SSOAR

Open Access Repository
www.ssoar.info

Sequential conditional correlations: Inference and evaluation

Palandri, Alessandro

Postprint / Postprint
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
www.peerproject.eu

Empfohlene Zitierung / Suggested Citation:
Palandri, A. (2009). Sequential conditional correlations: Inference and evaluation. *Journal of Econometrics*, *153*(2), 122-132. https://doi.org/10.1016/j.jeconom.2009.05.002

Nutzungsbedingungen:
Dieser Text wird unter dem "PEER Licence Agreement zur Verfügung" gestellt. Nähere Auskünfte zum PEER-Projekt finden Sie hier: http://www.peerproject.eu Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:
This document is made available under the "PEER Licence Agreement ". For more Information regarding the PEER-project see: http://www.peerproject.eu This document is solely intended for your personal, non-commercial use.All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.
By using this particular document, you accept the above-stated conditions of use.

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

Diese Version ist zitierbar unter / This version is citable under:
https://nbn-resolving.org/urn:nbn:de:0168-ssoar-251154

Accepted Manuscript

Sequential conditional correlations: Inference and evaluation

Alessandro Palandri

PII: S0304-4076(09)00127-4
DOI: 10.1016/j.jeconom.2009.05.002
Reference: ECONOM 3204

To appear in: *Journal of Econometrics*

Received date: 26 April 2006
Revised date: 27 April 2009
Accepted date: 5 May 2009

Please cite this article as: Palandri, A., Sequential conditional correlations: Inference and evaluation. *Journal of Econometrics* (2009), doi:10.1016/j.jeconom.2009.05.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Sequential Conditional Correlations: Inference and Evaluation

Alessandro Palandri*
University of Copenhagen and CREATES†

April 2009

Abstract

This paper presents a new approach to the modeling of conditional correlation matrices within the multivariate GARCH framework. The procedure, which consists in breaking the matrix into the product of a sequence of matrices with desirable characteristics, in effect converts a highly dimensional and intractable optimization problem into a series of simple and feasible estimations. This in turn allows for richer parameterizations and complex functional forms for the single components. An empirical application involving the conditional second moments of 69 selected stocks from the NASDAQ100 shows how the new procedure results in strikingly accurate measures of the conditional correlations.

Keywords: Multivariate GARCH, High Dimensional GARCH Models, Conditional Correlations, Sequential Estimation.

JEL classification: C51, C52, C61, G1.

*Mailing address: Department of Economics, University of Copenhagen, Studiestræde 6, DK-1455 København K, Danmark, Email: alessandro.palandri@econ.ku.dk

†CREATES, Center for Research in Econometric Analysis of Time Series, is funded by the Danish National Research Foundation.

ACCEPTED MANUSCRIPT

1 Introduction

Modeling the temporal dependence in the second order moments and forecasting future volatility have key relevance in many financial-econometric issues such as portfolio management and selection, risk analysis and hedging and the pricing of assets and derivatives. As a consequence, many multivariate GARCH (MGARCH) models have been developed in the recent years to model the conditional second moments. However, all of them must make the trade-off between parameters' parsimony and richness in the description of the second order dynamics. In fact, the number of parameters of a fairly rich multivariate volatility model soon becomes large enough to render estimation infeasible.

In this paper I introduce a methodology that allows to handle large cross-sectional dimensions without imposing unnecessary restrictions: the Sequential Conditional Correlations (SCC), a new and feasible approach to the modeling of correlations in the MGARCH framework. The key feature of the SCC is the decomposition of the conditional correlation matrix into the product of a sequence of matrices with desirable characteristics. In particular, it is possible to separately model the conditional correlations and partial correlations following an internally consistent procedure that automatically delivers a positive definite correlation matrix. This allows for a multi step estimation procedure, thus converting a highly dimensional and intractable optimization problem into a series of simple and feasible estimations. Furthermore, SCC allows for a wide range of parameterizations for the conditional variances[1] and the pairwise conditional correlations and partial correlations[2].

The fundamental issue for any MGARCH model [3] is how to guarantee positive definiteness of the conditional variance-covariance matrix. Current solutions, however, have led to models that, for non trivial cross-sectional dimensions M, cannot be feasibly estimated. In particular, for MGARCH specifications such as the VEC of Bollerslev, Engle and Wooldridge (1988) and the VEC-like models of Gallant and Tauchen (2002) and Kawakatsu (2006), feasible estimation can be an issue, even when the number of series M is relatively small, as the number of parameters grows at the rate M^4. In order to mitigate this problem, models like the BEKK of Engle and Kroner (1995) and the Factor-GARCH of Diebold and Nerlove (1989) and Engle,

[1] These include but are not restricted to GARCH models, Stochastic Volatility, Nonparametric models and Realized Volatility.

[2] The econometrician can employ any specification satisfying the constraint that the model-generated correlations be within the $(-1, 1)$ bounds.

[3] For a detailed literature review of MGARCH models: Bauwens, Laurent and Rombouts (2006).

ACCEPTED MANUSCRIPT

Ng and Rothschild (1990) specify particular structures for the dynamics that reduce the parameters' dimensionality to the order M^2. For these models, the parameters' dimensionality can be further reduced to the order M when they are associated with the *variance targeting*[4] procedure of Engle and Mezrich (1996).

In order to handle "serious" cross-sectional dimensions M, recent works have moved toward models that can be estimated in two steps. Patton (2006) and Jondeau and Rockinger (2006) propose the copula-GARCH models, for which univariate GARCH processes describe the conditional variances and a copula function joins the marginal distributions to form a multivariate distribution function. Ledoit, Santa-Clara and Wolf (2003) propose the Flexible-GARCH in which the variances are modeled by univariate GARCH processes and the covariances by bivariate MGARCH models. The estimated coefficients are then *transformed* to achieve positive definiteness of the whole variance-covariance matrices. Both approaches, however, do not solve the dimensionality problem[5]. The copula-GARCH shifts it to the parameterization of the copula and the associated optimization problem, while the Flexible-GARCH shifts it to the minimization of the Frobenius norm that will enforce positivity.

The proposed method builds on the Constant Conditional Correlations (CCC) of Bollerslev (1990) and the Dynamic Conditional Correlations (DCC) of Engle (2002), which decompose the conditional variance-covariance matrix by separating the conditional variances from the conditional correlations. SCC extends this approach by further decomposing the conditional correlation matrix into its constituting components while preserving positive definiteness. Thus, SCC eliminates the dimensionality problem by making it possible to separately model and estimate such constituting elements without violating positivity and without imposing parameters' constraints.

The paper is organized as follows. Section 2 presents the Sequential Conditional Correlations methodology. The estimation strategy is described in Section 3 while its conformity to the standard GMM framework is derived in Section 4. Section 5 introduces the Autoregressive Conditional Correlations ACC(1,1) model used within the SCC steps, presents Monte Carlo results and shows, through an empirical application involving the conditional second moments of 69 selected stocks from the

[4]It consists in setting a model's unconditional variance-covariance matrix equal to its sample counterpart, thus eliminating $M(M+1)/2$ parameters from the optimization procedure.

[5]Specifications that eliminate the dimensionality problem altogether do exist in the literature, such as the scalar versions of the BEKK and DCC and the Orthogonal-GARCH of Alexander and Chibumba (1997), but they achieve such result under very strong assumptions about the dynamics of the elements of the conditional variance-covariance matrix.

ACCEPTED MANUSCRIPT

NASDAQ100, how the new procedure results in strikingly accurate measures of the conditional correlations. Section 6 concludes. Appendix A contains the proofs related to SCC's correlation matrix decomposition and Appendix B lays out the stationarity conditions for the ACC(1,1) model.

2 Sequential Conditional Correlations

In general, an M-dimensional vector stochastic process y_t, can be described in terms of its location and scale:

$$y_t = \mu_t(\lambda) + H_t^{1/2}(\lambda) \cdot \epsilon_t$$

where $\mu_t(\lambda)$ is a function describing the evolution of the mean vector conditional on the information set $\mathcal{I}_{t-1}$, $H_t^{1/2}(\lambda)$ is a function such that $H_t(\lambda)$ is the conditional variance-covariance matrix of the process y_t, λ is a vector of parameters, and ϵ_t is a vector of innovations satisfying the following moment conditions: $\mathbb{E}[\epsilon_t] = 0$ and $\mathbb{V}[\epsilon_t] = I_M$. For simplicity, throughout the paper, μ_t is assumed to be zero [6].

The Sequential Conditional Correlations fully exploit the approach of the Conditional Correlations model, introduced by Bollerslev (1990), of separating the various components of the conditional variance-covariance matrix. While the CCC model of Bollerslev (1990) and the DCC proposed by Engle (2002) separate the variances from the correlations, allowing for a two-step estimation procedure, SCC further separates the correlations and partial correlations, thus allowing for a true multi-step estimation procedure. In this setup, the conditional variance-covariance matrix is rewritten in the following form:

$$H_t = D_t K_{1,2,t} K_{1,3,t} \dots K_{M-1,M,t} K'_{M-1,M,t} \dots K'_{1,3,t} K'_{1,2,t} D_t \tag{1}$$

or in a more compact manner as:

$$H_t = D_t \Bigg(\prod_{i=1}^{M-1} \prod_{j=i+1}^{M} K_{i,j,t} \Bigg) \Bigg(\prod_{i=1}^{M-1} \prod_{j=i+1}^{M} K_{i,j,t} \Bigg)' D_t \tag{2}$$

where D_t is the $(M \times M)$ diagonal matrix of time-varying standard deviations. Notice the similarities, of the specification in (2) with the DCC, for which $H_t = D_t R_t D_t$, and the CCC, where R_t is constant. In order to guarantee positive definiteness of

[6] Alternatively, y_t is defined to be the residual from some filtration of the data. In general, the moment conditions corresponding to such filtration need to be included among those of SCC for the correct determination of its parameters' standard errors.

ACCEPTED MANUSCRIPT

the conditional correlations matrix R_t, the DCC needs to parameterize and estimate its elements jointly. On the other hand, SCC further decomposes the conditional correlations matrix down to its constituting elements, the K-matrices, thus achieving positivity and symmetry by construction. The matrices $K_{i,j,t}$ are lower triangular and their generic element $[row, col]$ is given by:

$$\mathbf{K_{i,j,t}[row, col]} = \begin{cases} \rho_{i,j,t} & \text{if } row = j \text{ and } col = i \\ (1 - \rho_{i,j,t}^2)^{1/2} & \text{if } row = j \text{ and } col = j \\ \mathbf{I}[row, col] & \text{otherwise} \end{cases}$$

where $\mathbf{I}$ is the identity matrix. The element $\rho_{i,j,t}$ is the time t correlation ($i = 1$) or partial correlation ($i > 1$) between $y_{i,t}$ and $y_{j,t}$. Appendix A, contains the proof of how any correlation matrix can be expressed as the product of a sequence of K-matrices. Further, it is shown that the product of any sequence of K-matrices with $|\rho| < 1$ is a correlation matrix.

3 Estimation

The complex estimation of an MGARCH is here translated into a sequence of simple estimations. This is done by working from the outside toward the inside of the specification in equation (1): estimate the elements of D_t and use it to standardize the data, estimate the elements of $K_{1,2,t}$ and use it to standardize the data, estimate $K_{1,3,t}$ and standardize, etc.

To properly unravel the $K-$matrix decomposition, let y_t be an M dimensional vector of observations at time t, and H_t be its variance-covariance matrix. The first step is to model the conditional variances of the M series $\forall t$ and use them to standardize the data:

$$\widehat{\epsilon}_{i,t} = \frac{y_{i,t}}{\widehat{h}_{i,t}^{1/2}} \qquad \forall i = 1, ..., M \text{ and } \forall t = 1, ..., T$$

This is equivalent to filling the D_t matrix with $\{\widehat{h}_{i,t}^{1/2}\}_{i=1}^{M}$ and pre-multiply y_t by $\widehat{D}_t^{-1}$. The *partially*-standardized vector $\widehat{\epsilon}_t$ is homoscedastic with unit variance but will still exhibit time varying correlations.

The second step involves the sequential estimation of the conditional correlations and partial correlations. Modeling the correlation of the first series with the remaining $M-1$ produces the following correlations estimates $\{\widehat{\rho}_{1,2,t}\}_{t=1}^{T}, \{\widehat{\rho}_{1,3,t}\}_{t=1}^{T}, ..., \{\widehat{\rho}_{1,M,t}\}_{t=1}^{T}$

ACCEPTED MANUSCRIPT

which will be used to partial-out the effect of $\widehat{\epsilon}_{1,t}$ from all the other series or, in other words, to further standardize the data:

$$\widetilde{\epsilon}_{i,t} = \frac{\widehat{\epsilon}_{i,t} - \widehat{\rho}_{1,i,t}\widehat{\epsilon}_{1,t}}{(1 - \widehat{\rho}^2_{1,i,t,})^{1/2}} \qquad \forall i > 1 \tag{3}$$

This is equivalent to filling the $K_{1,i,t}$ matrices with $\widehat{\rho}_{1,i,t}$ and carry out the following matrix multiplications: $\widehat{K}^{-1}_{1,M,t} \cdot \widehat{K}^{-1}_{1,M-1,t} \cdot ... \cdot \widehat{K}^{-1}_{1,2,t} \cdot \widehat{\epsilon}_t$. Since the transformed series $2, 3, ..., M$ are orthogonal to the first series, the claim of further standardization is justified.

The following steps consist in the estimation of the correlations of the second series with the remaining $M-2$ and the standardization of the latter, then the third with the remaining $M-3$ and standardization, and so on until the correlation of series $M-1$ and M is estimated.

The now *fully*-standardized series will exhibit a variance-covariance equal to the identity matrix $\forall t$. Furthermore, the estimated pairwise correlations and partial correlations, substituted in the corresponding $K-$matrices, will yield the conditional correlation matrix R_t and together with D_t the conditional variance-covariance matrix H_t.

Of the possible decompositions of the positive definite matrix H_t, the $K-$matrix of SCC is extremely convenient as it allows for the separate modeling of the variances and all the correlations components. The construction of R_t one correlation at the time reduces the highly dimensional optimization problem of MGARCH models of serious dimensions into a sequence of simple optimizations. In fact, every step of SCC requires the estimation of the time varying correlation of only two series. Such simplicity allows for more richly parameterized dynamics and functional forms for the volatilities and correlations.

Since SCC allows for the separate modeling of the variances and correlations of the conditional variance-covariance matrix H_t, standard estimation techniques can be employed for the first and will not be discussed here. As for the latter, they can be estimated by GMM based on the score of a Gaussian likelihood function or, equivalently, by Gaussian Quasi-Maximum-Likelihood (QML). Since at any step of the SCC's K-matrix decomposition the series $\epsilon_{i,t}$ and $\epsilon_{j,t}$, whose correlation is being modeled, are homoscedastic with unit variance, the concentrated Gaussian log-

ACCEPTED MANUSCRIPT

likelihood L_c is:

$$\begin{aligned} L_c &= -\sum_{t=1}^{T}\left[\ln(1-\rho_t^2)+\frac{\epsilon_{i,t}^2-2\rho_t\epsilon_{i,t}\epsilon_{j,t}+\epsilon_{j,t}^2}{1-\rho_t^2}\right] \\ &= -\sum_{t=1}^{T}\left[\ln(1-\rho_t^2)+\frac{(\epsilon_{j,t}-\rho_t\epsilon_{i,t})^2}{1-\rho_t^2}+\epsilon_{i,t}^2\right] \\ &= -\sum_{t=1}^{T}\left[\ln(1-\rho_t^2)+\frac{(\epsilon_{j,t}-\rho_t\epsilon_{i,t})^2}{1-\rho_t^2}\right] \end{aligned} \tag{4}$$

The elimination of the sum of the $\epsilon_{i,t}^2$ terms from the concentrated log-likelihood is justified by the fact that these terms do not depend on ρ_t. Notice how equation (4) reflects the *standardization* that takes place in the K-matrix decomposition (3), where the difference between $\epsilon_{j,t}$ and its conditional mean $\rho_t\epsilon_{i,t}$ is divided by the standard deviation $(1-\rho_t^2)^{1/2}$.

Since within SCC, the estimation of any bivariate model for the pairwise correlations occurs $M(M-1)/2$ times, it is convenient to employ an optimization scheme that does not fail to converge. In this sense, *MCMC* and *Simulated Annealing* are very robust as they do not rely on a particular shape of the objective function (concave and with non-zero second derivatives) for convergence.

4 Asymptotic Properties

Modeling the M conditional variances as univariate processes and estimating their parameters by GMM based on the score of a Gaussian likelihood leads to the following moment conditions:

$$\begin{aligned} m(\underline{y_{i,T}};\boldsymbol{\lambda_i}) &= \frac{1}{T}\sum_t m(y_{i,t};\boldsymbol{\lambda_i}) \\ &= -\frac{1}{T}\sum_t \frac{\partial}{\partial\boldsymbol{\lambda_i}}\left(\ln h_{i,t}+\frac{y_{i,t}^2}{h_{i,t}}\right) \end{aligned}$$

where λ_i for $i=1,...,M$ is the vector of coefficients of the volatility model for the time series i. Similarly, the moment conditions associated with the models for the conditional correlations are:

$$\begin{aligned} m(\underline{\epsilon_{i,T}},\underline{\epsilon_{j,T}};\boldsymbol{\lambda_{i,j}}) &= \frac{1}{T}\sum_t m(\epsilon_{i,t},\epsilon_{j,t};\boldsymbol{\lambda_{i,j}}) \\ &= -\frac{1}{T}\sum_t \frac{\partial}{\partial\boldsymbol{\lambda_{i,j}}}\left[\ln(1-\rho_{i,j,t}^2)+\frac{(\epsilon_{j,t}-\rho_{i,j,t}\epsilon_{i,t})^2}{1-\rho_{i,j,t}^2}\right] \end{aligned}$$

ACCEPTED MANUSCRIPT

where $\lambda_{i,j}$, for $i = 1, ..., M-1$ and $j = i+1, ..., M$, is the vector of parameters modeling the correlation or partial correlation between the series $\epsilon_{i,t}$ and $\epsilon_{j,t}$. The latter are functions of the vector of observables $\underline{y_T}$, the conditional variances $\underline{h_T}$ and the conditional correlations and partial correlations $\underline{\rho_{k,l,T}}$ with $k < i$ and $l = k+1, ..., j$. Therefore, in terms of the model's parameters:

$$\begin{aligned} \underline{\epsilon_{i,T}} &= \underline{\epsilon_{i,T}}\Big(\underline{y_T}; \lambda_1, ..., \lambda_M; \lambda_{1,2}, ..., \lambda_{i-1,i}\Big) \\ \underline{\epsilon_{j,T}} &= \underline{\epsilon_{j,T}}\Big(\underline{y_T}; \lambda_1, ..., \lambda_M; \lambda_{1,2}, ..., \lambda_{i-1,j}\Big) \end{aligned}$$

which substituted into the moment conditions for $\underline{\rho_{i,j,T}}$ give:

$$m(\underline{\epsilon_{i,T}}, \underline{\epsilon_{j,T}}; \boldsymbol{\lambda_{i,j}}) = m(\underline{y_T}; \lambda_1, ..., \lambda_M; \lambda_{1,2}, ..., \lambda_{i-1,j}; \boldsymbol{\lambda_{i,j}})$$

Collecting all the moment conditions, necessary for the estimation of the parameters of the variance and correlation processes in SCC, yields a block-triangular matrix:

$$m(\underline{y_T}; \boldsymbol{\lambda}) = \begin{cases} m(\underline{y_T}; \boldsymbol{\lambda_1}) \\ \vdots \\ m(\underline{y_T}; \boldsymbol{\lambda_M}) \\ m(\underline{y_T}; \lambda_1, ..., \lambda_M; \boldsymbol{\lambda_{1,2}}) \\ \vdots \\ m(\underline{y_T}; \lambda_1, ..., \lambda_M; \boldsymbol{\lambda_{1,M}}) \\ m(\underline{y_T}; \lambda_1, ..., \lambda_M; \lambda_{1,2}, ..., \lambda_{1,M}; \boldsymbol{\lambda_{2,3}}) \\ \vdots \\ m(\underline{y_T}; \lambda_1, ..., \lambda_M; \lambda_{1,2}, ..., \lambda_{1,M}; \boldsymbol{\lambda_{2,M}}) \\ \vdots \\ \vdots \\ m(\underline{y_T}; \lambda_1, ..., \lambda_M; \lambda_{1,2}, ..., \lambda_{M-2,M}; \boldsymbol{\lambda_{M-1,M}}) \end{cases}$$

The estimates are obtained by setting all the equations to zero and simultaneously solving for λ: $m(\underline{y_T}; \widehat{\boldsymbol{\lambda}}) = 0$. However, given the particular structure of moment conditions this is the same as solving for one set of moments at the time and substituting the so obtained estimates in the following sets as they appear from the above ordering. The equality of the step-by-step solution to the simultaneous solution allows the SCC estimation to fall within the GMM framework. Hence:

$$\sqrt{T}(\widehat{\boldsymbol{\lambda}} - \boldsymbol{\lambda_0}) \xrightarrow{d} \boldsymbol{N}\Big(0, G^{-1} J (G^{-1})'\Big) \tag{5}$$

ACCEPTED MANUSCRIPT

where:

$$G = \mathbb{E}\Big[\frac{\partial}{\partial\lambda} m(y_t;\boldsymbol{\lambda_0})\Big] \qquad \text{and} \qquad J = \mathbb{V}\Big[m(y_t;\boldsymbol{\lambda_0})\Big]$$

which can be estimated by their sample counterparts:

$$\widehat{G} = \frac{1}{T}\sum_t \frac{\partial}{\partial\lambda} m(y_t;\widehat{\boldsymbol{\lambda}}) \qquad \text{and} \qquad \widehat{J} = \frac{1}{T}\sum_t m(y_t;\widehat{\boldsymbol{\lambda}})\cdot m(y_t;\widehat{\boldsymbol{\lambda}})' \tag{6}$$

The assumptions required for consistency and asymptotic normality of the estimator are very difficult to check for most MGARCH models[7], as well as some univariate GARCH specifications, and SCC is no exception. In Section 5.2 a Monte Carlo experiment has been conducted to asses the goodness of the estimates' correspondence to standard asymptotic theory.

If *targeting* is employed the fact that the population mean of the process has been set equal to its sample counterpart needs to be included among the moment conditions. Specifically, let $\lambda^{\bullet}_{i,j}$ be the element of the vector $\lambda_{i,j}$ that the econometrician might choose to target and $\lambda^{\circ}_{i,j}$ be the vector containing the remaining parameters. Then, the moment conditions associated with a bivariate model for the conditional correlations will depend on whether $\lambda^{\bullet}_{i,j}$ is treated as a parameter:

$$m(\underline{\epsilon_{i,T}}, \underline{\epsilon_{j,T}}, \lambda_{i,j}) = -\frac{1}{T}\sum_t \frac{\partial}{\partial\lambda_{i,j}}\left[\ln(1-\rho^2_{i,j,t}) + \frac{(\epsilon_{j,t}-\rho_{i,j,t}\epsilon_{i,j})^2}{1-\rho^2_{i,j,t}}\right]$$

where:

$$\lambda'_{i,j} = \left[\lambda^{\bullet}_{i,j}, \left(\lambda^{\circ}_{i,j}\right)'\right]$$

or is being targeted to the sample correlation $\overline{\rho}_{ij}$:

$$m(\underline{\epsilon_{i,T}}, \underline{\epsilon_{j,T}}, \lambda^{\bullet}_{i,j}, \lambda^{\circ}_{i,j}) = \begin{cases} \frac{-1}{T}\sum_t \frac{\partial}{\partial\lambda^{\circ}_{i,j}}\left[\ln(1-\rho^2_{i,j,t}) + \frac{(\epsilon_{j,t}-\rho_{i,j,t}\epsilon_{i,j})^2}{1-\rho^2_{i,j,t}}\right] \\ \frac{1}{T}\sum_t \rho_{i,j,t}\left(\lambda^{\bullet}_{i,j}, \lambda^{\circ}_{i,j}\right) - \overline{\rho}_{i,j} \end{cases}$$

Despite the theoretical convenience of being able to refer to standard GMM asymptotic results, in practice the computation of the parameters' variance-covariance matrix within the SCC methodology parallels that of other MGARCH models proposed in the literature where no computational short-cuts are available. Furthermore, since the dimensions of the parameters' variance-covariance matrix are at least[8] of order

[7] For example, Comte and Lieberman (2003) and Hafner and Preminger (2009) verify the conditions for consistency and asymptotic normality of the QMLE, but Chen and Fan (2006) and Kawakatsu (2006) do not.

[8] A lower bound arises from the minimal presence of at least one parameter per correlation.

ACCEPTED MANUSCRIPT

M^2, the matrix $\widehat{J}$, computed as the average of T outer-products of the moment conditions, will not be full-rank for large M. It is important to realize that this is a feature inherent to all MGARCH models, whether it is evident or hidden in a multi-step procedure, targeting, orthogonalization, etc.[9]

5 Application

5.1 Autoregressive Conditional Correlations

The proposed bivariate model, to be used in the SCC steps, is a model for the correlations rather than for the corresponding variance-covariance matrix. Modeling the correlations ρ_t directly will require the imposition of certain constraints on the parameters that govern their dynamics in order to have correlations that are bounded between plus and minus one. A more convenient approach is to model their Fisher transformation χ_t:

$$\chi_t = \frac{1}{2}\ln\left(\frac{1+\rho_t}{1-\rho_t}\right)$$

which maps the interval $(-1,1)$ into $(-\infty,+\infty)$, thus allowing the model's parameters to exclusively reflect the dynamic behavior of the process. Letting ϕ_t be some measure of the *realized* correlation at time t, its Fisher transformation ψ_t is given by:

$$\psi_t = \frac{1}{2}\ln\left(\frac{1+\phi_t}{1-\phi_t}\right)$$

The general ACC(p,q) model will then take the following form:

$$\chi_t = \omega + \sum_{j=1}^{p}\delta_j\chi_{t-j} + \sum_{i=1}^{q}(\theta_i + \beta_i d_{t-i})\psi_{t-i}$$

where d_{t-i} are dummy variables that allow for an asymmetric response to the realizations through the coefficients β_i. The conditional correlation ρ_t is then given by the inverse-Fisher transformation of χ_t:

$$\rho_t = \frac{\exp(2\chi_t)-1}{\exp(2\chi_t)+1}$$

The realized correlations ϕ_t are not observed and therefore require to be extracted from the data through some method. In this paper they are computed using an

[9]For example, for $M = 69$ the dimensions of the parameters' variance covariance matrix are: $2,691$ for the CCC, $2,694$ for the scalar DCC with asymmetry parameter, $12,065$ for the full SCC and would be $11,666,865$ for the full VEC(1,1). Thus, for a sample $T = 2517$ even a constant correlation model does not yield a full-rank and invertible variance-covariance matrix of the parameters.

ACCEPTED MANUSCRIPT

exponential smoothing, with parameter α, of past realizations. Letting Q_{t-2} be the realized bi-dimensional variance-covariance matrix at time $t-2$, the corresponding realized measure at time $t-1$ will be given by:

$$Q_{t-1} = \alpha Q_{t-2} + (1-\alpha) u_{t-1} u'_{t-1}$$

where u_{t-1} is the vector containing the standardized realizations of the two series whose time-varying correlation is being modeled. From Q_{t-1} it is straightforward to compute the realized correlation ϕ_{t-1}

$$\phi_{t-1} = \frac{Q_{t-1}[1,2]}{\sqrt{Q_{t-1}[1,1] \cdot Q_{t-1}[2,2]}}$$

The optimal value of the smoothing parameter will be determined by the data by simultaneously estimating α with the parameters $\delta_j, \theta_i, \beta_i$ of the ACC(p,q). Alternatively, α could be chosen *a priori* by the econometrician or be the result of a preliminary and separate estimation. Other measures of the realized correlations, besides the exponential smoothing, may be employed.

In this empirical application, the conditional correlations are modeled by an ACC(1,1) with asymmetry parameter:

$$\chi_t = \omega + \delta \chi_{t-1} + (\theta + \beta d_{t-1}) \psi_{t-1}$$

where:

$$d_{t-1} = \begin{cases} 1 & \text{if } u_{1,t-1} < 0 \text{ and } u_{2,t-1} < 0 \\ 0 & \text{otherwise} \end{cases}$$

Furthermore, targeting of the parameter ω will be employed so that the long-run predictions of the process match the sample unconditional value:

$$\chi_t = (\overline{\chi} - \delta \overline{\chi} - \theta \overline{\psi} - \beta \overline{d\psi}) + \delta \chi_{t-1} + (\theta + \beta d_{t-1}) \psi_{t-1}$$

with $\overline{\psi} = (1/T) \sum_{t=1}^{T} \psi_t$, $\overline{d\psi} = (1/T) \sum_{t=1}^{T} d_t \psi_t$, $\overline{\chi} = (1/2) \ln[(1+\overline{\rho})/(1-\overline{\rho})], \overline{\rho} = (1/T) \sum_{t=1}^{T} u_{1,t} u_{2,t}$. In Appendix B, conditions on the model's parameters are derived for an ACC(1,1) to be mean and covariance stationary under the set of assumptions A1-A4. Mean stationarity requires that:

$$|\delta + \theta + \beta \mathbb{E}[d]| < 1$$

and variance stationarity:

$$(\delta + \theta + \beta \mathbb{E}[d])^2 + \beta^2 \mathbb{E}[d](1 - \mathbb{E}[d]) < 1$$

ACCEPTED MANUSCRIPT

Since $\mathbb{E}[d]$ is not known, stationarity will be imposed by replacing the population value with its sample counterpart $\overline{d} = (1/T)\sum_{t=1}^{T} d_t$. Appendix B further shows how forecasts, conditional on a given information set, can be computed iteratively using the following recursive formula:

$$\mathbb{E}[\chi_{t+1+j}|\mathcal{I}_{t-1}] = (1 - \delta - \theta - \beta\overline{d})\overline{\chi} + (\delta + \theta + \beta\overline{d})\mathbb{E}[\chi_{t+j}|\mathcal{I}_{t-1}] \qquad \forall j \geq 0$$

5.2 Monte Carlo Evidence

Theoretical results for the population properties of most MGARCH models proposed in the literature are quite hard to derive and SCC is no exception. A Monte Carlo experiment has been conducted to evaluate the goodness of standard asymptotic approximations. Artificial data is generated with time dimension $T = 3000$, cross-sectional dimension $M = 3$ and Monte Carlo replications $S = 10000$. Since the purpose of this simulation is to asses the behavior of the SCC methodology and in particular the multi-step procedure arising from the proposed correlation matrix decomposition, simpler models for the variances and correlations have been selected as data generating processes. In particular, the conditional variances h_t have been generated by a GARCH(1,1):

$$h_t = \omega + \beta h_{t-1} + \alpha y_{t-1}^2$$

and the conditional correlations ρ_t by a basic ACC(1,1) with $\beta = 0$ and $\alpha = \delta$.

The simulation results are reported in Table 2 where it can be seen that the Monte Carlo averages of the parameters' estimates $\overline{\lambda}_S$ are very close to the true values λ_0. Furthermore, the theoretical standard errors σ, computed according to equations (5) and (6), are also in agreement with the standard errors of the Monte Carlo experiment σ_S. Coverage of the Gaussian distribution is overall reasonable when testing at 95% while it shows an average downward bias of 2% when testing at 90% and an average upward bias of 1% when testing at 99%. These discrepancies between the finite sample distribution and the asymptotic approximation are in line with those of similar models such as the univariate GARCH and are the result, among others, of stationarity constraints affecting the estimates only from one side.

5.3 Evaluation

A natural way to evaluate MGARCH specifications would be to compare the model's predictions with the realizations. Unfortunately the realizations of the conditional variance-covariance matrices are not observable and the outer products $y_t y_t' =$

ACCEPTED MANUSCRIPT

$H_t^{1/2}\epsilon_t\epsilon_t'(H_t^{1/2})'$, while unbiased estimators of the quantities of interest, are very noisy measurements because of the idiosyncratic term $\epsilon_t\epsilon_t'$. Andersen and Bollerslev (1998) propose the use of *Integrated Volatility* as a better measure of the realized conditional variance-covariance matrix. In a setting where high frequency data is not available, a less precise but still useful proxy of the Integrated Volatility can be employed as in Ledoit, Santa-Clara, and Wolf (2003). With respect to correlations, the idea is to construct a measure $\widehat{\Upsilon}$ of the *realized correlation* from the *realized variance-covariance matrix* $\widehat{Q}$ by taking a rolling-window average over K periods of the outer products of the standardized returns $\widehat{\epsilon}_{i,t} = y_{i,t}/\sqrt{h_{i,t}}$:

$$\begin{aligned}\widehat{Q}_t^{(K)} &= \frac{1}{K}\sum_{k=1}^{K}\widehat{\epsilon}_{t-1+k}\cdot\widehat{\epsilon}_{t-1+k}' \qquad \forall t = 1, ..., T-K+1\\ \widehat{\Upsilon}_t^{(K)} &= \widetilde{Q}_t^{(K)}\cdot\widehat{Q}_t^{(K)}\cdot\widetilde{Q}_t^{(K)}\end{aligned}$$

where $\widetilde{Q}_t^{(K)}$ is a diagonal matrix containing the inverse of the square root of the elements on the main diagonal of $\widehat{Q}_t^{(K)}$. The model's prediction $\widehat{R}$ for the average correlations over the same period and window will be given by:

$$\widehat{R}_t^{(K)} = \frac{1}{K}\sum_{k=1}^{K}\left(\widehat{R}_{t-1+k}|\mathcal{I}_{t-1}\right) \qquad \forall t = 1, ..., T-K+1$$

where the term in parenthesis is the model's predicted correlation matrix conditional on the information set $\mathcal{I}_{t-1}$. The distance between *predictions* and *realizations* can be measured by mean square error MSE and mean absolute deviation MAD:

$$\begin{aligned}\text{MSE} &= \frac{1}{T-K+1}\sum_{t=1}^{T-K+1}\left[\sum_{r=1}^{M}\sum_{c=r+1}^{M}\left(\widehat{\Upsilon}_{r,c,t}^{(K)} - \widehat{R}_{r,c,t}^{(K)}\right)^2\right]\\ \text{MAD} &= \frac{1}{T-K+1}\sum_{t=1}^{T-K+1}\left[\sum_{r=1}^{M}\sum_{c=r+1}^{M}\left|\widehat{\Upsilon}_{r,c,t}^{(K)} - \widehat{R}_{r,c,t}^{(K)}\right|\right]\end{aligned}$$

While in the MGARCH literature the MAD tends to be preferred to the MSE on the grounds of robustness, Patton (2008) showed that using the MAD can lead to the conclusion that the perfect forecast is inferior to a biased one. In this paper, in order to evaluate conditional correlations' predictions, both the MSE and the MAD will be computed even though only the first is to be considered a reliable indicator.

For $K \to T$ the models' predictions converge to the long-run mean ($\widehat{R}_t^{(K)} \to \overline{R}$) and the realized measures converge to the unconditional mean ($\widehat{\Upsilon}_t^{(K)} \to \overline{R}$). This implies that the MSE and MAD are decreasing in K, giving the misleading impression

ACCEPTED MANUSCRIPT

that the models' performances increase with K. Hence, the measure of the goodness of a model, relative to a *constant*, is downward biased and the size of the bias increases with K. Since the variances of the *realized* measures $\widehat{\Upsilon}_t^{(K)}$, and therefore those of the MSE and MAD, decrease in K, it follows that the choice of K implies a trade-off between the variance and bias of the MSE and MAD.

Predicted variances-covariances and correlations can either be out of sample or in sample. In the second case, the whole sample is used for estimation and to generate the K-steps ahead predictions. In other words, $\widehat{H}_t^{(K)}$ and $\widehat{R}_t^{(K)}$ will be computed starting from the fitted values $\widehat{H}_{t-1}$ and $\widehat{R}_{t-1}$ followed by the model's predictions for the next K periods. For every t both $\widehat{H}_t^{(K)}$ and $\widehat{R}_t^{(K)}$ do not depend on the information available after $t-1$ except for that already contained in the parameters' estimates. It must be emphasized how the in sample predictions, like the out of sample predictions or forecasts, cannot be improved simply by increasing the number of parameters as it happens for the fit. In order to understand the reason behind this, it is enough to realize that the predictions $\widehat{H}_t^{(K)}$ and $\widehat{R}_t^{(K)}$ depend on i) the initial fitted values $\widehat{H}_{t-1}$ and $\widehat{R}_{t-1}$ and ii) the rate of convergence to the unconditional values $\overline{H}$ and $\overline{R}$. Therefore, an over-parameterized model will produce larger MSE and MAD due to the fact that its predictions are generated by more noisy rates of convergence and more noisy initial conditions.

The K-matrix decomposition is unique for a given correlation matrix R. However, a permutation of the order of the series which is reflected in a permutation of the elements of R does lead to a different sequence of $K_{1,2}, K_{1,3}, ..., K_{M-1,M}$ matrices. If the true correlations and conditional correlations were known, such permutation would not affect the results. On the other hand, when correlations are estimated the ordering of the series might have some impact on the results. A closer look at the K-matrix decomposition, however, does suggest an appropriate ordering method. The first $M-1$ correlations are estimated on the data standardized by the conditional variances, while the following $M-2$ correlations are estimated on the same data appropriately transformed to remove the effect of ϵ_1. The latter transformation will inevitably introduce some noise as it is a function of the results of the first $M-1$ estimations. From the K-matrix decomposition, correlations on the second row of $\widehat{R}_t$ have the following expression: $(1-\widehat{\rho}_{1,i,t}^2)^{1/2} \cdot \widehat{\rho}_{2,i,t}$. If $\widehat{\rho}_{1,i}$ is big in modulus, $(1-\widehat{\rho}_{1,i}^2)^{1/2}$ is small and so is the effect of $\widehat{\rho}_{2,i}$ on the product. Since similar structures do occur for all the rows of R_t and since the noise in the estimates is expected to increase as SCC moves from the top to the bottom rows of R_t, it appears sensible to counterbalance

ACCEPTED MANUSCRIPT

such noise by ordering the series in the data set in decreasing order of total correlation:

$$\sum_{j=1}^{M} \overline{\rho}_{1,j}^2 > \sum_{j=1}^{M} \overline{\rho}_{2,j}^2 > ... > \sum_{j=1}^{M} \overline{\rho}_{M,j}^2$$

SCC based on this ordering will be denoted by SCC(↓).

5.4 Data

The data employed in this paper consists of $M = 69$ series from the NASDAQ100. The sample is 10 years of daily observations (from 9/1/1994 to 8/31/2004) for a total of 2517 returns. 31 series have not been included in this study because there were not enough observations available: from Apollo Group (2454 obs.) to Kmart Holding (340 obs.). A list of symbols of the included stocks can be found in Table 1.

Returns have been calculated as log-differences of closing prices. Unconditional means were subtracted from each series of returns before proceeding to the estimation of the conditional second moments. SCC's conditional variances have been modeled using Zakoian's (1994) specification of univariate threshold GARCH(1,1) which not only provides a good fit to the data but also goes along well with automated procedures as it is relatively easy to implement and converge. Conditional Correlations have been modeled by Autoregressive Conditional Correlations ACC(1,1) processes with asymmetry.

5.5 Results

To better evaluate its performance, SCC will be compared to the CCC and the scalar asymmetric DCC. The reason why more sophisticated MGARCH models are not introduced in the comparison is because they cannot be feasibly estimated for such a large set of data. For example, the DCC where correlations are modeled by a diagonal BEKK requires the simultaneous optimization of the objective function (Quasi-Gaussian-Likelihood) with respect to $3M = 207$ parameters.

Table 3 reports the differences between the models' predicted conditional correlations and their realizations. Such differences are measured by the MSE and the MAD. While DCC reduces the MSE with respect to CCC of no more than 9.00%, SCC(↓) improves it by more than 30%. It has been discussed at length in Section 5.3 how the measures of the models' performances are contaminated by *noisy* measurements of the *realized correlations*, over short horizons, and how the models' predictions converge to the long-run means, over long horizons. This is confirmed by the apparent

ACCEPTED MANUSCRIPT

performances of both DCC and SCC with respect to CCC: increasing with the length of the horizon up to a certain point and decreasing after that.

The striking performance of SCC, which on average can explain 30% of the variation of the 2346 time-varying correlations cannot be simply attributed to the fact that with its $12,075$ parameters it is over-parameterized and thus over-fitting. If this was the case the estimated rates of convergence to the unconditional values and the initial values of the predictions would be contaminated by additional noise, thus compromising the measure of the average correlations implied by the model over the given horizon, as discussed in Section 5.3. In order to better asses the ability of SCC to track the correlations, it is worth noticing that the 30% reduction in MSE is comparable to the performances of univariate GARCH models in tracking variances. In particular, if the same in sample criteria was used to evaluate a univariate GARCH, results in terms of MSE reduction would be around 40%. This not only implies that correlations are time varying but also that their movements can be accuratelly described by the SCC and ACC(1,1) combination.

Table 4 shows the results for some permutations of the series' order in SCC. While for SCC(A) they are ordered alphabetically, in SCC(Z) such order is reversed. In SCC(1) through SCC(8) the series have been randomly ordered. Both the MSE and the MAD are affected by the ordering even though in every case their reduction with respect to CCC and DCC is substantial. The Mean Squared Error improvement ranges from 29% to 33% while for the Mean Absolute Deviation it goes from 14% to 18%.

Since it is common practice to evaluate multivariate volatility models in terms of likelihood-based statistics[10] these have been reported in Table 5. The p-values of the Likelihood-Ratio-Test (LRT) show clear rejection of CCC when compared to both DCC and SCC. In terms of the Schwartz Information Criteria (SIC), the preferred model is DCC while SCC is outperformed by CCC. This is the results of the high likelihood value of SCC being offset by a heavy penalty for the large number of model's parameters. However, if SIC is used for model selection it makes sense that it should also be used to determine the best SCC specification. In particular, which conditional correlations and partial correlations should be modeled as time varying and which should be set constant. Therefore, the SCC($\downarrow$) has been re-estimated and at each step

[10]Corresponding results should be taken *cum grano salis* since i) QML is, at best, based on approximations of the distribution of financial data and ii) estimates of models that are carried out in two or more steps, such as CCC, DCC and SCC, do not actually maximize the Quasi-Likelihood function.

ACCEPTED MANUSCRIPT

the ACC(1,1) has been compared to a constant correlation model[11] in terms of the SIC, yielding the SCC_SIC specification. This is a very parsimonious model that can still explain very well the variations in the time-varying correlations, as can be seen in Table 4. Furthermore, Table 5 shows that in terms of the information criterion it is the preferred model. The LRT, on the other hand, rejects the SCC_SIC in favor of the full SCC. This is due to the heavy penalty placed by the SIC on the parameters' dimensionality, corresponding to a LRT conducted at a significance level of less than 10^{-14}.

6 Conclusions

This paper develops a new approach for the modeling of conditional correlation matrices. The structure of the SCC allows for a sequential estimation of the various components of such matrices. This makes it possible to eliminate the traditional intractability of MGARCH models of non-trivial dimensions. The multi step procedure translates the highly dimensional optimization problem associated with MGARCH models into a series of simple univariate and bivariate estimations. Because of the particular decomposition of the variance-covariance matrix in the SCC, the latter is guaranteed to be symmetric and positive definite by construction, without imposing any parametric restriction.

Another appealing feature of the SCC is its flexibility in the modeling of the various correlations: not only there are no constraints on the choice of the model, but different bivariate models can be used at different steps. This paper has introduced the bivariate ACC which allows for a data-dependent choice of the best proxy for the lagged realizations, for an immediate interpretation of the model's parameters and guarantees that the correlations are bounded between plus and minus one without imposing any parametric restriction. The empirical results on the selected 69 stocks from the NASDAQ100 show how SCC and ACC(1,1) can significantly improve the tracking of the conditional correlations compared to CCC and DCC. While the improvement with respect to CCC is only 9% for DCC, it increases dramatically to more than 30% for SCC in terms of MSE.

The Sequential Conditional Correlations methodology, developed in this paper, could be extended to incorporate with ease exogenous explanatory variables. The

[11] A finer approach would require that at each step the constant correlation, the full ACC(1,1) and the ACC(1,1) with various combinations of parameters' restrictions be evaluated.

ACCEPTED MANUSCRIPT

inclusion of exogenous regressors is a relevant feature as it allows the researcher to uncover potentially important economic relationships that until now could not be effectively investigated. Another obvious extension of the proposed methodology is to the multivariate estimation of *realized* quantities in the high frequency environment. Estimators of the constituting elements of the correlation matrix, that are consistent in the presence of microstructure noise and asynchronous trading, could be coupled with the SCC methodology which, starting from these building blocks, would deliver a positive definite matrix. These extensions and their applications are left as an area for future research.

A K-Matrix Decomposition

Definition A.1 *A correlation matrix R is a symmetric positive definite matrix giving the correlations between all pairs of data sets.*

Necessary but not sufficient conditions for a matrix to be a correlation matrix are that all the elements on the main diagonal must be equal to 1 (a random variable co-varies perfectly with itself) and that all the off-diagonal elements must be less than 1 in modulus[12].

Definition A.2 *$K_{p,q}^{-1}$, with $p < q$, is a lower triangular matrix of the form:*

$$\mathbf{K_{p,q}^{-1}[i,j]} = \begin{cases} -\rho(1-\rho^2)^{-1/2} & \text{if } i = q \text{ and } j = p \\ (1-\rho^2)^{-1/2} & \text{if } i = q \text{ and } j = q \\ \mathbf{I}[i,j] & \text{otherwise} \end{cases}$$

where ρ is the (p,q) element of the correlation matrix that is pre-multiplied by $K_{p,q}^{-1}$. It follows from the necessary conditions of a correlation matrix that $|\rho| < 1$. In other words, $K_{p,q}^{-1}$ is an operator that "generates" a matrix whose elements are function of the (p,q) element of the correlation matrix it is being applied to.

Theorem A.1 *For any $(M \times M)$ correlation matrix R, there exists a unique sequence of matrices $K_{1,2}^{-1}, K_{1,3}^{-1}, ..., K_{M-1,M}^{-1}$ such that:*

$$I = K_{M-1,M}^{-1} \cdot ... \cdot K_{1,3}^{-1} \cdot K_{1,2}^{-1} \cdot R \cdot (K_{1,2}^{-1})' \cdot (K_{1,3}^{-1})' \cdot ... \cdot (K_{M-1,M}^{-1})'$$

Proof of Theorem A.1:

[12]This excludes the case of perfectly correlated variables.

ACCEPTED MANUSCRIPT

Definition A.3 *Let $R_{p,q}$, with $p < q$, be an $(M \times M)$ correlation matrix of the following form:*

$$\mathbf{R_{p,q}} = \begin{pmatrix} & & & 0 & & & & & \\ & \mathbf{I}_{(p-1)} & & 0 & & & \mathbf{0} & & \\ & & & \vdots & & & & & \\ 0 & 0 & \dots & 1 & 0 & \dots & \rho & \# & \# \\ & & & 0 & & & & & \\ & & & \vdots & & & & & \\ & \mathbf{0} & & \rho & & & \mathbf{F}_{(M-p)} & & \\ & & & \# & & & & & \\ & & & \# & & & & & \end{pmatrix}$$

where ρ is the $[p,q]$ element, $\mathbf{I}$ is the identity matrix, $\mathbf{F}$ is a non-trivial correlation matrix, and $\#$ indicates (generally) non-zero elements.

Lemma A.1 *$S_{p,q} = K_{p,q}^{-1} R_{p,q} (K_{p,q}^{-1})'$ is a correlation matrix such that:*

$$\mathbf{S_{p,q}}[\mathbf{i},\mathbf{j}] = 0 \qquad \text{if } i \leq p,\ j \leq q \text{ and } i \neq j$$

Proof:

$$\mathbf{S_{p,q}} = \begin{pmatrix} & & 0 & \dots & & s_1 & \dots & 0 \\ & \mathbf{I}_{(p-1)} & 0 & \dots & & s_2 & \dots & 0 \\ & & \vdots & & & \vdots & & \vdots \\ 0 & 0 \dots & 1 & 0 & \dots & s_p & \# & \# \\ \vdots & \vdots & 0 & F_{1,1} & \dots & s_{p+1} & \dots & F_{1,M-p} \\ & & \vdots & \vdots & & \vdots & & \vdots \\ s_1 & s_2 \dots & s_p & s_{p+1} & \dots & s_q & \dots & s_M \\ \vdots & \vdots & \# & \vdots & & \vdots & & \vdots \\ 0 & 0 \dots & \# & F_{1,M-p} & \dots & s_M & \dots & F_{M-p,M-p} \end{pmatrix}$$

where:

$$s_i = \begin{cases} -\frac{\rho}{(1-\rho^2)^{1/2}} R_{pi} + \frac{1}{(1-\rho^2)^{1/2}} R_{qi} & \text{if } i \neq q \\ \frac{\rho^2}{(1-\rho^2)} R_{pp} - \frac{2\rho}{(1-\rho^2)} R_{pq} + \frac{1}{(1-\rho^2)} R_{qq} & \text{if } i = q \end{cases}$$

ACCEPTED MANUSCRIPT

allowing to easily compute the following values of interest:

$$\begin{aligned} i < p \;&\Rightarrow\; R_{pi} = 0, R_{qi} = 0 \;&\Rightarrow\; s_i = 0 \\ i = p \;&\Rightarrow\; R_{pp} = 1, R_{qp} = \rho \;&\Rightarrow\; s_p = 0 \\ i = q \;&\Rightarrow\; R_{pp} = 1, R_{qq} = 1, R_{pq} = \rho \;&\Rightarrow\; s_q = 1 \end{aligned}$$

which substituted back in $S_{p,q}$ yield:

$$\mathbf{S_{p,q}} = \begin{pmatrix} & & & 0 & \dots & & 0 & \dots & 0 \\ & \mathbf{I}_{(p-1)} & & 0 & \dots & & 0 & \dots & 0 \\ & & & \vdots & & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 & \# & \# \\ \vdots & \vdots & & 0 & F_{1,1} & \dots & s_{p+1} & \dots & F_{1,M-p} \\ & & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & s_{p+1} & \dots & 1 & \dots & s_M \\ \vdots & \vdots & & \# & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & \# & F_{1,M-p} & \dots & s_M & \dots & F_{M-p,M-p} \end{pmatrix}$$

The $F-$block of $S_{p,q}$ has every element on the main diagonal equal to 1. Furthermore, since $S_{p,q}$ is a quadratic form of a positive definite matrix, it is itself positive definite and therefore also the $F-$block must be positive definite. Hence, it is a correlation matrix:

$$\mathbf{S_{p,q}} = \begin{pmatrix} & & & 0 & \dots & & 0 & \dots & 0 \\ & \mathbf{I}_{(p-1)} & & 0 & \dots & & 0 & \dots & 0 \\ & & & \vdots & & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 & \# & \# \\ \vdots & \vdots & & 0 & & & & & \\ & & & \vdots & & & & & \\ 0 & 0 & \dots & 0 & & & \widehat{\mathbf{F}}_{(M-p)} & & \\ \vdots & \vdots & & \# & & & & & \\ 0 & 0 & \dots & \# & & & & & \end{pmatrix}$$

Definition A.4 *$R_{p,q+1} = S_{p,q}$ if $q+1 \leq M$ or $R_{p+1,M} = S_{p,q}$ if $q+1 > M$.*

Lemma A.2 *Given a correlation matrix $R_{p,q}$, successive pre- and post-multiplication by the matrices $K_{p,q}^{-1}, K_{p,q+1}^{-1}, ..., K_{M-1,M}^{-1}$ will yield the identity matrix.*

ACCEPTED MANUSCRIPT

Proof:

The application of the matrix $K_{r,s}^{-1}$ and its transpose has the property of changing the correlation in the r-row and s-column (and therefore the s-row and r-column) into a zero without affecting the correlations that have been set to zero by the matrices $K_{i,j}^{-1}$ with $i \leq r$ or $j \geq s$:

$$I = K_{M-1,M}^{-1} \cdot K_{M-2,M}^{-1} \cdot \ldots \cdot K_{p,q}^{-1} \cdot R_{p,q} \cdot (K_{p,q}^{-1})' \cdot \ldots (K_{M-2,M}^{-1})' \cdot (K_{M-1,M}^{-1})'$$

To complete the proof of **Theorem A.1** the results so obtained for the $R_{p,q}$ type correlation matrix need to be extended to *any correlation matrix R*. This is immediate, once it is recognized that $R_{1,2}$ is indeed *any correlation matrix R*.

Definition A.5 *$K_{p,q} = (K_{p,q}^{-1})^{-1}$ is a lower triangular matrix of the form:*

$$\mathbf{K_{p,q}[i,j]} = \begin{cases} \rho & \text{if } i = q \text{ and } j = p \\ (1-\rho^2)^{1/2} & \text{if } i = q \text{ and } j = q \\ \mathbf{I}[i,j] & \text{otherwise} \end{cases}$$

where ρ is the (p,q) element of the $K_{p,q}^{-1}$ matrix and therefore $|\rho| < 1$.
In other words, $K_{p,q}$ is the inverse of the matrix given by the application of the operator $K_{p,q}^{-1}$ to a correlation matrix.

Corollary A.1 *to* **Theorem A.1:** *any $(M \times M)$ correlation matrix R can be expressed as the product of a sequence of matrices $K_{1,2}, K_{1,3}, ..., K_{M-1,M}$:*

$$R = K_{1,2} \cdot K_{1,3} \cdot ... \cdot K_{M-1,M} \cdot K'_{M-1,M} \cdot ... \cdot K'_{1,3} \cdot K'_{1,2}$$

Corollary A.1 states that every correlation matrix can be expressed as the product of a sequence of K matrices. What it does not say is whether the product of *any* sequence of K matrices is a correlation matrix.

Theorem A.2 *The product of any sequence of K matrices, taken in the appropriate order, is a correlation matrix.*

Proof of Theorem A.2:

Let $R_{p,q}$ be the correlation matrix of Definition A.3 and without loss of generality let $q - 1 > p$.

ACCEPTED MANUSCRIPT

Lemma A.3 *$R_{p,q-1} = K_{p,q-1} R_{p,q} K'_{p,q-1}$ is a correlation matrix.*
Proof:

$$\mathbf{R_{p,q-1}} = \begin{pmatrix} & & & 0 & \dots & & x_1 & \dots & 0 \\ & \mathbf{I}_{(p-1)} & & 0 & \dots & & x_2 & \dots & 0 \\ & & & \vdots & & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & x_p & \# & \# \\ \vdots & \vdots & & 0 & F_{1,1} & \dots & x_{p+1} & \dots & F_{1,M-p} \\ & & & \vdots & \vdots & & \vdots & & \vdots \\ x_1 & x_2 & \dots & x_p & x_{p+1} & \dots & x_{q-1} & \dots & x_M \\ \vdots & \vdots & & \# & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & \# & F_{1,M-p} & \dots & x_M & \dots & F_{M-p,M-p} \end{pmatrix}$$

where:

$$\mathbf{x_i} = \begin{cases} \rho R_{pi} + (1-\rho^2)^{1/2} R_{(q-1)i} & \text{if } i \neq q-1 \\ \rho^2 R_{pp} + 2\rho(1-\rho^2)^{1/2} R_{(q-1)p} + (1-\rho^2) R_{(q-1)(q-1)} & \text{if } i = q-1 \end{cases}$$

allowing to easily compute the following values of interest:

$$\begin{array}{lcccl} i < p & \Rightarrow & R_{pi} = 0, R_{(q-1)i} = 0 & \Rightarrow & x_i = 0 \\ i = p & \Rightarrow & R_{pp} = 1, R_{(q-1)p} = 0 & \Rightarrow & x_p = \rho \\ i = q-1 & \Rightarrow & R_{pp} = 1, R_{(q-1)p} = 0, R_{(q-1)(q-1)} = 1 & \Rightarrow & x_{q-1} = 1 \end{array}$$

Therefore:

$$\mathbf{R_{p,q-1}} = \begin{pmatrix} & & & 0 & \dots & & 0 & \dots & 0 \\ & \mathbf{I}_{(p-1)} & & 0 & \dots & & 0 & \dots & 0 \\ & & & \vdots & & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & \rho & \# & \# \\ \vdots & \vdots & & 0 & F_{1,1} & \dots & x_{p+1} & \dots & F_{1,M-p} \\ & & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & \rho & x_{p+1} & \dots & 1 & \dots & x_M \\ \vdots & \vdots & & \# & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & \# & F_{1,M-p} & \dots & x_M & \dots & F_{M-p,M-p} \end{pmatrix}$$

ACCEPTED MANUSCRIPT

$R_{p,q-1}$ is positive definite because it is a quadratic form of the positive definite matrix $R_{p,q}$. Furthermore, since every element on its main diagonal is equal to 1, $R_{p,q-1}$ is a correlation matrix.

Among others, pre- and post-multiplication of the matrix $R_{p,q}$ by the matrix $K_{p,q-1}$ and its transpose have the effect of filling the zero element in position $(p, q-1)$ with $|\rho| < 1$.

Lemma A.4 *Given any sequence of K matrices, their product $K_{1,2} \cdot K_{1,3} \cdot ... \cdot K_{M-1,M} \cdot K'_{M-1,M} \cdot ... \cdot K'_{1,3} \cdot K'_{1,2}$ is a correlation matrix.*

Proof:

Consider performing the matrix multiplication starting from the central term - $K_{M-1,M} \cdot K'_{M-1,M}$ and moving out by pre- and post-multiplying the successive matrices, then at every step the result is a correlation matrix of the type $R_{p,q}$. Thus, the result is an $R_{1,2}$ correlation matrix or simply a correlation matrix.

B Properties of the ACC Model

B.1 Assumptions

$$\begin{aligned}
&\textbf{A1:} \quad \mathbb{E}[d_t|\mathcal{I}_{t-1}] = \mathbb{E}[d_t] = \mathbb{E}[d_1] \qquad \forall t \\
&\textbf{A2:} \quad \mathbb{COV}[d_t, \eta_t] = 0 \\
&\textbf{A3:} \quad \mathbb{COV}[d_t, \eta_t^2] = 0 \\
&\textbf{A4:} \quad \mathbb{E}[\eta_t|\mathcal{I}_{t-1}] = \mathbb{E}[\eta_t] = \mathbb{E}[\eta_1] \qquad \forall t
\end{aligned}$$

B.2 Covariance Stationarity

In order to determine the conditions under which the ACC(1,1) is covariance stationary, it is convenient to rewrite it as an ARMA process for the Fisher transformation of the realized correlations ψ_t. These can be defined to be equal to the Fisher transformation of the *true* underlying correlation χ_t plus the error term η_t:

$$\psi_t = \omega + (\delta + \theta + \beta d_{t-1})\psi_{t-1} - \delta\eta_{t-1} + \eta_t$$

Taking expectations on both sides and using assumptions A1 and A2:

$$\mathbb{E}[\psi_t] = \omega + (\delta + \theta + \beta\mathbb{E}[d])\mathbb{E}[\psi_{t-1}] + (1 - \delta)\mathbb{E}[\eta]$$

ACCEPTED MANUSCRIPT

Thus, the process is mean stationary if and only if:

$$\left|\delta + \theta + \beta\mathbb{E}[d]\right| < 1$$

In order to analyze the variance of the process, it is convenient to introduce the more compact notation $a_{t-1} \equiv \delta + \theta + \beta d_{t-1})$ and $\epsilon_t \equiv -\delta\eta_{t-1} + \eta_t$:

$$\psi_t = \omega + a_{t-1}\psi_{t-1} + \epsilon_t$$

Hence, the variance of ψ_t is equal to:

$$\mathbb{V}[\psi_t] = \mathbb{V}[a_{t-1}\psi_{t-1}] + 2\mathbb{COV}[a_{t-1}\psi_{t-1}, \epsilon_t] + \mathbb{V}[\epsilon_t]$$

Using A1-A4, the first term on the right hand side is equal to:

$$\begin{aligned}\mathbb{E}[a_{t-1}^2\psi_{t-1}^2] - \mathbb{E}[a_{t-1}\psi_{t-1}]^2 &= \mathbb{E}[a_{t-1}^2]\mathbb{E}[\psi_{t-1}^2] - \mathbb{E}[a_{t-1}]^2\mathbb{E}[\psi_{t-1}]^2 \\ &= \mathbb{E}[a_{t-1}^2]\mathbb{V}[\psi_{t-1}] + \mathbb{E}[\psi_{t-1}]^2\mathbb{V}[a_{t-1}]\end{aligned}$$

while the second term is equal to:

$$\begin{aligned}2\mathbb{E}[a_{t-1}\psi_{t-1}\epsilon_t] - 2\mathbb{E}[a_{t-1}\psi_{t-1}]\mathbb{E}[\epsilon_t] &= 2\mathbb{E}[a_{t-1}]\mathbb{COV}[\psi_{t-1}, \epsilon_t] \\ &= 2\mathbb{E}[a_{t-1}]\mathbb{COV}[\psi_{t-1}, -\delta\eta_{t-1} + \eta_t] \\ &= 2\mathbb{E}[a_{t-1}]\mathbb{COV}[\psi_{t-1}, -\delta\eta_{t-1}] \\ &= 2\mathbb{E}[a_{t-1}]\mathbb{COV}[\eta_{t-1}, -\delta\eta_{t-1}] \\ &= -2\delta\mathbb{E}[a_{t-1}]\mathbb{V}[\eta]\end{aligned}$$

Therefore, the expression for the variance of ψ_t takes the following form:

$$\mathbb{V}[\psi_t] = \mathbb{E}[a_{t-1}^2]\mathbb{V}[\psi_{t-1}] + \mathbb{E}[\psi_{t-1}]^2\mathbb{V}[a_{t-1}] + \Big(1 + \delta^2 - 2\delta\mathbb{E}[a_{t-1}]\Big)\mathbb{V}[\eta]$$

It follows, that the process is variance stationary if and only if $\mathbb{E}[a_{t-1}^2] < 1$:

$$\begin{aligned}\mathbb{E}[a_{t-1}^2] &= \mathbb{E}[(\delta + \theta)^2 + 2(\delta + \theta)\beta d_{t-1} + \beta^2 d_{t-1}^2] \\ &= (\delta + \theta)^2 + 2(\delta + \theta)\beta\mathbb{E}[d] + \beta^2\mathbb{E}[d] \\ &= (\delta + \theta + \beta\mathbb{E}[d])^2 + \beta^2\mathbb{E}[d](1 - \mathbb{E}[d])\end{aligned}$$

where $\mathbb{E}[d^2] = \mathbb{E}[d]$ is due to the fact that d is a binary random variable. The model is then variance stationary when its parameters satisfy:

$$(\delta + \theta + \beta\mathbb{E}[d])^2 + \beta^2\mathbb{E}[d](1 - \mathbb{E}[d]) < 1$$

Notice how variance stationarity implies mean stationarity. Furthermore, it can be shown that the condition for mean stationarity implies stationarity of the auto-covariances.

ACCEPTED MANUSCRIPT

B.3 Forecasts

The unconditional expectation of the stationary process χ_t is equal to:

$$\mathbb{E}[\chi] = \omega + \delta\mathbb{E}[\chi] + \theta\mathbb{E}[\psi] + \beta\mathbb{E}[d\psi]$$

which implies that:

$$\omega = (1-\delta)\mathbb{E}[\chi] - \theta\mathbb{E}[\psi] - \beta\mathbb{E}[d\psi]$$

Taking the expectation of χ_{t+1}, conditional on the information set $\mathcal{I}_{t-1}$, and using A1 and A4:

$$\begin{aligned}
\mathbb{E}[\chi_{t+1}|\mathcal{I}_{t-1}] &= \omega + \delta\chi_t + \mathbb{E}\Big[(\theta+\beta d_t)(\chi_t+\eta_t)\Big] \\
&= \omega + (\delta+\theta+\beta\mathbb{E}[d])\chi_t + \theta\mathbb{E}[\eta_t] + \beta\mathbb{E}[d_t\eta_t] \\
&= \omega + (\delta+\theta+\beta\mathbb{E}[d])\chi_t + \theta\mathbb{E}[\psi_t-\chi_t] + \beta\mathbb{E}[d_t\psi_t - d_t\chi_t] \\
&= \omega + (\delta+\theta+\beta\mathbb{E}[d])\chi_t + \theta\mathbb{E}[\psi] - \theta\mathbb{E}[\chi] + \beta\mathbb{E}[d\psi] - \mathbb{E}[d]\mathbb{E}[\chi]
\end{aligned}$$

and substituting the expression for ω yields:

$$\mathbb{E}[\chi_{t+1}|\mathcal{I}_{t-1}] = (1-\delta-\theta-\beta\mathbb{E}[d])\mathbb{E}[\chi] + (\delta+\theta+\beta\mathbb{E}[d])\chi_t$$

In order for the forecasts[13] to be mean reverting it must be that:

$$\Big|\delta+\theta+\beta\mathbb{E}[d]\Big| < 1$$

which is true whenever the conditions for covariance stationarity are satisfied.

Acknowledgements

I thank John Geweke, the associate editor and two anonymous referees for their insightful comments and detailed suggestions which have greatly improved the paper. I am grateful to Tim Bollerslev, George Tauchen, A. Ronald Gallant and Barbara Rossi for their guidance, constant moral support and valuable comments. I also thank all the participants to the Duke Financial Econometrics Group for their helpful comments and suggestions.

[13] For a discussion on multi-step ahead forecasts for non-linear processes based on approximations: Engle and Sheppard (2002).

ACCEPTED MANUSCRIPT

References

Alexander, C. (2001), "A Primer on the Orthogonal GARCH Model.", ISMA Center, Mimeo.

Alexander, C. and A. Chibumba (1997), "Multivariate Orthogonal Factor GARCH.", University of Sussex, Mimeo.

Andersen, T.G. and T. Bollerslev (1998), "Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts." *International Economic Review* 39, 885-905.

Bauwens, L., S. Laurent and J.V.K. Rombouts (2006), "Multivariate GARCH Models: A survey." *Journal of Applied Econometrics* 21/1, 79-109.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity." *Journal of Econometrics* 31, 307-327.

Bollerslev, T. (1987), "A conditional Heteroscedastic Time Series Model for Speculative Prices and Rates of Return." *Review of Economics and Statistics* 69, 542-547.

Bollerslev, T. (1990), "Modeling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH model." *Review of Economics and Statistics* 72, 498-595.

Bollerslev, T. and R. Engle (1986), "Modeling the Persistence of Conditional Variances" *Econometric Reviews* 5, 1-50.

Bollerslev, T., R. Engle, and J. Wooldridge (1988), "A Capital Asset Pricing Model with Time Varying Covariances" *Journal of Political Economy* 96, 116-131.

Brooks, C., S.P. Burke, and G. Persand (2003), "Multivariate GARCH Models: Software Choice and Estimation Issues." *Journal of Applied Econometrics* 18, 725-734.

Cappiello, L., R. Engle, and K. Sheppard (2006), "Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns." *Journal of Financial Econometrics* 4(4), 537-572.

ACCEPTED MANUSCRIPT

Chen, X. and Y. Fan (2006), "Estimation and Model Selection of Semiparametric Copula-Based Multivariate Dynamic Models Under Copula Misspecification." *Journal of Econometrics* 135, 125-154.

Comte, F. and O. Lieberman (2003), "Asymptotic Theory for Multivariate GARCH Processes." *Journal of Multivariate Analysis* 84(1), 61-84.

Diebold, F.X. and M. Nerlove (1989), "The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor Arch Model." *Journal of Applied Econometrics* 1, 1-21.

Ding, Z., R. Engle, and C.W.J. Granger (1993), "A Long Memory Property of Stock Market Returns and a New Model." *Journal of Empirical Finance* 1, 83-106.

Engle, R. (2002), "Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models" *Journal of Business and Economic Statistics* vol. 17, n. 5.

Engle, R. and F. Kroner (1995), "Multivariate Simultaneous Generalized ARCH." *Econometric Theory* 11, 122-150.

Engle, R. and J. Mezrich (1996), "GARCH for groups" *RISK* 9, 36-40.

Engle, R., V. Ng, and M. Rothschild (1990), "Asset Pricing with a Factor-ARCH Covariance Structure: Empirical Estimates for Treasury Bills." *Journal of Econometrics* 45, 213-238.

Engle, R. and K. Sheppard (2002), "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH" *Stern Finance Working Paper Series* FIN-01-027.

Gallant, A.R. and G. Tauchen (1989), "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications." *Econometrica* 57, 1091-1120.

Gallant, A.R. and G. Tauchen (2002), "SNP: A Program for Nonparametric Time Series Analysis", User's Guide Version 8.8

Geweke, J. (1986), "Modeling the persistence in conditional variances: A comment" *Econometric Review* 5, 57-61.

ACCEPTED MANUSCRIPT

Glosten, L., R. Jagannathan, and D. Runkle (1993), "On the Relationship Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *Journal of Finance*, 48, 1779-1801.

Gouriéroux, C., A. Monfort and A. Trognon (1984), "Pseudo-Maximum Likelihood Methods: Theory." *Econometrica* 52, 681-700.

Hafner, C.M. and A. Preminger (2009), "Asymptotic Theory for a Factor GARCH Model." *Econometric Theory* 25(2), 336-363.

Harvey, A., E. Ruiz and E. Sentana (1992), "Unobserved component time series models with ARCH disturbances" *Journal of Econometrics* 52, 129-157.

Hentschel, L. (1995), "All in the family: Nesting symmetric and asymmetric GARCH models" *Journal of Financial Economics*, 39, 71-104.

Higgins, M.L. and A.K. Bera (1992), "A Class of Nonlinear ARCH Models." *International Economic Review* 33, 137-158.

Jeantheau, T. (1998), "Strong Consistency of Estimators for Multivariate ARCH Models." *Econometric Theory* 14, 70-86.

Jondeau, E. and M. Rockinger (2006), "The copula-GARCH model of conditional dependencies: An international stock-market application." *Journal of International Money and Finance* 25(5), 827-853.

Kawakatsu, H. (2006), "Matrix Exponential GARCH" *Journal of Econometrics* 134(1), 95-128.

Ledoit, O., P. Santa-Clara, and M. Wolf (2003), "Flexible Multivariate GARCH Modeling With an Application to International Stock Markets." *The Review of Economics and Statistics* vol. 8, issue 3, 735-747.

Ling, S. and M. McAleer (2003), "Asymptotic Theory for a Vector ARMA-GARCH Model" *Econometric Theory* 19, 280-310.

Nelson, D.B. (1991), "Conditional Heteroscedasticity in Asset Returns: A New Approach." *Econometrica* 59, 347-370.

Pantula, S.G. (1986), "Modeling the Persistence of Conditional Variances: A Comment" *Econometric Reviews* 5, 71-74.

ACCEPTED MANUSCRIPT

Patton, A.J. (2006), "Modeling Asymmetric Exchange Rate Dependence." *International Economic Review* 47(2), 527-556.

Patton, A.J. (2008), "Volatility Forecast Evaluation and Comparison Using Imperfect Volatility Proxies", forthcoming in *Journal of Econometrics.*

Sentana, E. (1995), "Quadratic ARCH models" *Review of Economic Studies* 62(4), 639-661.

Tse, Y. and A. Tsui (2002), "A Multivariate GARCH Model with Time-Varying Correlations." *Journal of Business and 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.

Zakoian, J.M. (1994), "Threshold Heteroscedastic Models." *Journal of Economic Dynamics and Control* 18, 931-955.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 1: Symbols of Included Stocks

ADBE	ALTR	APCC	AMGN	AAPL	AMAT	BBBY
BIIB	BMET	CDWC	CEPH	CHIR	CTAS	CSCO
CMCSA	CPWR	CMVT	COST	DELL	XRAY	ERTS
EXPD	ESRX	FAST	FHCC	FISV	FLEX	GNTX
GENZ	GILD	IACI	INTC	INTU	JDSU	KLAC
LRCX	LNCR	LLTC	MXIM	MEDI	MERQ	MCHP
MSFT	MOLX	NXTL	NVLS	ORCL	PCAR	PDCO
PTEN	PAYX	PSFT	PETM	QLGC	QCOM	ROST
SANM	SIAL	SSCC	SPLS	SBUX	SUNW	SYMC
SNPS	TLAB	TEVA	VRTS	WFMI	XLNX	

Table 2: Monte Carlo Results: true parameter values λ_0, M.C. average of estimated parameter values λ_S, M.C. average of simulated theoretical parameter standard error σ, M.C. parameter standard error σ_S and 10%, 5% and 1% rejection frequencies assuming that the estimated parameters have a gaussian distribution.

	λ_0	$\overline{\lambda}_S$	σ	σ_S	10%	5%	1%
$\omega_{1,2}$	0.035	0.038	$1.12 \cdot 10^{-2}$	$1.14 \cdot 10^{-2}$	8.5%	5.3%	2.5%
$\delta_{1,2}$	0.900	0.895	$2.04 \cdot 10^{-2}$	$2.05 \cdot 10^{-2}$	9.7%	5.5%	1.9%
$\theta_{1,2}$	0.050	0.051	$9.60 \cdot 10^{-3}$	$9.40 \cdot 10^{-3}$	9.5%	4.6%	1.0%
$\omega_{1,3}$	0.004	0.006	$5.49 \cdot 10^{-3}$	$5.86 \cdot 10^{-3}$	5.3%	3.9%	2.3%
$\delta_{1,3}$	0.980	0.973	$1.74 \cdot 10^{-2}$	$1.88 \cdot 10^{-2}$	6.6%	4.8%	2.9%
$\theta_{1,3}$	0.010	0.012	$6.23 \cdot 10^{-3}$	$5.95 \cdot 10^{-3}$	8.6%	5.3%	2.0%
$\omega_{2,3}$	0.002	0.002	$1.51 \cdot 10^{-3}$	$1.35 \cdot 10^{-3}$	5.5%	3.6%	1.7%
$\delta_{2,3}$	0.950	0.947	$1.50 \cdot 10^{-2}$	$1.68 \cdot 10^{-2}$	9.6%	5.7%	2.6%
$\theta_{2,3}$	0.030	0.029	$7.80 \cdot 10^{-3}$	$7.82 \cdot 10^{-3}$	10.6%	5.4%	1.3%

ACCEPTED MANUSCRIPT

Table 3: Results on the 69-Dimensional Model: In Sample Predictions. MSE and percentage improvement w.r.t. CCC; MAD in brackets and percentage improvement w.r.t. CCC.

Days	CCC	DCC		SCC(↓)	
20	149.800	143.126	-4.46%	127.851	-14.65%
	[482.677]	[471.557]	-2.30%	[443.174]	-8.18%
40	93.505	87.573	-6.34%	73.737	-21.14%
	[379.799]	[367.437]	-3.25%	[334.809]	-11.85%
60	72.796	67.356	-7.47%	54.315	-25.39%
	[334.479]	[321.713]	-3.82%	[286.843]	-14.24%
80	62.013	56.955	-8.16%	44.444	-28.33%
	[308.321]	[295.566]	-4.14%	[259.299]	-15.89%
100	55.142	50.400	-8.60%	38.300	-30.54%
	[290.266]	[277.661]	-4.34%	[240.553]	-17.13%
120	50.246	45.779	-8.89%	33.999	-32.33%
	[276.624]	[264.186]	-4.50%	[226.471]	-18.13%
140	46.719	42.508	-9.01%	31.080	-33.47%
	[266.354]	[254.236]	-4.55%	[216.387]	-18.76%
160	44.020	40.046	-9.03%	28.958	-34.21%
	[258.168]	[246.417]	-4.55%	[208.718]	-19.15%
180	41.836	38.080	-8.98%	27.263	-34.83%
	[251.317]	[239.922]	-4.53%	[202.279]	-19.51%
200	40.036	36.493	-8.85%	26.013	-35.03%
	[245.502]	[234.524]	-4.47%	[197.408]	-19.59%
220	38.561	35.229	-8.64%	25.103	-34.90%
	[240.694]	[230.173]	-4.37%	[193.736]	-19.51%
240	37.342	34.221	-8.36%	24.464	-34.49%
	[236.697]	[226.704]	-4.22%	[191.124]	-19.25%

ACCEPTED MANUSCRIPT

Table 4: Results on the 69-Dimensional Model: In Sample Predictions for Permutations. MSE and percentage improvement w.r.t. CCC; MAD in brackets and percentage improvement w.r.t. CCC.

Days	120		160		200	
SCC_SIC	35.272	-29.80%	29.863	-32.16%	26.576	-33.62%
	[231.010]	-16.49%	[212.258]	-17.78%	[199.834]	-18.60%
SCC(A)	34.899	-30.54%	30.147	-31.52%	27.385	-31.60%
	[230.030]	-16.84%	[213.486]	-17.31%	[203.005]	-17.31%
SCC(Z)	35.539	-29.27%	30.566	-30.56%	27.656	-30.92%
	[231.702]	-16.24%	[214.662]	-16.85%	[203.839]	-16.97%
SCC(1)	35.068	-30.21%	30.367	-31.01%	27.643	-30.95%
	[230.564]	-16.65%	[214.125]	-17.06%	[203.674]	-17.04%
SCC(2)	35.408	-29.53%	30.659	-30.35%	27.916	-30.27%
	[231.866]	-16.18%	[215.090]	-16.69%	[204.385]	-16.75%
SCC(3)	34.671	-31.00%	29.874	-32.14%	27.108	-32.29%
	[228.913]	-17.25%	[211.935]	-17.91%	[201.308]	-18.00%
SCC(4)	34.439	-31.46%	29.554	-32.86%	26.723	-33.25%
	[228.612]	-17.36%	[211.602]	-18.04%	[200.963]	-18.14%
SCC(5)	35.544	-29.26%	30.967	-29.65%	28.334	-29.23%
	[231.760]	-16.22%	[215.763]	-16.43%	[205.734]	-16.20%
SCC(6)	34.521	-31.30%	29.789	-32.33%	27.025	-32.50%
	[228.370]	-17.44%	[211.740]	-17.98%	[200.995]	-18.13%
SCC(7)	34.905	-30.53%	30.121	-31.57%	27.377	-31.62%
	[229.650]	-16.98%	[213.063]	-17.47%	[202.662]	-17.45%
SCC(8)	34.627	-31.09%	29.986	-31.88%	27.315	-31.77%
	[229.051]	-17.20%	[212.596]	-17.65%	[202.267]	-17.61%

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 5: Likelihood-Based Model Evaluations: Likelihood Ratio Test and Schwartz Information Criterion.

	Parameters	Likelihood	LRT p-value CCC	LRT p-value SCC_SIC	SIC
CCC	2346	$-5.6910 \cdot 10^4$	-	-	$1.3219 \cdot 10^5$
DCC	2349	$-5.6874 \cdot 10^4$	$< 1.00 \cdot 10^{-14}$	-	$1.3214 \cdot 10^5$
SCC_SIC	2614	$-5.4533 \cdot 10^4$	$< 1.00 \cdot 10^{-14}$	-	$1.2954 \cdot 10^5$
SCC(↓)	11720	$-4.8107 \cdot 10^4$	$< 1.00 \cdot 10^{-14}$	$< 1.00 \cdot 10^{-14}$	$1.8801 \cdot 10^5$