fitted. Furthermore, Figure 4 shows that the estimates of conditional standard deviations, covariances and correlations are very similar regardless of whether the DBEKK, SBEKK, D-RBEKK or S-RBEKK models are fitted.

1.3 Conditional correlation models

Instead of modeling directly the conditional covariance matrix, $\mathbf{H}_t$, many authors propose specifying it as the following product of conditional variances and correlations

$$\mathbf{H}_t = \text{diag}(\mathbf{H}_t)^{1/2}\mathbf{R}_t\text{diag}(\mathbf{H}_t)^{1/2}, \tag{8}$$

where $\text{diag}(\mathbf{H}_t) = \text{diag}(h_{11,t}, \cdots, h_{NN,t})$ is a diagonal matrix whose elements are the conditional variances of each series and $\mathbf{R}_t$ contains the conditional correlations between the series. In this way, assuming that conditional variances and correlations are not related, it is possible to simplify the estimation, estimating first the conditional variances and second the conditional correlations. This two-step (2s) estimation procedure is relatively simple allowing working with high dimensional systems. Carnero and Eratalay (2014) show that, if the innovations are Gaussian, estimating the parameters in multiple steps has a very similar performance to that of the QML estimator.

The original conditional correlation models, like the CCC and DCC, assume that each conditional variance has a GARCH-type specification. Of course, it is very important to choose accurate GARCH-type specifications for these individual volatilities. For example, Audrino (2006) and Laurent *et al.* (2012) compare different specifications focusing on their impact on the accuracy of the conditional covariance estimates. In this paper, we focus on the GARCH(1,1) model.

There are two main types of models for conditional correlations depending on whether they are assumed to be constant or time-varying, namely the CCC and DCC models.

1.3.1 CCC models

Bollerslev (1990) introduces the CCC model, assuming that the conditional correlation matrix is constant over time, that is, $\mathbf{R}_t = \mathbf{R}$, where, $\mathbf{R}$ is a symmetric positive definite matrix; see, for example, Laurent *et al.* (2013) and Amado and Teräsvirta (2014) for recent empirical applications. The matrices $\mathbf{H}_t$ are definite positive if and only if all the conditional variances h_{iit}, $i = 1, \cdots, N$, are positive and $\mathbf{R}$ is a positive definite matrix.

Allowing for feedback among volatilities, Jeantheau (1998) proposes an extension of the original CCC model, the extended CCC (ECCC) model, whose individual volatilities are fitted by a vector GARCH(1,1) model, as follows

$$\mathbf{h}_t = \mathbf{c} + \mathbf{A}\ \mathbf{r}_{t-1}^2 + \mathbf{B}\ \mathbf{h}_{t-1}, \tag{9}$$

where $\mathbf{c}$ is an N-dimensional vector, $\mathbf{A}$ and $\mathbf{B}$ are $N \times N$ parameter matrices, $\mathbf{r}_{t-1}^2 = (r_{11,t}^2 \dots r_{NN,t}^2)'$ and $\mathbf{h}_t = (h_{11,t} \dots h_{NN,t})'$; see also Caporin (2007) who allows for volatility feedback. The ECCC model is covariance and strictly stationary if the moduli of the eigenvalues of $\mathbf{A} + \mathbf{B}$ are less than one; see He and Teräsvirta (2004). Aue *et al.* (2009) establish a sufficient condition for strict stationarity and the existence of fourth-order moments. On the other hand, Nakatani and Teräsvirta (2008) establish sufficient positivity conditions. Nakatani and Teräsvirta (2009) suggest a procedure for testing the hypotesis of a diagonal structure against the hypothesis of volatility feedbacks. Jeantheau (1998) also proves the strong consistency of the G-QML estimator for the ECCC model, while Ling and McAleer (2003) prove its asymptotic normality. More recently, Francq and Zakoïan (2012) establish the strong consistency and asymptotic normality of the G-QML estimator under mild conditions that coincide with the minimal ones in the univariate case. Finally, Francq *et al.* (2014) establish the strong consistency and the asymptotic normality of the VT estimator.

The first column of Figure 5 plots the estimated conditional standard deviations, covariances and correlations obtained after fitting the CCC and ECCC models to the same simulated system described in subsection 1.1, while the second column plots the real values versus the estimated values considering alternative procedures. The errors in the conditional correlation estimates of the CCC and ECCC models are notably greater than the errors of the full model. However, when compared with the errors plotted in Figure 3, we can observe that the errors are smaller than when the BEKK model is fitted and similar to those plotted in Figure 4 for the DBEKK and SBEKK models. Furthermore, Figure 5 illustrates that the errors when fitting the CCC and ECCC models are similar when looking at conditional correlations, but the ECCC model have smaller errors than the CCC model for the conditional standard deviations of the second series and for the conditional covariances.

Conrad and Karanasos (2010) propose a further extension that also allows for negative feedback among volatilities and derive necessary and sufficient conditions for the positive definiteness of the covariance matrix. It is important to point out that Conrad and Karanasos (2010) claim that their results are also valid for models in which the correlations are time-varying which will be considered latter in this paper. Most of the results obtained by them are referred to bivariate models. Whether they are useful in large systems is an open question.

1.3.2 DCC models

Assuming constant conditional correlations is not reasonable in many practical situations; see, for example, Longin and Solnik (1995) and Tse (2000) for early references. Consequently, several authors suggest models in which the correlations are time-varying; see Engle (2002), Tse and Tsui (2002) and Silvennoinen and Teräsvirta (2009b).

Because of its popularity, in this paper, we focus on the (scalar) DCC model of Engle (2002) and its consistent correction by Aielli (2013), known as cDCC, which is given by

$$\mathbf{R}_t = \text{diag}(\mathbf{Q}_t)^{-1/2}\mathbf{Q}_t \ \text{diag}(\mathbf{Q}_t)^{-1/2}, \tag{10}$$

$$\mathbf{Q}_t = (1 - a - b)\mathbf{S} + a \ \text{diag}(\mathbf{Q}_{t-1})^{1/2}\mathbf{u}_{t-1}\mathbf{u}'_{t-1}\text{diag}(\mathbf{Q}_{t-1})^{1/2} + b\mathbf{Q}_{t-1}. \quad (11)$$

where $\mathbf{Q}_t$ is an $N \times N$ positive definite matrix, $\mathbf{u}_t = (u_{1t}, \cdots, u_{Nt})'$ with $u_{it} = r_{it}/\sqrt{h_{iit}}$ being the standardized correlated returns, $\mathbf{S}$ is the unconditional covariance matrix of $\text{diag}(\mathbf{Q}_{t-1})^{1/2}\mathbf{u}_t$ and a and b are scalars. In order for $\mathbf{H}_t$ to be a positive definite matrix, it is sufficient that all conditional variances be positive, $a, b > 0$ and $a + b < 1$. Aielli (2013) also proves that if $\mathbf{S}$ is positive definite and $a + b < 1$, the correlated and the standardized return process are strictly and covariance stationary. The cDCC model in equations (10)-(11) is very popular to describe and forecast multivariate conditional heteroscedastic time series; see, for example, Engle and Kelly (2012), Bauwens *et al.* (2013), Aielli and Caporin (2014) and Audrino (2014) for empirical applications.

The comparison of the scalar versions of the BEKK and DCC models has been carried out by Caporin and McAleer (2008). Empirically, they found that both models are very similar in forecasting conditional variances, covariances and correlations. Latter, Caporin and McAleer (2012) compare the BEKK and DCC models from a theoretical point of view and conclude that the BEKK model could be used to obtain consistent estimates of the DCC model with a direct link to the indirect DCC model suggested in Caporin and McAleer (2008); see also Caporin and McAleer (2014) who carry out an empirical comparison.

The parameters in equations (8)-(11) can be estimated by the three-step (3s) estimator described by Aielli (2013); see also Caporin and McAleer (2012) for a description of available asymptotic results on the estimation of the parameters of DCC models. Although the cDCC model avoids the bias problem when estimating $\mathbf{S}$, the three-step estimator is downward biased. Pakel *et al.* (2014) propose the composite likelihood (CL) estimator which is based on summing up the quasi-likelihood functions of subsets of assets, allowing to estimate models even when the cross-sectional dimension is larger than the sample size. Alternatively, Hafner and Reznikova (2012) suggest a reduction of the bias by using shrinkage techniques applied to the sample covariance matrix of the standardized residuals.

Noureldin *et al.* (2014) also apply the rotation technique to the DCC model of Engle (2002), resulting in the rotated DCC (RDCC) model. They conclude empirically that the RDCC model, which is feasible in large dimensions, performs better than the orthogonal GARCH (OGARCH) of Alexander (2001), the generalized OGARCH (GOGARCH) of Boswijk and van der Weide (2011) and the RBEKK models.

Next, we illustrate the performance of the cDCC and RDCC models when fitted to the simulated system described in subsection 1.1. The first column of Figure 5 plots, together with the simulated conditional standard deviations, covariances and correlations and the estimates of the CCC and ECCC models, the corresponding estimates obtained after fitting the cDCC and RDCC models, while the second column plots the real values versus the estimated values.

Note that the conditional standard deviations of the CCC, cDCC and RDCC models are exactly the same given that they are estimated separately in the first step. The cDCC and RDCC models seem to have slightly larger errors than the VECH model, but lower errors than the CCC and ECCC models. Also, the RDCC and cDCC models tend to underestimate the correlations when they are large and underestimate when they are small. On the other hand, the estimated conditional covariances of the cDCC and RDCC models seem to be robust to the misspecification.

Equation (11) imposes a common dynamic structure for all conditional correlations, governed by the parameters a and b. This might not be realistic when pairwise correlations between different returns have different behaviors. To avoid this constraint, several generalizations of the DCC model of Engle (2002) have been proposed. For example, Billio *et al.* (2006) introduce a block-diagonal structure, where the dynamics are restricted to be equal only among certain groups of variables, and a BEKK structure on the conditional correlations is proposed. Bauwens *el al.* (in press) estimate the cDCC model in (11) when a and b are replaced by matrices using a Bregman-proximal trust-region method and conclude empirically that the use of richly parametrized models have better performance than the scalar case. Hafner and Franses (2009) also extend the DCC model by allowing the parameters to vary across assets and Otranto (2010) proposes a clustering algorithm to identify similar structures of correlation dynamics in the DCC models. Finally, several authors propose different short- and long-run sources that affect correlations; see, for example, Colacito *et al.* (2011), Rangel and Engle (2012), Audrino and Trojani (2011) and Audrino (2014). Also, Silvennoinen and Teräsvirta (2009b, 2015) propose the smooth transition conditional correlation (STCC) model allowing the correlations to vary smoothly between two different states.

The cDCC model has been extended by Bauwens and Otranto (2013) to include volatility as determinant of correlations by introducing measures of volatility as exogenous variables. On the other hand, Palandri (2009) proposes breaking the conditional correlation matrices into the product of a sequence of matrices in such a way that they preserve positive definiteness without imposing constraints on the parameters. The sequential conditional correlations (SCC) model separates the correlations and partial correlations, allowing for a multi-step estimation procedure. Consequently, very complex optimization problems are converted into a series of mere univariate and bivariate estimations, which enables working with very high dimensional systems and at the same time complex functional forms for the conditional correlation process. However, the SCC model still assumes that variances and correlations are not related between them. Finally, Boudt *et al.* (2013) propose a robust extension of the model, known as BIP-cDCC model, for forecasting correlations in the presence of one-off events which cause large changes in prices whilst not affecting the volatility dynamics. They apply the new model to daily returns of the EUR/USD and Yen/USD exchange rates and conclude that the BIP-cDCC model is always better or have similar performance in relation to the cDCC model when forecasting future covariance matrices at different forecast horizons.

2 Monte Carlo simulation

In this section, we carry out Monte Carlo experiments to analyze the finite sample properties of the estimated conditional variances, covariances and correlations obtained after fitting a restricted specification to systems in which all variances and covariances are interrelated with each other. We also consider alternative estimators available in the literature to estimate the pseudo-parameters in the restricted specifications considered. We simulate 500 replicates of sizes $T = 1000$ and 2000 by the bivariate VECH model described in subsection 1.1 with Gaussian and Student-7 errors. Table 1 summarizes the models fitted for each replicate, all of them including just one lag of past returns, as well the alternative estimators considered. The O-GARCH model of Alexander (2001) and the GO-GARCH model of Boswijk and van der Weide (2011) are include in the experiments, since there is a growing literature on multivariate conditionally heteroscedastic factor models[7]. In the cases where the constraints to ensure covariance stationarity and/or positivity are non-linear in the parameters, we do not restrict the parameters to maximize the likelihood. In the cDCC and RDCC models, we restrict the maximization by imposing $a + b < 1$ and $0 < a, b < 1$ in equation 11, in addition to the usual restrictions to estimate the GARCH models. In the DBEKK model estimated by QML, it is imposed $a_{ii}^2 + b_{ii}^2 < 1$, $0 < a_{ii}, b_{ii} < 1$, for $i = 1, 2$, and $\mathbf{C} > 0$, where a_{ii} and b_{ii} are the diagonal elements of the $\mathbf{A}_1$ and $\mathbf{B}_1$ matrices in (5). In the SBEKK model estimated by QML, it is imposed that $a^2 + b^2 < 1$, $0 < a, b < 1$ and $\mathbf{C} > 0$, where a and b are given in (5), when the $\mathbf{A}_1$ and $\mathbf{B}_1$ matrices are replaced by these scalars. In the R-DBEKK model, it is imposed $a_{ii}^2 + b_{ii}^2 < 1$ and $0 < a_{ii}, b_{ii} < 1$, for $i = 1, 2$, where a_{ii} and b_{ii} are the diagonal elements of the $\mathbf{A}$ and $\mathbf{B}$ matrices in (7). Finally, in the S-RBEKK model, it is imposed $a^2 + b^2 < 1$ and $0 < a, b < 1$, where a and b are given in (7), when $\mathbf{A}$ and $\mathbf{B}$ matrices are replaced by these scalars. All restrictions mentioned above are sufficient conditions to ensure covariance stationarity and positivity of the corresponding models.

After estimating the parameters, Table 1 reports the number of replicates in which stationarity and positivity of covariances matrices are not empirically satisfied. We consider that a fitted model is not positive for some replicate if at least one of the conditional covariance matrices is not positive defined. On the other hand, we verify whether the parameter estimates satisfy the sufficient restrictions to ensure covariance stationarity. Note that, the estimated parameters of all rotated models, the (G)O-GARCH model and the models based on representing conditional correlations, except for the ECCC model, always satisfy the stationarity and positivity restrictions, as we impose it in the estimation process. Moving on to the results of the VECH, BEKK and ECCC models, the stationarity conditions are violated in a relatively large number of replicates if the parameters are estimated by QML. The number of violations of covariance stationarity increases when the DGP is Student, regardless of whether the Gaussian or the Student likelihoods are maximized. Still, the number of covariance stationarity violations decreases with the sample size. One alternative to ensure covariance stationarity of the VECH model would be to consider the Bregman divergences based optimization method defined in Chrétien and Ortega (2014). However, we do not consider it because of the high computational effort. On the other hand, when the parameters of the VECH, BEKK and ECCC models are estimated by VT, the estimates always satisfy the covariance stationarity, except for one particular case in the ECCC model. However, the estimates of covariance matrices by VT are not positive in a relative large number of replicates, while QML estimates are in the majority of the cases. Finally, when the DVECH, DBEKK or SBEKK models are fitted, we can observe that, regardless of the estimator, they always satisfy covariance stationarity and positivity restrictions, except when the DGP is Student and the DVECH model is estimated by QML-G or QML-S, where covariance stationarity conditions are violated in some cases. It is important to emphasize that although the BEKK, DBEKK and SBEKK models are positive by definition, when they are estimated by VT without restrictions, positivity is not ensured; see Caporin and McAleer (2012). An alternative procedure to estimate BEKK models by VT would be to consider the Constrained Hamiltonian Monte Carlo defined by Burda (2015).

Table 1 also reports the average computer time involved in the estimation. In each estimation that does not converge for an initial value, we try alternative initial values until it converges; see Asai (2013) and Chrétien and Ortega (2014) for methods to choose initial values for the BEKK and VECH models, respectively.[8] It is obvious that VT is faster than G-QML and S-QML with G-QML being faster than S-QML. In the DVECH model, the LSW estimator is even faster and does not have problems with convergence. When comparing the DBEKK and SBEKK models with their rotated versions, we can observe that the computer time involved in the estimation of the rotated cases is lower and, in addition, they ensure positivity and stationarity. Therefore, it seems that it is worth using the rotated BEKK models rather than non-rotated BEKK models.

For each replicate and estimator considered, the performance of the estimated conditional covariance and correlation matrices is measured by the following Frobenius norms

$$LF_1 = \frac{\sum_{t=1}^{T} \mathrm{Tr}[(\hat{\mathbf{H}}_t - \mathbf{H}_t)'(\hat{\mathbf{H}}_t - \mathbf{H}_t)]}{T}, \quad LF_2 = \frac{\sum_{t=1}^{T} \mathrm{Tr}[(\hat{\mathbf{R}}_t - \mathbf{R}_t)'(\hat{\mathbf{R}}_t - \mathbf{R}_t)]}{T} \qquad (12)$$

where $\hat{\mathbf{H}}_t$ and $\hat{\mathbf{R}}_t$ are the estimated conditional covariance and correlation matrices at time t; see Laurent *et al.* (2013) for a comprehensive list of different loss functions and their impacts on ranking forecasting performances of MGARCH models.

Table 2, which reports the medians of LF_1 and LF_2, shows that, regardless of the estimation method, these statistics have similar values when the true DGP is fitted, with VT being, in general, slightly worse estimating correlations and better estimating covariances. Moving on to the restricted misspecified models, we observe that, for a given model, different estimators lead to similar results. Therefore, if the objective is the estimation of variances, covariances and correlations the estimator should be chosen in terms of computational advantages.

Hereafter we focus on comparing different restricted models. First, we observe that although the DVECH model has a LF_1 statistic much larger than when the true VECH model is fitted, this difference is not so large when estimating correlations. Also note that increasing the sample size from $T = 1000$ to 2000, reduces the distance when the true model is fitted but not necessarily when the DVECH model if fitted. Second, the errors corresponding to the BEKK model are much larger compared to the DVECH model. However, when fitting the more restrictive DBEKK model, the results are improved with respect to its full version. Surprisingly, when the even more restrictive SBEKK model is fitted the results are even better for conditional covariance matrices. We also note that the performance of the DVECH model is similar to the restricted BEKK-type models when estimating covariances and better when estimating correlations. Third, the rotation of BEKK-type models only improves marginally the estimated covariance matrices of the DBEKK and SBEKK models. However, it is important to note that the computer time involved in the estimation of the rotated cases is lower and the rotated models ensure positivity of covariance matrices and covariance stationarity; see Table 1. Finally, when the O-GARCH and GO-GARCH models are fitted, we observe that the restrictions imposed in these models generate estimates of conditional variances, covariances and correlations with larger distances than those of the DVECH or restricted BEKK-type models.

Turning now to the results of the models that represent the dynamics of conditional correlations, we can see that the errors corresponding to the cDCC and RDCC models are very similar between them and similar, in terms of correlations, to the DVECH model. On the other hand, the cDCC and RDCC are much better than the DVECH model and BEKK-type models in terms of covariances. When estimating covariances, the ECCC and cDCC models are comparable and the difference with respect to the true DGP is 50%. Finally, when estimating correlations, the cDCC and RDCC models remarkably outperform the constant correlation models.

As conclusion, for the set of parameters considered in the simulation, the restrictions imposed by a misspecified model are much more relevant than the choice of the estimation method. Moreover, all alternative models have notably inferior performances in relation to the VECH model, specially when $T = 2000$, and among the alternative models, the cDCC and RDCC models have, in general, a superior performance. Indeed, Laurent *et al.* (2012) conclude empirically that it is very difficult to outperform the DCC model after comparing 125 models fitted to forecast the correlations of a system of 10 assets from the New York Stock Exchange. Caporin and

McAleer (2014) also support empirically the preference for correlation models over covariance models when analyzing different cross-sectional dimensions from 5 up to 89 assets.

3 Empirical application

In this section, we compare empirically the performance of several MGARCH models fitted to a system of five exchange rate returns, namely Euro (EUR), British Pound (GBP), Swiss Franc (CHF), Australian Dollar (AUD) and Japanese Yen (JPY) against the US Dollar (USD). The data are daily closing exchange rates observed at 12:00 AM (New York time) from January 2, 2004 to December 31, 2013, with a total of 2582 daily observations. The full-sample period is split into estimation in-sample and forecast out-of-sample periods. We consider three different splits in such way that the out-of-sample periods are the last one, six and five years, respectively. In the first case, the in-sample period contains the crises period and has large heteroscedasticity and the out-of-sample period is very calm and has low heteroscedasticity. In the second split, the estimation period is a low volatility period and has low heteroscedasticity, whereas the forecasting period has extreme market conditions and large heteroscedasticity. Finally, in the last split, both the in-sample and out-of-sample periods have periods of low and high volatilities. We define the exchange rate returns as usual by $r_{t,j} = 100 \times \log(y_{t,j}/y_{t-1,j})$, $j = 1, 2, \cdots, 5$, where $y_{t,j}$ is the daily exchange rate of the j-th series at time t. Figure 6 plots the returns for the full sample returns. The vertical lines show the three splits considered. Table 3 reports descriptive statistics and the 20-lag Ljung-Box statistics for returns ($Q(20)$) and squared returns ($Q_2(20)$) for the overall period and all in-sample and out-of-sample periods considered. The traditional features of returns like zero mean, skewness and excess of kurtosis are present in all the currencies for most periods. According to the Ljung-Box statistics, the returns and squared returns are significantly autocorrelated for the overall, in-sample and out-of-sample periods, except for the first out-of-sample period. Although the Ljung-Box statistic for serial correlation of returns is significant for the first and third splits, an analysis of the sample autocorrelation functions shows that the magnitudes of the correlations are very small, and generally not significant in the first two lags and multiple of five lags. Consequently, we fit only MGARCH models without any dependence in the conditional mean. In particular, we consider the DVECH model estimated by the LSW procedure, the DBEKK model estimated by VT, the SBEKK model estimated by G-QML, S-QML and VT, the D-RBEKK and S-RBEKK models estimated by VT and S-VT and the CCC, cDCC and RDCC models estimated by Gaussian and Student multi-step procedures. As an illustration, Figure 7 plots the estimated pairwise conditional correlation after fitting the R-DBEKK and cDCC models. We can observe that the correlations estimated by these models are rather similar, which is consistent with Caporin and McAleer (2008, 2014).

The out-of-sample predictions are evaluated using a the rolling window procedure as proposed by Giacomini and White (2006). For each model and estimation method, W_1 is the first window which incorporates the first T_1 sample observations. Denote by $\hat{\theta}_1$ the parameter estimates for this window, and compute h-step-ahead forecasts for $h = 1, 5, 20$. For the next four windows, W_2 to W_5, add a new observation and use the estimate θ_1 from the previous window, i.e., $\hat{\theta}_5 = \cdots = \hat{\theta}_2 = \hat{\theta}_1$ to compute the forecasts. For window W_6 we add the $(T_1 + 5) - th$ observation, the first five observations are dropped and the new parameter estimate is obtained, $\hat{\theta}_6$. For the next four windows repeat the same procedure used for windows W_2 to W_5. This whole procedure is repeated until windows W_{258}, W_{1550} and W_{1290} in the first, second and third splits, respectively. However for the last four observations of each case, only one-step-ahead forecasts are obtained, as we cannot compare the forecasts with the observed values for five- and twenty-step-ahead.

In order to compare the out-of-sample forecasts, we consider the out-of-sample negative log-likelihood (NL), used by Audrino (2014), which is given by

$$\mathrm{NL}(h) = -\sum_{i=0}^{n_{pred,h}-1} \log[L(\mathbf{r}_{T+i+h}; \hat{\theta}_{i+1}, \hat{\mathbf{H}}_{T+i}(h))], \tag{13}$$

where $L(.;.)$ is the conditional likelihood, $n_{pred,h}$ is the number of h-step-ahead predictions, $\hat{\theta}_{i+1}$ is the parameter estimate of the window which includes $\mathbf{r}_{T+i}$ in the last five observations of the sample, and $\hat{\mathbf{H}}_{T+i}(h)$ is the h-step-ahead forecast of $\mathbf{H}_{T+i+h}$. The prediction $\hat{\mathbf{H}}_{T+i}(h)$ is obtained using observations up to the $T + i$ and evaluated at $\hat{\theta}_{i+1}$. This method is reasonable to compare the predictive accuracy of models since the parameters of the models are estimated in-sample using the same functions. Furthermore, the NL has the strength of not considering any proxies for the unobservable covariance matrices.

The comparison of models is also carried out using the following Frobenius loss

$$LF_1(h) = \frac{\sum_{i=0}^{n_{pred,h}-1} \mathrm{Tr}[(\mathbf{H}_{T+i+h} - \hat{\mathbf{H}}_{T+i}(h))'(\mathbf{H}_{T+i+h} - \hat{\mathbf{H}}_{T+i}(h))]}{n_{pred,h}}, \tag{14}$$

$$LF_2(h) = \frac{\sum_{i=0}^{n_{pred,h}-1} \mathrm{Tr}[(\mathbf{R}_{T+i+h} - \hat{\mathbf{R}}_{T+i}(h))'(\mathbf{R}_{T+i+h} - \hat{\mathbf{R}}_{T+i}(h))]}{n_{pred,h}}. \tag{15}$$

As the true covariance and correlation matrices, $\mathbf{H}_t$ and $\mathbf{R}_t$, respectively, are unobservable, we use the realized covariances and correlations as a proxy. The LF has the advantage of belonging to the class of loss functions robust to noises in volatility proxy; see Patton and Sheppard (2009). It is widely recognized that the estimation of realized covariances and correlations suffers from asynchronous trading and market microstructure noise, causing the covariance and correlation estimators to be biased and inconsistent; see, for instance, McAleer and Medeiros (2008), Patton (2011) and Corsi and Audrino (2012). We sample the intraday returns into 288 five-minute intervals[9] to avoid the asynchronous effect and compute the realized covariance by the Realized

Outlyingness Weighted Covariance (rOWCov) of Boudt et al. (2011)[10]; see Barndorff-Nielsen *et al.* (2011), Corsi and Audrino (2012) and Boudt *et al.* (2012) for alternative proxies.

Finally, we perform the superior predictive ability (SPA) test of Hansen (2005) and the model confidence set (MCS) of Hansen *et al.* (2011) to verify whether the performance of alternative models and methods are significantly different according to the NL and LF criteria.[11] The first test allows for multiple comparison against a pre-specified benchmark model, with the null hypothesis being that each model is not outperformed by at least one of the other competing models. The second one chooses from an initial set, a subset of forecasts that outperforms all the other alternatives.

Table 4 reports the observed NL values, the corresponding p-values of the SPA tests and the MCS with $\alpha = 10\%$ significance level, corresponding to the the first split for the different models and estimation methods considered. For the three forecasting horizons, the models estimated by maximizing the Student likelihood have notably better performance than models estimated by maximizing the Gaussian likelihood. Indeed, with the exception of the CCC model, all the models estimated by Student likelihood are not significantly outperformed in the SPA test at 5% confidence level, while all the models estimated by Gaussian likelihood are outperformed for one- and twenty-step-ahead. The SBEKK estimated by S-QML has the best performance for one- and five-step-ahead, while the RDCC estimated by S-3s is the better for twenty-step-ahead. The CCC estimated by G-QML and the DVECH estimated by LSW are the worst fits.

Table 4 also reports the performance of the models and methods based on the LF statistics. Although the models estimated by Student likelihood have in general a better performance in relation to the models estimated by Gaussian likelihood, the difference between the estimation methods are not so large as in the case of LF. Contrary to the NL statistics, the time-varying conditional correlation models and the R-DBEKK model have the better performance according to the LF_1 criterion and the time-varying conditional correlation models according to the LF_2 criterion. Furthermore, using the LF_1 and LF_2 criteria, the only model included in the MSC for all forecasting horizons is the RDCC estimated by S-2s.

Table 5 reports the performance of the alternative models and estimators considered when implemented to forecast volatilities, covariances and correlations during the out-of-sample periods corresponding to the second and third splits. Given that we do not have high frequency data for these periods, Table 5 only reports the results corresponding to the out-of-sample negative log-likelihood. Once again, all the models estimated by maximizing the Student likelihood have notably better performances than those estimating by maximizing the Gaussian likelihood. The SBEKK model estimated by S-QML has the best performance for the three horizons considered and is the only model included in the MCS for all forecasting horizons in the third split and for the 1-step-ahead in the second split. Considering 5- and 20-step-ahead in the second period, the SBEKK estimated by S-QML has again the better performance, but the R-DBEKK model estimated by S-VT is also included in the MSC.

Therefore, it seems that regardless of the particular split analyzed and, consequently, of the behavior of the volatility in the in-sample and out-of-sample periods, the SBEKK model estimated by S-QML has a reasonable forecast performance when this is measured used the negative log-likelihood criterion; see also the empirical results of Burda (2015). However, when the forecasting performance is measured using the LF criteria, it seems that the R-DCC model is better during the low-volatility out-of-sample period. This result is in concordance with our simulation results.

4 Concluding Remarks

In this paper we discuss the main strengths and limitations of the most popular symmetric multivariate GARCH models available in the literature. A simulation study is carried out to compare different restricted specifications and estimators, when the series are generated by the general VECH model. The R-SBEKK, R-DBEKK, DCC and R-DCC models explain adequately the evolution of volatilities, conditional covariances and correlations generated by the general model. However, the BEKK fails to estimate conditional covariances and correlations. Furthermore, the performance of the forecasts of volatilities, conditional covariances and correlations are similar for a given model regardless of the particular procedure used to estimate the parameters. Therefore, it seems that the estimation method should be chosen using computational advantages. An empirical application to a five-dimensional system of exchange rates returns is carried out. We conclude that the SBEKK model estimated by S-QML has a reasonable forecast performance according to the negative log-likelihood criterion, while the R-DCC model outperforms the other models according to the LF criteria during the low-volatility out-of-sample period.

It is also of interest to compare the models considered in these paper with those based on copulas; see, for example, Patton (2006), Lee and Long (2009), So and Yeung (2014) and Creal and Tsay (in press). Furthermore, in this paper, we focus on symmetric models. However, there is a strong empirical evidence of asymmetries in the responses of conditional variances and covariances to positive versus negative past returns; see, for example, Bollerslev *et al.* (2006) for a comprehensive list of references with empirical evidence about the asymmetric response of volatility to past returns and Kroner and Ng (1998), Cappiello *et al.* (2006) and Caporin and McAleer (2011) for asymmetric response to simultaneous negative returns and simultaneous positive ones. Further research should focus on studying the economic empirical implications of the restrictions imposed on MGARCH models to reduce the number of parameters and/or to guarantee covariance stationarity and/or positiveness.

References

Aboura, S. and J. Chevallier. 2015. Volatility returns with vengeance: Financial markets vs. commodities. *Research in International Business and Finance* 33: 334–354.

Aielli, G.P. 2013. Dynamic conditional correlations: on properties and estimation. *Journal of Business & Economic Statistics* 31: 282–299.

Aielli, G.P. and M. Caporin. 2014. Variance clustering improved dynamic conditional correlation MGARCH estimators. *Computational Statistics & Data Analysis* 76: 556–576.

Alexander, C.O. 2001. Orthogonal GARCH. *Mastering Risk* 2: 21–38.

Alexander, C.O. and A.M. Chibumba. 1997. *Multivariate Orthogonal Factor GARCH*, Mimeo, University of Sussex.

Amado, C. and T. Teräsvirta. 2014. Conditional correlation models of autoregressive conditional heteroskedasticity with nonstationary GARCH equations. *Journal of Business & Economic Statistics* 32: 69–87.

Andersen, T.G., T. Bollerslev, P.F. Christoffersen and F.X. Diebold. 2007. Practical volatility and correlation modeling for financial market risk management. In M. Carey and R. Stulz (eds.), *The Risks of Financial Institutions*, 513–548. Chicago: University of Chicago Press.

Asai, M. 2013. Initial Values on Quasi-Maximum Likelihood Estimation for BEKK Multivariate GARCH Models. *Soka University, working paper.*

Attanasio, O.P. 1991. Risk, time-varying second moments and market efficiency. *The Review of Economic Studies* 58: 479–494.

Audrino, F. 2006. The impact of general non-parametric volatility functions in multivariate GARCH models. *Computational Statistics & Data Analysis* 50: 3032–3052.

Audrino, F. 2014. Forecasting correlations during the late-2000s financial crisis: short-run component, long-run component, and structural breaks. *Computational Statistics & Data Analysis* 76: 43–60.

Audrino, F. and F. Trojani. 2011. A general multivariate threshold GARCH model with dynamic conditional correlations. *Journal of Business & Economic Statistics* 29: 138–149.

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

Avarucci, M., E. Beutner and P. Zaffaroni. 2013. On moments conditions for quasi-maximum likelihood estimation of multivariate ARCH models. *Econometric Theory* 29: 545–566.

Bai, J. and Z. Chen. 2008. Testing multivariate distributions in GARCH models. *Journal of Econometrics* 143: 19–36.

Baillie, R.T. and R.J. Myers. 1991. Bivariate GARCH estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics* 6: 109–124.

Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde and N. Shephard. 2011. Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics* 162: 149–169.

Bauwens, L. and S. Laurent. 2005. A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models. *Journal of Business & Economic Statistics* 23: 346–354.

Bauwens, L. and E. Otranto. 2013. Modeling the dependence of conditional correlations on volatility (No. 2013014), Université Catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Bauwens, L., S. Laurent and J.V.K. Rombouts. 2006. Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21: 79-109.

Bauwens, L., C.M. Hafner and D. Pierret. 2013. Multivariate volatility modeling of electricity futures. *Journal of Applied Econometrics* 28: 743–761.

Bauwens, L., C.M. Hafner and J.V.K. Rombouts. 2007. Multivariate mixed normal conditional heteroskedasticity. *Computational Statistics & Data Analysis* 51: 3551–3566.

Bauwens, L., L. Grigoryeva and J.P. Ortega. In press. Estimation and empirical performance of non-scalar dynamic conditional correlation models. *Computational Statistics & Data Analysis.*

Beirne, J., G.M. Caporale, M. Schulze-Ghattas and N. Spagnolo. 2013. Volatility spillovers and contagion from mature to emerging stock markets. *Review of International Economics* 21: 1060–1075.

Billio, M., M. Caporin and M. Gobbo. 2006. Flexible dynamic conditional correlation multivariate GARCH models for asset allocation. *Applied Financial Economics Letters* 2: 123–130.

Bollerslev, T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31: 307–327.

Bollerslev, T. 1990. Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics* 72: 498–505.

Bollerslev T. and J.M. Wooldridge. 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews* 11: 143–172.

Bollerslev, T., R.F. Engle and J.M. Wooldridge. 1988. A capital asset pricing model with time varying covariances. *Journal of Political Economy* 96: 116–131.

Bollerslev, T., J. Litvinova and G. Tauchen. 2006. Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics* 4: 353–384.

Boswijk, H.P. and R. van der Weide. 2011. Method of moments estimation of GO-GARCH models. *Journal of Econometrics* 163: 118–126.

Boudt, K. and C. Croux. 2010. Robust M-estimation of multivariate GARCH models. *Computational Statistics & Data Analysis* 54: 2459–2469.

Boudt, K., C. Croux and S. Laurent. 2011. Outlyingness weighted covariation. *Journal of Financial Econometrics* 9: 657–684.

Boudt, K., J. Cornelissen and C. Croux. 2012. Jump robust daily covariance estimation by disentangling variance and correlation components. *Computational Statistics & Data Analysis* 56: 2993–3005.

Boudt, K., J. Danielsson and S. Laurent. 2013. Robust forecasting of dynamic conditional correlation GARCH models. *International Journal of Forecasting* 29: 244–257.

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

Burda, M. 2015. Constrained Hamiltonian Monte Carlo in BEKK GARCH with Targeting. *Journal of Time Series Econometrics* 7: 95-113.

Caporin, M. 2007. Variance (non) causality in multivariate GARCH. *Econometric Reviews* 26: 1–24.

Caporin M. and M. McAleer. 2008. Scalar BEKK and indirect DCC. *Journal of Forecasting* 27: 537–549.

Caporin M. and M. McAleer. 2011. Thresholds, news impact surfaces and dynamic asymmetric multivariate GARCH. *Statistica Neerlandica* 65: 125–163.

Caporin, M. and M. McAleer. 2012. Do we really need both BEKK and DCC? A tale of two multivariate GARCH models. *Journal of Economic Surveys* 26: 736–751.

Caporin, M. and M. McAleer. 2014. Robust ranking of multivariate GARCH models by problem dimension. *Computational Statistics & Data Analysis* 76: 172–185.

Caporin, M. and P. Paruolo. 2015. Proximity-structured multivariate volatility models. *Econometric Reviews* 34: 559–593.

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

Carnero, M. and M.H. Eratalay. 2014. Estimating VAR-MGARCH models in multiple steps. *Studies in Nonlinear Dynamics & Econometrics* 18: 339–365.

Chrétien, S. and J.P. Ortega. 2014. Multivariate GARCH estimation via a Bregman-proximal trust-region method. *Computational Statistics & Data Analysis* 76: 210–236.

Colacito, R., R.F. Engle, and E. Ghysels. 2011. A component model for dynamic correlations. *Journal of Econometrics* 164: 45–59.

Comte, F. and O. Lieberman. 2000. Second-order noncausality in multivariate GARCH processes. *Journal of Time Series Analysis* 21: 535–557.

Comte, F. and O. Lieberman. 2003. Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis* 84: 61–84.

Conrad, C. and M. Karanasos. 2010. Negative volatility spillovers in the unrestricted ECCC-GARCH model. *Econometric Theory* 26: 838–862.

Corsi, F. and F. Audrino. 2012. Realized covariance tick-by-tick in presence of rounded time stamps and general microstructure effects. *Journal of Financial Econometrics* 10: 591–616.

Creal, D.D. and R.S. Tsay. In press. High dimensional dynamic stochastic copula models. *Journal of Econometrics*.

Cripps, E. and W. Dunsmuir. 2003. Modeling the variability of Sydney Harbor wind measurements. *Journal of Applied Meteorology* 42: 1131–1139.

De Santis, G. and B. Gerard. 1997. International asset pricing and portfolio diversification with time-varying risk. *The Journal of Finance* 52: 1881–1912.

Diebold, F.X. and K. Yilmaz. 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal* 119: 158–171.

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

Engle, R.F. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50: 987–1007.

Engle, R.F. 2002. Dynamic conditional correlation - a simple class of multivariate GARCH models. *Journal of Business & Economic Statistics* 20: 339–350.

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

Engle, R.F. and K.F. Kroner. 1995. Multivariate simultaneous generalized ARCH. *Econometric Theory* 11: 122–150.

Engle, R.F. and J. Mezrich. 1996. GARCH for groups. *Risk* 9: 36–40.

Engle, R.F. and R. Colacito. 2006. Testing and valuing dynamic correlations for asset allocation. *Journal of Business & Economic Statistics* 24: 238–253.

Engle, R. and K. Sheppard. 2008. Evaluating the specification of covariance models for large portfolios. *New York University, working paper.*

Engle, R.F. and B. Kelly. 2012. Dynamic equicorrelation. *Journal of Business & Economic Statistics* 30: 212–228.

Engle R.F., V.K. Ng and M. Rothschild. 1990. Asset pricing with a factor ARCH covariance structure: empirical estimates for treasury bills. *Journal of Econometrics* 45: 213–238.

Fiorentini G, E. Sentana and G. Calzolari. 2003. Maximum likelihood estimation and inference in multivariate conditionally heteroskedastic dynamic regression models with Student t innovations. *Journal of Business & Economic Statistics* 21: 532–546.

Francq, C. and J.M. Zakoïan. 2012. QML estimation of a class of multivariate asymmetric GARCH models. *Econometric Theory* 28: 179–206.

Francq, C., L. Horvath and J.M. Zakoian. 2011. Merits and drawbacks of variance targeting in GARCH models. *Journal of Financial Econometrics* 9: 619–656.

Francq, C., L. Horvath and J.M. Zakoian. 2014. Variance targeting estimation of multivariate GARCH models. *Journal of Financial Econometrics.*

Gardebroek, C. and M.A. Hernandez. 2013. Do energy prices stimulate food price volatility? Examining volatility transmission between US oil, ethanol and corn markets. *Energy Economics* 40: 119–129.

Giacomini, R. and H. White. 2006. Tests of conditional predictive ability. *Econometrica* 74: 1545–1578.

Gourieroux, C. 1997. *ARCH Models and Financial Applications.* New York: Springer-Verlag.

Gourieroux, C. 2007. Positivity conditions for a bivariate autoregressive volatility specification. *Journal of Financial Econometrics* 5: 624–636.

Granger, C.W.J., R.P. Robins and R.F. Engle. 1986. Wholesale and retail prices: Bivariate time series modeling with forecastable error variances. In D.A. Belsley and E. Kuh (eds.), *Model Reliability*, 1–17. Cambridge: MIT Press.

Hafner C.M. 2003. Fourth moment structure of multivariate GARCH models. *Journal of Financial Econometrics* 1: 26–54.

Hafner, C.M. and H. Herwartz. 2006. Volatility impulse responses for multivariate GARCH models: an exchange rate illustration. *Journal of International Money and Finance* 25: 719–740.

Hafner, C.M. and H. Herwartz. 2008a. Analytical quasi maximum likelihood inference in multivariate volatility models. *Metrika* 67: 219–239.

Hafner, C.M. and H. Herwartz. 2008b. Testing for causality in variance using multivariate GARCH models. *Annales d'Economie et de Statistique* 89: 215–241.

Hafner, C.M. and P.H. Franses. 2009. A generalized dynamic conditional correlation model: simulation and application to many assets. *Econometric Reviews* 28: 612–631.

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

Hafner, C.M. and O. Reznikova. 2012. On the estimation of dynamic conditional correlation models. *Computational Statistics & Data Analysis* 56: 3533–3545.

Haixia, W. and L. Shiping. 2013. Volatility spillovers in China's crude oil, corn and fuel ethanol markets. *Energy Policy* 62: 878–886.

Hansen, P.R. 2005. A test for superior predictive ability. *Journal of Business and Economic Statistics* 23: 365–380.

Hansen, P.R., A. Lunde and J.M. Nason. 2011. Model confidence sets. *Econometrica* 79: 453–497.

Hansson, B. and P. Hordahl. 1998. Testing the conditional CAPM using multivariate GARCH-M. *Applied Financial Economics* 8: 377–388.

Haug, S., C. Klüppelberg, A. Lindner and M. Zapp. 2007. Method of moment estimation in the COGARCH (1, 1) model. *The Econometrics Journal* 10: 320–341.

He, C. and T. Teräsvirta. 2004. An extended constant conditional correlation GARCH model and its fourth-moment structure. *Econometric Theory* 20: 904–926.

Hendrych, R. and T. Cipra. In press. On conditional covariance modelling: An approach using state space models. *Computational Statistics & Data Analysis*.

Jeantheau, T. 1998. Strong consistency of estimators for multivariate ARCH models. *Econometric Theory* 14: 70–86.

Jeon, J. and J.W. Taylor. 2012. Using conditional kernel density estimation for wind power density forecasting. *Journal of the American Statistical Association* 107: 66–79.

Kasch, M. and M. Caporin. 2013. Volatility threshold dynamic conditional correlation: an International analysis. *Journal of Financial Econometrics* 11: 706–742.

Kavussanos, M.G. and I.D. Visvikis. 2004. Market interactions in returns and volatilities between spot and forward shipping freight markets. *Journal of Banking & Finance* 28: 2015–2049.

Kawakatsu, H. 2006. Matrix exponential GARCH. *Journal of Econometrics* 134: 95–128.

Kristensen, D. and O. Linton. 2004 Consistent standard errors for target variance approach to GARCH estimation. *Econometric Theory* 20: 990–993.

Kroner, K.F. and V.K. Ng. 1998. Modeling asymmetric comovements of asset returns. *Review of Financial Studies* 11: 817–844.

Laurent, S., J.V.K. Rombouts and F. Violante. 2012. On forecasting accuracy of multivariate GARCH models. *Journal of Applied Econometrics* 27: 934–955.

Laurent, S., J.V.K. Rombouts and F. Violante. 2013. On loss functions and ranking forecasting performances of multivariate volatility models. *Journal of Econometrics* 173: 1–10.

Ledoit, O., P. Santa-Clara and M. Wolf. 2003. Flexible multivariate GARCH modeling with an application to international stock markets. *Review of Economics and Statistics* 85: 735–747.

Lee, T.H. and X. Long. 2009. Copula-based multivariate GARCH model with uncorrelated dependent errors. *Journal of Econometrics* 150: 207–218.

Lien, D. and Y.K. Tse. 2002. Some recent developments in futures hedging. *Journal of Economic Surveys* 16: 357–396.

Ling, S. and M. McAleer. 2003. Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory* 19: 278–308.

Longin, F. and B. Solnik. 1995. Is the correlation in international equity returns constant. *Journal of International Money and Finance* 14: 3–26.

Longin, F. and B. Solnik. 2001. Extreme correlation of international equity markets. *Journal of Finance* 56: 649–676.

McAleer, M. and M.C. Medeiros. 2008. Realized volatility: A review. *Econometric Reviews* 27: 10–45.

McAleer, M., F. Chan, S. Hoti and O. Lieberman. 2008. Generalized autoregressive conditional correlation. *Econometric Theory* 24: 1554–1583.

McNeil, A.J., R. Frey and P. Embrechts. 2010. *Quantitative Risk Management: Concepts, Techniques, and Tools.* Princeton: Princeton University Press.

Nakatani, T. and T. Teräsvirta. 2008. Positivity constraints on the conditional variances in the family of conditional correlation GARCH models. *Finance Research Letters* 5: 88–95.

Nakatani, T. and T. Teräsvirta. 2009. Testing for volatility interactions in the constant conditional correlation GARCH model. *The Econometrics Journal* 12: 147–163.

Noureldin, D., N. Shephard and K. Sheppard. 2014. Multivariate Rotated ARCH Models. *Journal of Econometrics* 179: 16–30.

Okimoto, T. 2008. New evidence of asymmetric dependence structures in international equity markets. *Journal of Financial and Quantitative Analysis* 43: 787–815.

Otranto, E. 2010. Identifying financial time series with similar dynamic conditional correlation. *Computational Statistics & Data Analysis* 54: 1–15.

Pakel, C., N. Shephard, K. Sheppard and R.F. Engle. 2014. Fitting vast dimensional time-varying covariance models. *aeaweb.org*.

Palandri, A. 2009. Sequential conditional correlations: Inference and evaluation. *Journal of Econometrics* 153: 122–132.

Patton, A.J. 2006. Modelling asymmetric exchange rate dependence. *International Economic Review* 47: 527–556.

Patton, A.J. 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* 160: 246–256.

Patton, A.J. and K. Sheppard. 2009. Evaluating volatility and correlation forecasts. In T.G. Andersen, R.A. Davis, J.P. Kreiss and T. Mikosch (eds), *Handbook of Financial Time Series*, 801–838. Berlin: Springer.

Pedersen, R.S. and A. Rahbek. 2014. Multivariate variance targeting in the BEKK-GARCH model. *The Econometrics Journal* 17: 24–55.

Rangel, J.G. and R.F. Engle. 2012. The factor-spline-GARCH model for high and low frequency correlations. *Journal of Business & Economic Statistics* 30: 109-124.

Ramchand, L. and R. Susmel. 1998. Volatility and cross correlation across major stock markets. *Journal of Empirical Finance* 5: 397–416.

Rombouts, J., L. Stentoft and F. Violante. 2014. The value of multivariate model sophistication: An application to pricing Dow Jones Industrial Average options. *International Journal of Forecasting* 30: 78–98.

Rossi, E. and F. Spazzini. 2010. Model and distribution uncertainty in multivariate GARCH estimation: a Monte Carlo analysis. *Computational Statistics & Data Analysis* 54: 2786–2800.

Santos, A.A.P., E. Ruiz and F.J. Nogales. 2013. Comparing univariate and multivariate models to forecast portfolio value-at-risk. *Journal of Financial Econometrics* 11: 400–441.

Santos, A.A.P., F.J. Nogales, E. Ruiz, and D.V. Dijk. 2012. Optimal portfolios with minimum capital requirements. *Journal of Banking & Finance* 36: 1928–1942.

Scherrer, W. and E. Ribarits. 2007. On the parametrization of multivariate GARCH models. *Econometric Theory* 23: 464–484.

Serban, M., A. Brockwell, J. Lehoczky and S. Srivastava. 2007. Modelling the dynamic dependence structure in multivariate financial time series. *Journal of Time Series Analysis* 28: 763–782.

Silvennoinen, A. and T. Teräsvirta. 2009a. Multivariate GARCH models. In T.G. Andersen, R.A. Davis, J-P. Kreiss and T. Mikosch (eds.), *Handbook of Financial Times Series*, 201–232. New York: Springer.

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.

Silvennoinen, A. and T. Teräsvirta. 2015. Modeling conditional correlations of asset returns: A smooth transition approach. *Econometric Reviews* 34: 174–197.

So, M.K. and C.Y. Yeung. 2014. Vine-copula GARCH model with dynamic conditional dependence. *Computational Statistics & Data Analysis* 76: 655–671.

Stelzer, R. 2008. On the relation between the VEC and BEKK multivariate GARCH models. *Econometric Theory* 24: 1131–1136.

Tsay, R.S. 1987. Conditional heteroscedastic time series models. *Journal of the American Statistical Association* 82: 590–604.

Tsay, R.S. 2002. *Analysis of Financial Time Series*, 357–394. New York: John Wiley.

Tse, Y.K. 2000. A test for constant correlations in a multivariate GARCH model. *Journal of Econometrics* 98: 107–127.

Tse, Y.K. and A.K.C. Tsui. 2002. A multivariate GARCH model with time-varying correlations. *Journal of Business & Economic Statistics* 20: 351–362.

van der Weide, R. 2002. GO-GARCH: a multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics* 17: 549–564.

Notes

[1]For the sake of simplicity, in this paper, we focus on the simplest specification which only includes one lag of past returns, conditional variances and covariances.

[2]We refer to the observed vector of time series as vector of returns. Nevertheless, it should be understood that we may also refer to the residuals of a multivariate model fitted to represent the conditional means when dealing with non-financial data.

[3]All the programming used in this paper has been developed by the first author using MATLAB codes.

[4]Note that, as the DGP has Gaussian errors, the QML is in fact ML.

[5]Kawakatsu (2006) proposes an alternative model based on $\exp\{\mathbf{H}_t\}$ that also guarantees positivity. Other models that guarantee the symmetry and positive definiteness of the conditional covariance matrices are given in Tsay (2002) and Bai and Chen (2008), who consider a Cholesky decomposition of $\mathbf{H}_t$, and Hendrych and Cipra (in press), who propose a formulation based on the LDL decomposition.

[6]Hereafter, the BEKK model refers to the BEKK(1,1,1) model.

[7]See Engle *et al.* (1990), Alexander and Chibumba (1997) and van der Weide (2002) for early references and Hafner and Preminger (2009) for the asymptotic theory of factor MGARCH models. Haug *et al.* (2007) propose a method of moments estimator for a COGARCH model.

[8]See Engle and Sheppard (2008) for methods that ensure that the conditional variance, covariance and correlation estimates are not biased due to the use of parameter estimates located at local rather than global optima.

[9]We also compute the realized covariances using ten-minute and thirty-minute intervals with similar results.

[10]The proxies are estimated in R by the package *highfrequency*.

[11]We compute the SPA test and MSC using the Sheppard's MFE Toolbox package, written in MATLAB, by considering 10,000 bootstrap replications and block length of 6. The results do not change when we change the block length to 3 or 9. The MSC is constructed using the R method within the MFE Toolbox package.

Table 1: Summary of the four Monte Carlo simulations: when the DGP is the Gaussian VECH and Student-7 VECH models with $T_1 = 1000$ and $T_2 = 2000$ and considering 500 replications. In the second and third columns are the number of cases where the fitted model is not stationary and not positive, respectively; in the fourth column, the mean time in seconds required to estimate the model. *estimation* means that the stationarity (or positivity) is ensured by the estimation procedure; *model* means that the positivity is ensured by the model parametrization.

		Stationarity				Positivity				Computer time			
Model	Estimator	Gaussian		Student		Gaussian		Student		Gaussian		Student	
		T_1	T_2	T_1	T_2	T_1	T_2	T_1	T_2	T_1	T_2	T_1	T_2
VECH	G-QML	15	5	65	30	2	0	5	0	28.8	57.2	31.4	58.4
	S-QML	19	6	50	18	2	0	6	0	25.6	49.1	55.2	104.7
	VT	0	0	0	0	41	31	85	67	20.5	40.7	22.7	41.8
DVECH	G-QML	0	0	21	3	0	0	0	0	9.4	20.3	11.8	21.6
	S-QML	0	0	13	2	0	0	0	0	20.8	41.8	25.2	48.3
	VT	0	0	0	0	0	0	0	0	7.8	17.2	7.5	18.4
	LSW	estimation				estimation				0.27	0.43	0.26	0.39
BEKK	G-QML	15	9	96	57	model				17.2	34.7	17.0	30.5
	S-QML	13	8	60	37	model				24.3	51.5	28.3	54.5
	VT	0	0	0	0	19	12	47	32	11.5	22.7	13.2	23.5
DBEKK	G-QML	estimation				model				3.7	7.7	4.3	8.4
	S-QML	estimation				model				6.2	12.5	9.5	17.2
	VT	0	0	0	0	0	0	0	0	3.6	4.8	3.0	5.3
SBEKK	G-QML	estimation				model				2.7	6.2	2.7	5.8
	S-QML	estimation				model				3.4	7.3	4.7	8.9
	VT	0	0	0	0	0	0	0	0	1.3	2.7	1.6	3.1
D-RBEKK	VT	estimation				estimation				1.6	3.1	1.6	3.0
	S-VT	estimation				estimation				3.2	6.6	4.2	7.8
S-RBEKK	VT	estimation				estimation				0.80	1.6	0.82	1.5
	S-VT	estimation				estimation				1.6	3.3	1.7	3.1
O-GARCH	G-QML	estimation				model				1.3	2.3	1.3	2.4
GO-GARGH	G-QML	estimation				model				3.1	5.9	3.1	5.9
CCC	G-2s	estimation				estimation				0.23	0.33	0.23	0.34
	S-2s	estimation				estimation				1.1	1.3	0.74	1.3
ECCC	G-2s	12	5	62	33	4	0	15	1	11.9	23.3	12.9	22.3
	S-2s	20	7	115	76	6	0	14	1	23.3	45.3	22.5	39.9
	VT	0	0	1	0	25	16	59	37	13.5	24.7	13.4	20.0
cDCC	G-3s	estimation				estimation				1.1	1.7	1.2	1.8
	S-3s	estimation				estimation				3.2	4.9	3.0	5.3
RDCC	G-3s	estimation				estimation				1.2	2.0	1.3	2.1
	S-3s	estimation				estimation				3.9	5.7	3.7	6.4

Table 2: Medians through the Monte Carlo replicates of the LF statistics for the conditional covariances (1) and correlations (2) matrices after fitting alternative MGARCH models by different estimation procedures when the DGP is the Gaussian VECH and Student-7 VECH models, considering sample sizes $T = 1000$ and 2000 and 500 replications. The LF_1 and LF_2 statistics are in the scale $\times 10^{-9}$ and $\times 10^{-2}$, respectively.

		Gaussian				Student			
Model	Estimator	$T = 1000$		$T = 2000$		$T = 1000$		$T = 2000$	
		LF_1	LF_2	LF_1	LF_2	LF_1	LF_2	LF_1	LF_2
VECH	G-QML	1.59	0.62	1.10	0.28	2.60	0.61	1.59	0.29
	S-QML	1.60	0.62	1.10	0.29	2.24	0.54	1.50	0.26
	VT	1.46	0.71	0.96	0.33	2.26	0.72	1.41	0.35
DVECH	G-QML	3.05	0.87	3.49	0.75	4.33	0.89	4.65	0.75
	S-QML	3.06	0.86	3.49	0.75	4.16	0.79	4.51	0.67
	VT	3.53	1.06	3.44	0.93	4.60	1.13	4.37	0.96
	LSW	2.61	1.42	3.00	1.45	4.12	1.50	4.38	1.39
BEKK	G-QML	4.41	2.99	4.53	3.02	6.52	3.49	6.08	3.22
	S-QML	4.31	2.94	4.51	2.95	6.66	3.12	6.10	2.74
	VT	4.38	3.11	4.52	3.02	6.39	3.56	5.92	3.22
DBEKK	G-QML	2.96	2.21	3.42	2.46	4.34	2.35	4.42	2.37
	S-QML	2.97	2.13	3.40	2.35	4.14	2.04	4.17	2.09
	VT	2.86	2.28	3.27	2.51	3.89	2.39	3.87	2.53
SBEKK	G-QML	2.43	2.49	2.97	2.64	3.47	2.51	3.88	2.51
	S-QML	2.45	2.39	2.93	2.54	3.26	2.20	3.77	2.22
	VT	2.46	2.67	2.84	2.80	3.08	2.81	3.59	2.86
D-RBEKK	VT	2.64	2.49	3.01	2.71	3.61	2.72	3.76	2.82
	S-VT	2.64	2.48	2.98	2.69	3.42	2.63	3.73	2.73
S-RBEKK	VT	2.46	2.67	2.84	2.80	3.08	2.81	3.59	2.86
	S-VT	2.47	2.65	2.84	2.77	3.08	2.73	3.64	2.79
O-GARCH	G-QML	3.82	3.09	4.95	3.17	5.58	3.13	6.57	3.15
GO-GARCH	G-QML	4.25	2.14	5.59	2.14	5.51	2.69	6.62	2.79
CCC	G-2s	2.30	1.72	2.73	1.73	3.52	1.88	3.84	1.91
	S-2s	2.30	1.73	2.71	1.73	3.36	1.89	3.82	1.91
ECCC	G-2s	2.19	1.76	2.14	1.74	3.98	1.94	3.29	1.92
	S-2s	2.24	1.76	2.15	1.74	3.73	1.93	3.40	1.92
	VT	1.92	1.76	1.83	1.75	3.04	1.96	2.73	1.96
cDCC	G-3s	1.82	0.97	2.08	0.86	3.10	1.20	3.14	1.13
	S-3s	1.83	0.96	2.12	0.85	2.93	1.19	2.95	1.10
RDCC	G-3s	1.83	0.97	2.08	0.84	3.10	1.17	3.14	1.11
	S-3s	1.82	0.96	2.10	0.84	2.91	1.16	2.97	1.08

Table 3: Descriptive Statistics of daily returns of the EUR, GBP, CHF, AUD and JPY against the USD: mean, minimum (Min), maximum (Max), variance (Var), skewness (Skew), excess of kurtosis (Kurt) and Ljung-Box for returns ($Q(20)$) and squared returns ($Q_2(20)$). The top panel corresponds to the full sample period; the second panel to the first in-sample period; the third panel to the first out-of-sample period; the fourth panel to the second in-sample period; the fiftieth panel to the second out-of-sample period; the sixth panel to the third in-sample period; and the bottom panel to the third out-of-sample period. *, ** and *** mean significant at 10%, 5% and 1% levels for the skewness, excess of kurtosis and Ljung-Box asymptotic tests.

Currency	Mean	Min	Max	Var	Skew	Kurt	$Q(20)$	$Q_2(20)$
Full sample - from January 2, 2004 to December 31, 2013								
EUR	0.0034	-3.84	4.62	0.41	0.123*	3.054***	16.9	39.2***
GBP	-0.0030	-3.92	4.47	0.40	-0.079	4.181***	77.0***	1359.6***
CHF	0.0128	-8.48	5.45	0.50	-0.514***	11.11***	36.5**	45.3***
AUD	0.0067	-8.83	6.70	0.86	-0.950***	12.35***	46.6***	2610.5***
JPY	0.0004	-5.17	3.81	0.46	0.072	4.270***	42.7***	738.1***
First in-sample - from January 2, 2004 to December 31, 2012, with $T = 2323$								
EUR	0.0019	-3.84	4.62	0.43	0.139*	2.987***	19.8	37.5**
GBP	-0.0042	-3.92	4.47	0.42	-0.072	4.103***	75.6***	1194.3***
CHF	0.0130	-8.48	5.45	0.52	-0.541***	11.28***	34.3**	46.4***
AUD	0.0138	-8.83	6.70	0.91	-0.955***	12.07***	49.5***	2353.6***
JPY	0.0089	-5.17	3.81	0.44	0.098	4.698***	54.3***	778.4***
First out-of-sample - from January 2, 2013 to December 31, 2013, with $H = 258$								
EUR	0.0171	-1.72	1.45	0.23	-0.187	0.963**	40.7***	26.9
GBP	0.0073	-1.27	1.36	0.21	-0.180	0.301	17.6	19.7
CHF	0.0112	-1.74	1.73	0.31	0.097	0.609*	33.1**	12.2
AUD	-0.0577	-3.77	1.52	0.40	-0.802***	4.118***	24.6	19.4
JPY	-0.0756	-3.42	2.81	0.59	-0.008	1.717***	13.3	15.5
Second in-sample - from January 2, 2004 to December 31, 2007, with $T = 1031$								
EUR	0.014	-1.738	1.959	0.274	0.010	0.754***	14.4	21.0
GBP	0.010	-2.014	2.010	0.277	-0.069	0.479***	16.8	101.7***
CHF	0.009	-2.245	1.945	0.354	0.143	0.597***	14.5	20.5
AUD	0.015	-5.104	2.440	0.487	-0.924***	4.822***	29.9*	72.3***
JPY	-0.004	-2.193	2.223	0.316	0.290***	1.115***	27.5	74.5***
Second out-of-sample - from January 2, 2008 to December 31, 2013, with $H = 1550$								
EUR	-0.004	-3.844	4.617	0.497	0.164*	3.034***	15.8	20.4
GBP	-0.012	-3.918	4.474	0.482	-0.063	4.468***	83.9***	808.8***
CHF	0.016	-8.479	5.453	0.601	-0.701***	12.589***	42.7***	31.8**
AUD	0.001	-8.828	6.701	1.103	-0.894***	11.402***	41.5***	1641.5***
JPY	0.004	-5.168	3.813	0.555	0.001	4.412***	47.4***	429.9***
Third in-sample - from January 2, 2004 to December 31, 2008, with $T = 1290$								
EUR	0.008	-3.844	4.027	0.374	-0.042	4.078***	34.6**	25.5
GBP	-0.017	-3.728	4.474	0.393	-0.185*	4.586***	47.9***	880.9***
CHF	0.012	-4.434	5.453	0.463	0.281***	4.878***	9.3	29.7*
AUD	-0.006	-8.828	6.701	0.985	-1.489***	17.171***	73.9***	1463.6***
JPY	0.013	-5.168	3.813	0.473	0.223**	5.270***	60.8***	743.3***
Third out-of-sample - from January 2, 2009 to December 31, 2013, with $H = 1550$								
EUR	-0.001	-2.180	4.617	0.442	0.254**	5.257***	14.2	18.6
GBP	0.011	-3.918	4.325	0.407	0.018	6.792***	54.1***	437.0***
CHF	0.014	-8.479	4.227	0.543	-1.137***	18.500***	52.9***	22.1
AUD	0.019	-4.108	4.974	0.728	-0.067	5.457***	30.2*	364.1***
JPY	-0.012	-3.423	3.610	0.446	-0.096	6.111***	25.4	56.3***

Table 4: Results based on the out-of-sample negative log-likelihood (NL) and Frobenius loss function (LF) for $h = 1$, 5 and 20 horizon forecasting of the system of the Euro, British Pound, Swiss Franc, Australian Dollar and Japanese Yen currency returns against the US Dollar currency. We evaluate the realized covariance using a five-minute frequency data and rOWCov method. In the top left panel is the NL values; in the top right panel is the LF_1; and in the bottom panel is the LF_2. In all the panel are the p-value for the superior predictive ability (SPA) tests for the null hypothesis that each model is not outperformed by any of the other models according with the corresponding criterion. In bold, the models included in the confidence sets (MCS) at 10% significance level. The out-of-sample periods goes from January 2, 2012 to December 31, 2013, for a total of 258 daily observations.

	NL						LF_1						LF_2					
Model-Estimator	1-step-ahead		5-step-ahead		20-step-ahead		1-step ahead		5-step ahead		20-step ahead		1-step ahead		5-step ahead		20-step ahead	
	NL	SPA	NL	SPA	NL	SPA	LF_1	SPA	LF_1	SPA	LF_1	SPA	LF_2	SPA	LF_2	SPA	LF_2	SPA
DVEC-Flex	869.9	0.00	868.5	0.00	848.0	0.00	0.319	0.00	0.372	0.00	0.407	0.00	0.485	0.00	0.499	0.00	0.502	0.00
DBEKK-VT	814.5	0.02	834.4	0.06	790.4	0.00	0.304	0.00	0.369	0.05	0.400	0.00	0.456	0.00	0.452	0.00	0.482	0.03
SBEKK-G-QML	804.9	0.02	**823.1**	0.06	789.2	0.01	0.295	0.03	0.366	0.04	0.396	0.00	0.461	0.00	0.445	0.00	0.499	0.00
-S-QML	**771.7**	1.00	**773.9**	1.00	**744.3**	0.58	0.304	0.00	0.367	0.00	0.396	0.00	0.470	0.00	0.435	0.00	0.489	0.04
-VT	815.2	0.00	820.8	0.00	792.9	0.00	0.299	0.00	0.381	0.00	0.405	0.00	0.449	0.00	0.451	0.00	0.495	0.00
D-RBEKK-VT	818.6	0.01	**840.5**	0.05	797.2	0.00	0.295	0.12	**0.364**	0.21	0.392	0.02	0.426	0.00	0.439	0.00	0.493	0.00
S-VT	**779.0**	0.14	**781.4**	0.25	**752.1**	0.20	**0.285**	1.00	**0.352**	1.00	0.387	0.09	0.435	0.00	0.430	0.02	0.494	0.00
S-RBEKK-VT	815.2	0.00	820.8	0.00	792.9	0.00	0.301	0.00	0.370	0.01	0.399	0.00	0.444	0.00	0.441	0.00	0.495	0.00
S-VT	**777.9**	0.29	**774.1**	0.92	747.6	0.13	0.299	0.01	0.370	0.02	0.398	0.00	0.441	0.00	0.449	0.00	0.492	0.00
CCC-G-2s	876.7	0.00	861.0	0.00	818.5	0.00	0.316	0.00	0.413	0.00	0.409	0.00	0.492	0.00	0.511	0.00	0.512	0.00
S-2s	811.9	0.03	**799.2**	0.07	755.1	0.03	0.318	0.00	0.421	0.00	0.412	0.00	0.493	0.00	0.504	0.00	0.516	0.00
cDCC-G-3s	819.4	0.01	828.9	0.02	789.0	0.00	0.304	0.00	0.372	0.03	0.392	0.00	0.410	0.05	**0.423**	0.08	**0.481**	0.11
S-3s	**779.2**	0.21	**776.6**	0.49	**742.8**	0.17	0.297	0.02	**0.367**	0.15	0.398	0.00	0.415	0.01	**0.412**	0.36	**0.470**	0.68
RDCC-G-3s	819.2	0.01	828.0	0.03	784.3	0.00	**0.288**	0.65	**0.363**	0.22	0.392	0.02	**0.402**	0.12	**0.416**	0.29	**0.478**	0.52
S-3s	**778.4**	0.25	**774.9**	0.89	**739.0**	1.00	**0.286**	0.89	**0.354**	0.91	**0.381**	1.00	**0.376**	1.00	**0.408**	1.00	**0.469**	1.00

Table 5: Results based on the out-of-sample negative log-likelihood (NL) for $h = 1$, 5 and 20 horizon forecasting of the system of the Euro, British Pound, Swiss Franc, Australian Dollar and Japanese Yen currency returns against the US Dollar currency. On the right of each observed NL is the corresponding p-value for the superior predictive ability (SPA) tests for the null hypothesis that each model is not outperformed by any of the other models. The left panel correspond to the second split and the right panel to the third split. In the bold, the models included in the model confidence set (MCS) at $\alpha = 10\%$.

	Second split						Third split					
Model-	1-step-ahead		5-step-ahead		20-step-ahead		1-step-ahead		5-step-ahead		20-step-ahead	
Estimator	NL	SPA	NL	SPA	NL	SPA	NL	SPA	NL	SPA	NL	SPA
DVEC-Flex	6948.1	0.00	7188.6	0.00	7755.3	0.00	5495.2	0.00	5609.2	0.00	6137.1	0.00
DBEKK-VT	6316.9	0.00	6609.8	0.00	7220.6	0.00	4840.1	0.00	4995.6	0.00	5303.3	0.00
SBEKK-G-QML	6268.3	0.00	6587.8	0.00	7143.2	0.00	4805.6	0.00	5008.1	0.00	5245.5	0.00
-S-QML	**5891.5**	1.00	**6019.5**	1.00	**6242.3**	1.00	**4514.6**	1.00	**4585.8**	1.00	**4849.5**	1.00
-VT	6366.2	0.00	6678.6	0.00	7080.1	0.00	4849.2	0.00	4942.7	0.00	5334.2	0.00
D-RBEKK-VT	6284.4	0.00	6543.0	0.00	7022.2	0.00	4816.6	0.00	4940.7	0.00	5311.8	0.00
S-VT	5924.2	0.02	**6033.6**	0.28	**6242.9**	0.54	4563.2	0.00	4655.3	0.02	4934.8	0.01
S-RBEKK-VT	6352.7	0.00	6633.2	0.00	7052.4	0.00	4887.2	0.00	4953.7	0.00	5334.2	0.00
S-VT	5957.2	0.00	6071.9	0.05	6320.1	0.03	4590.7	0.00	4649.9	0.02	4955.8	0.00
CCC-G-2s	6932.1	0.00	7101.2	0.00	7564.7	0.00	5389.6	0.00	5445.3	0.00	5947.0	0.00
S-2s	6473.0	0.00	6526.6	0.00	6661.8	0.00	5049.5	0.00	5049.9	0.00	5365.3	0.00
cDCC-G-3s	6307.6	0.00	6546.3	0.00	7088.9	0.00	4873.0	0.00	4997.3	0.00	5434.5	0.00
S-3s	5995.8	0.00	6105.6	0.01	6336.9	0.03	4628.6	0.00	4683.0	0.00	5026.2	0.00
RDCC-G-3s	6294.5	0.00	6529.7	0.00	7035.4	0.00	4862.9	0.00	4990.5	0.00	5428.4	0.00
S-3s	5987.4	0.00	6091.3	0.01	6307.4	0.02	4619.1	0.00	4669.9	0.00	5019.7	0.00

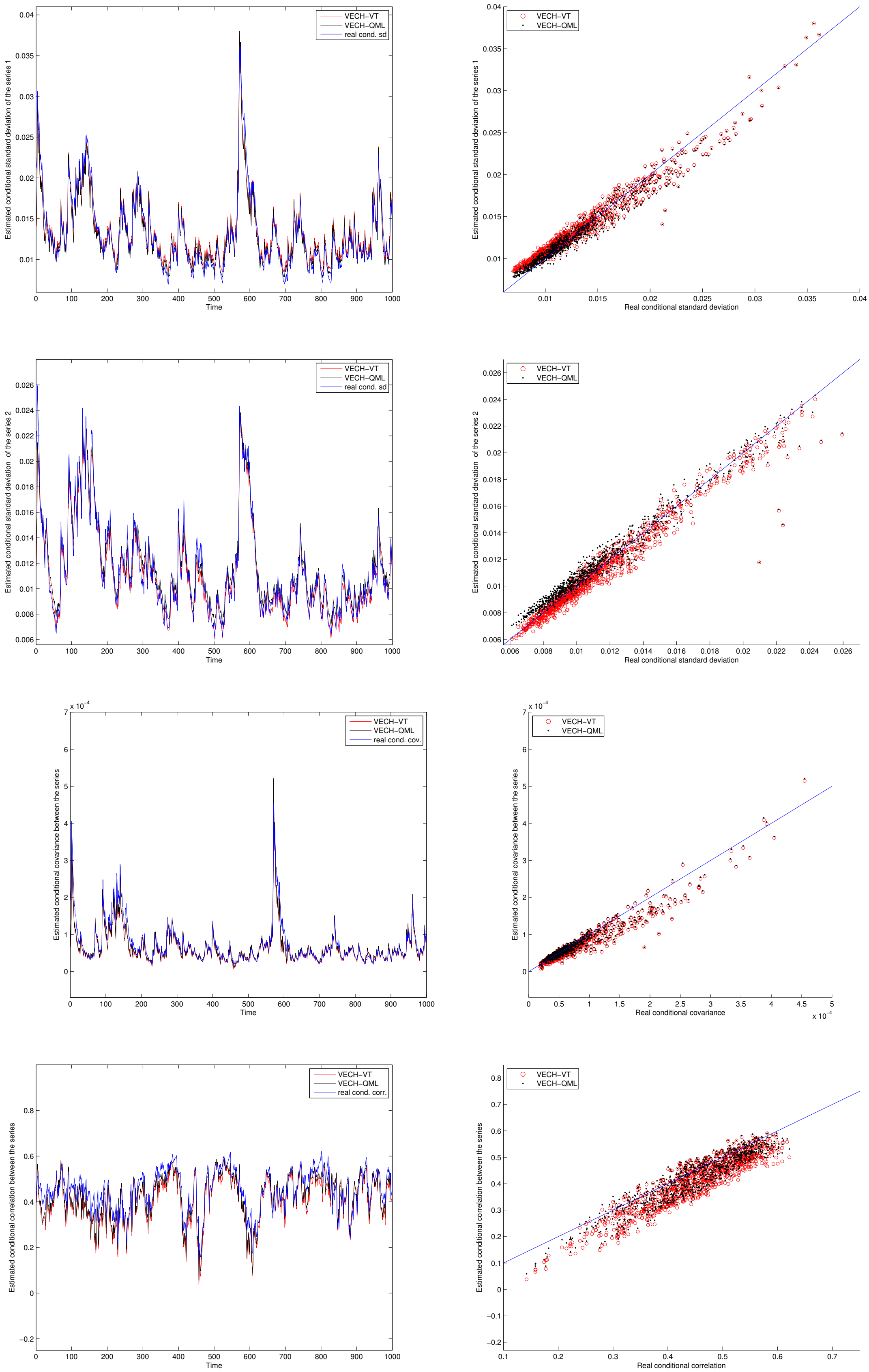

Figure 1: First column: Simulated conditional standard deviations (first two rows), covariances (third row) and correlations (fourth row) of the VECH series together with the corresponding estimates obtained after fitting the VECH model by QML and VT. Second column: corresponding simulated vs fitted values.

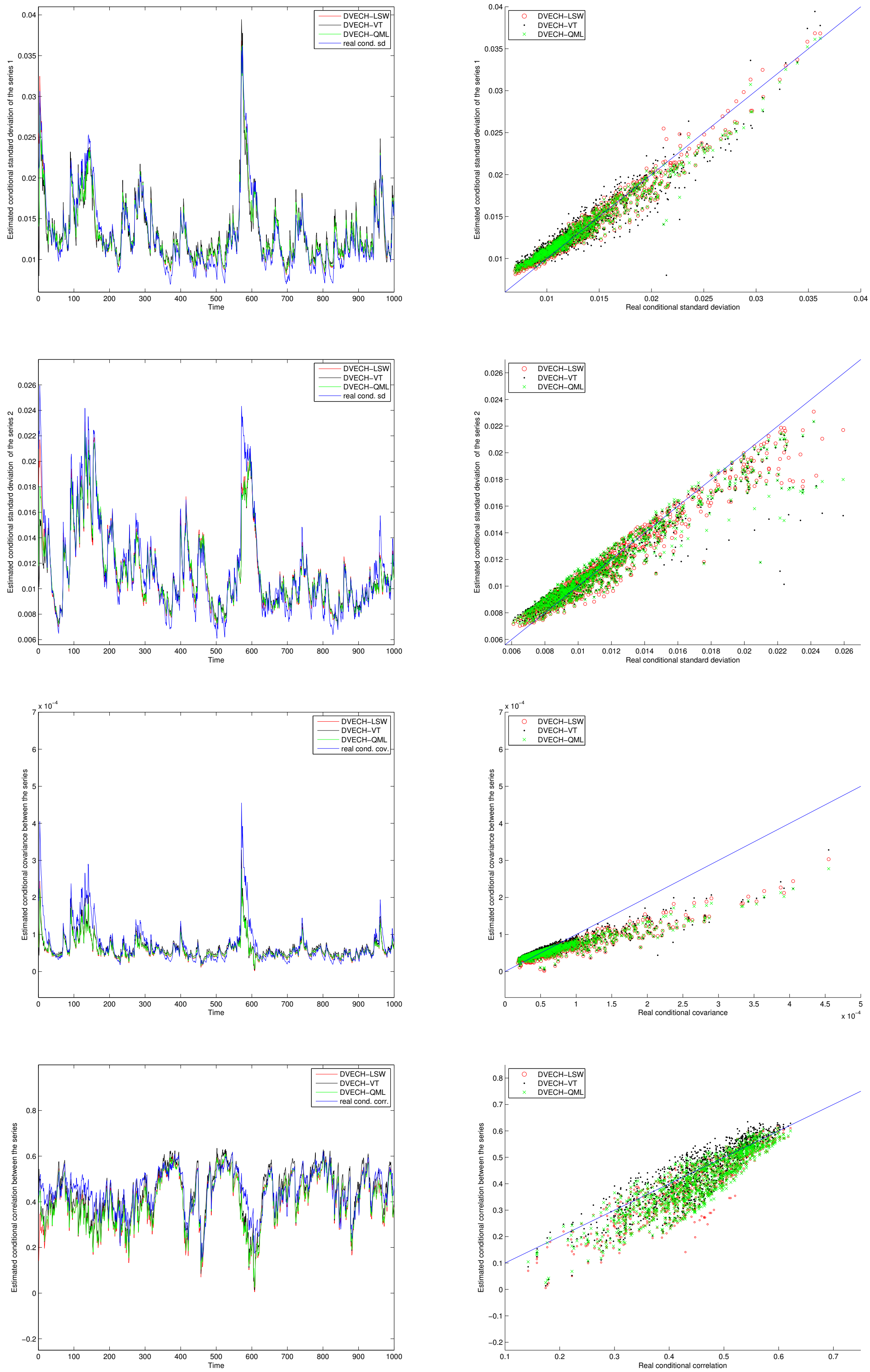

Figure 2: First column: Simulated conditional standard deviations (first two rows), covariances (third row) and correlations (fourth row) of the VECH series together with the corresponding estimates obtained after fitting the DVECH model by QML, VT and LSW. Second column: corresponding simulated vs fitted values.

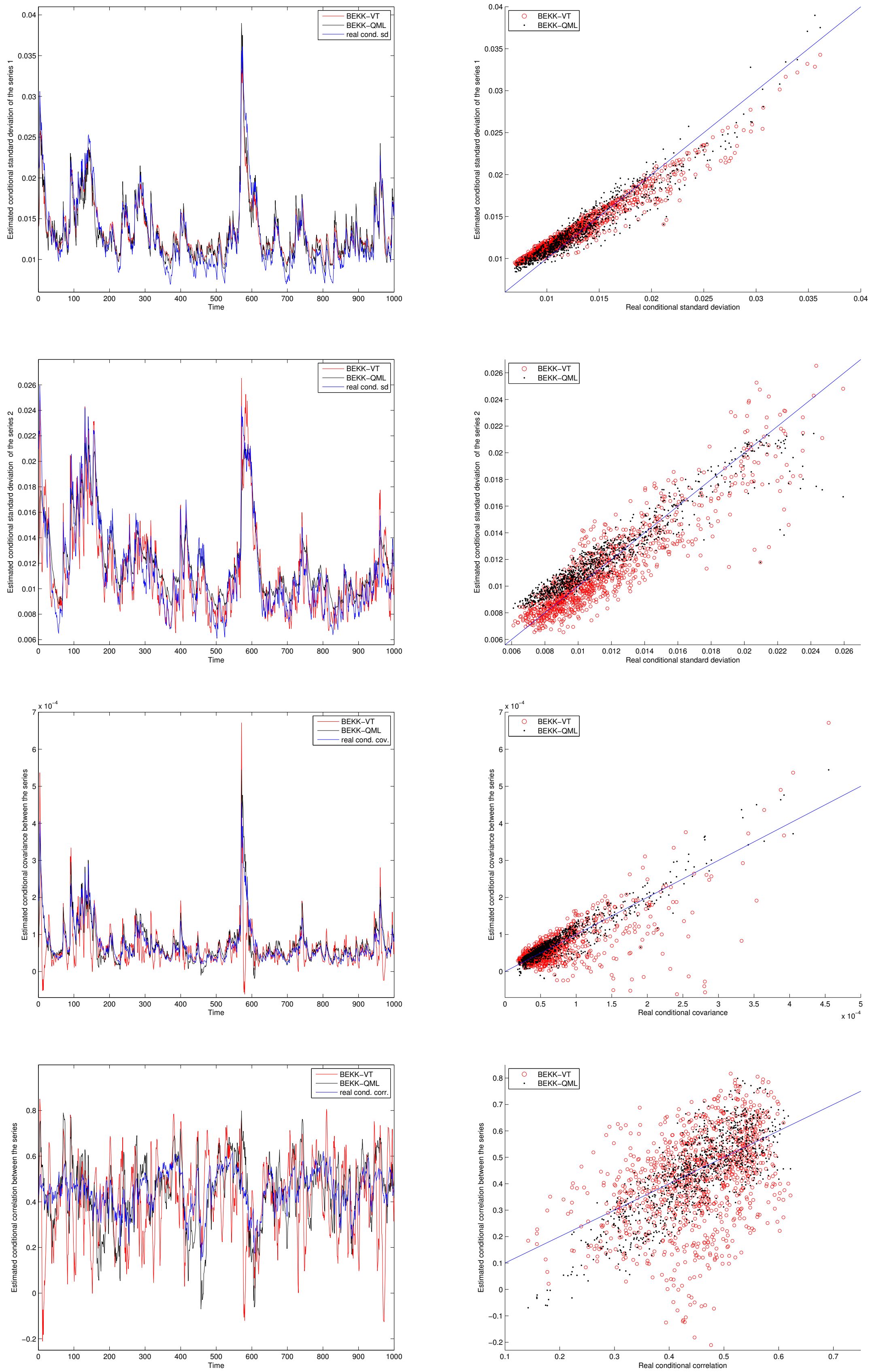

Figure 3: First column: Simulated conditional standard deviations (first two rows), covariances (third row) and correlations (fourth row) of the VECH series together with the corresponding estimates obtained after fitting the BEKK by QML and VT. Second column: corresponding simulated vs fitted values.

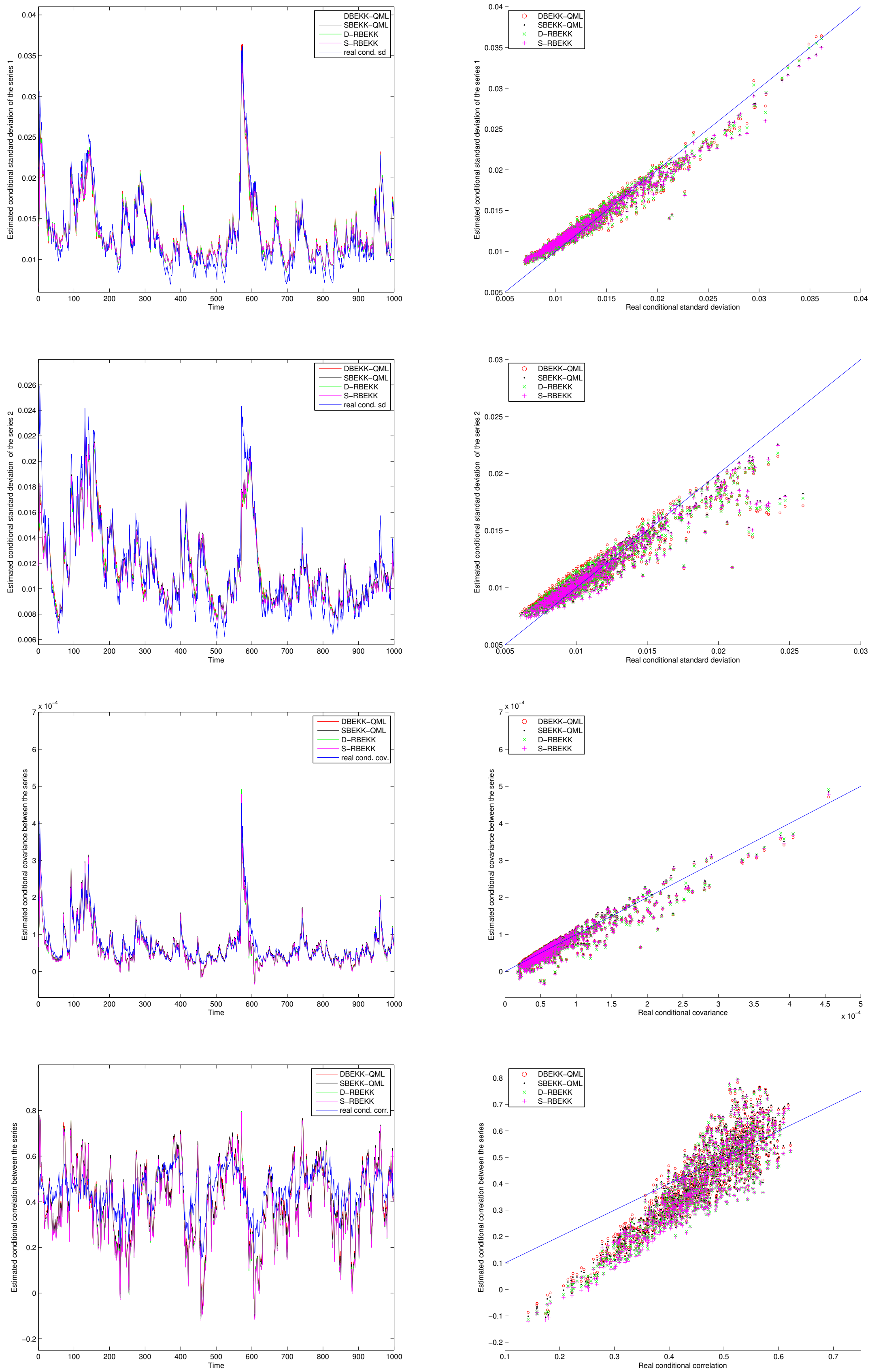

Figure 4: First column: Simulated conditional standard deviations (first two rows), covariances (third row) and correlations (fourth row) of the VECH series together with the corresponding estimates obtained after fitting the DBEKK, SBEKK and D-RBEKK and S-RBEKK models. Second column: corresponding simulated vs fitted values.

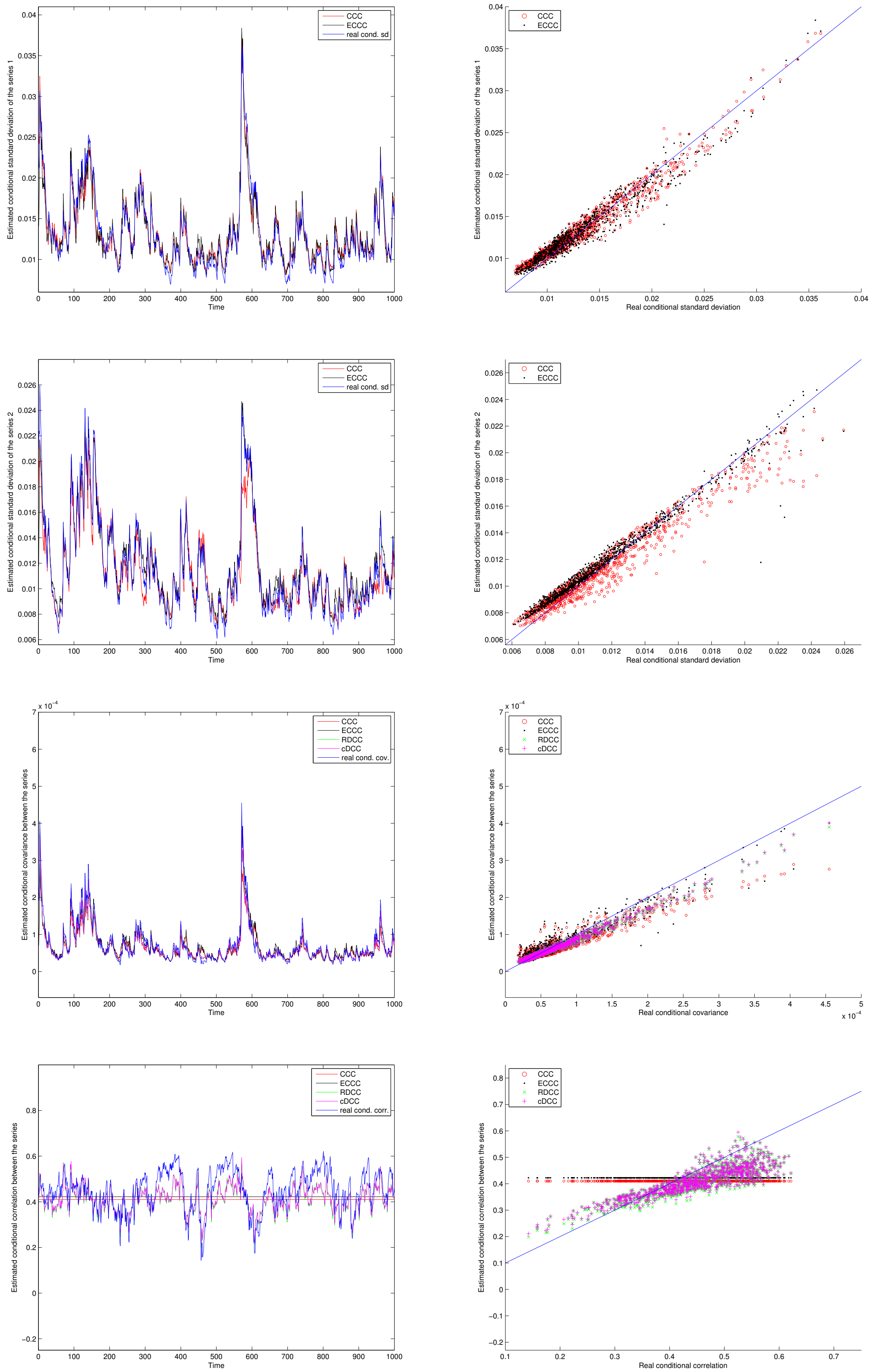

Figure 5: First column: Simulated conditional standard deviations (first two rows), covariances (third row) and correlations (fourth row) of the VECH series together with the corresponding estimates obtained after fitting the CCC, ECCC, cDCC and RDCC models. Second column: corresponding simulated vs fitted values.

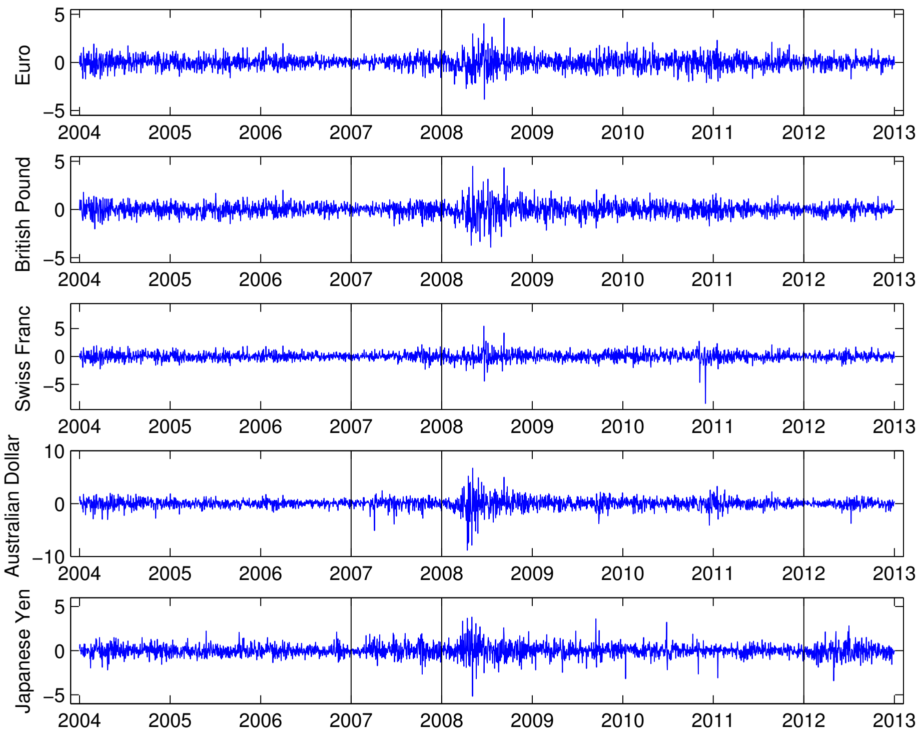

Figure 6: Returns series of the Euro (first row), British Pound (second row), Swiss Franc (third row), Australian Dollar (fourth row) and Japanese Yen currencies (fifth row) against the US Dollar currency in the period starting on January 1, 2004 and ending on December 31, 2013, for a total of 2581 daily observations.

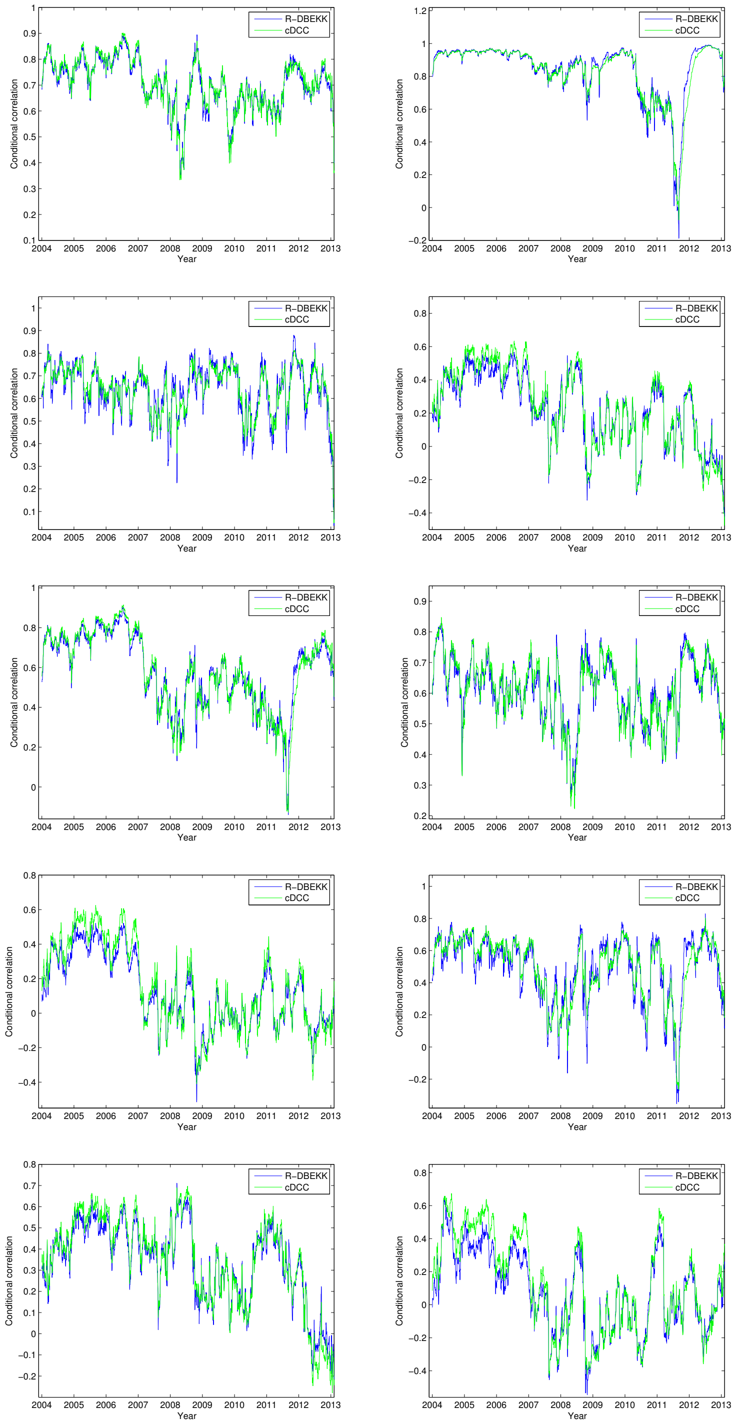

Figure 7: Pairwise conditional correlations estimated by the R-DBEKK and cDCC models by maximizing the Gaussian likelihood in the period starting on January 2, 2004 and ending on December 31, 2013. In the first row on the left (right), correlations between EUR vs GBP(CHF) returns; in the second, EUR vs AUD(JPY); in the third, GBP vs CHF(AUD); in the fourth, GBP vs JPY(CHF vs AUD); in the fifth, CHF vs JPY(AUD vs JPY).