

National Centre of Competence in Research
Financial Valuation and Risk Management

Working Paper No. 386

A general multivariate threshold GARCH model with dynamic conditional correlations

Francesco Audrino

Fabio Trojani

First version: April 2007
Current version: April 2007

This research has been carried out within the NCCR FINRISK project on
“New Methods in Theoretical and Empirical Asset Pricing”

A general multivariate threshold GARCH model with dynamic conditional correlations

Francesco Audrino^{a,b*} and Fabio Trojani^{b,c}

^aInstitute of Finance, University of Lugano

^bDepartment of Economics, University of St. Gallen

^cSwiss Institute of Banking and Finance, University of St. Gallen

Revised: April 2007

Abstract

We propose a new multivariate GARCH model with Dynamic Conditional Correlations that extends previous models by admitting multivariate thresholds in conditional volatilities and correlations. The model estimation is feasible in large dimensions and the positive definiteness of the conditional covariance matrix is easily ensured by the structure of the model. Thresholds in conditional volatilities and correlations are estimated from the data, together with all other model parameters. We study the performance of our model in three distinct applications to US stock and bond market data. Even if the conditional volatility functions of stock returns exhibit pronounced GARCH and threshold features, their conditional correlation dynamics depends on a very simple threshold structure with no local GARCH features. We obtain a similar result for the conditional correlations between government and corporate bond returns. On the contrary, we find both threshold and GARCH structures in the conditional correlations between stock and government bond returns. In all applications, our model improves significantly the in-sample and out-of-sample forecasting power for future conditional correlations with respect to other relevant multivariate GARCH models.

*We thank the editor, Torben Andersen, the associate editor and two anonymous referees for many valuable comments. Corresponding address: Institute of Finance, University of Lugano, Via Buffi 13, CH-6900 Lugano, e-mail: francesco.audrino@lu.unisi.ch, Phone: 0041 58 6664789. E-mail for Fabio Trojani: fabio.trojani@unisg.ch. Financial support by the Swiss National Science Foundation (grants 100012-103781, 100012-105745 and NCCR FINRISK) and by the Foundation for Research and Development of the University of Lugano is gratefully acknowledged.

Keywords: Multivariate GARCH models; Dynamic conditional correlations; Tree-structured GARCH models.

JEL codes: C12, C13, C51, C53, C61

1 Introduction

In this paper, we present a new multivariate GARCH model with Dynamic Conditional Correlations (DCC) that extends previous approaches by admitting multivariate thresholds in the conditional volatilities and correlations of multivariate time series. This extension allows us to account for rich asymmetric effects and dependencies of conditional volatilities and correlations, as they are often encountered - for instance - in financial real data applications. Similarly to the classical Engle (2002) DCC-model, our model estimation is numerically feasible in large dimensions. Moreover, the positive definiteness of the conditional covariance matrix is ensured in a natural way by the structure of the model. Finally, thresholds in volatilities and correlations of our model are not fixed *ex ante*, but are estimated from the data, together with all other parameters in the model.

To define the threshold function in our model, we extend the tree-structured state space partition in Audrino and Bühlmann (2001) to a setting with multivariate thresholds in both volatilities and correlations. As shown in Audrino and Trojani (2006) and Audrino (2006) the tree-structured threshold construction can incorporate in a parsimonious way a potentially large number of multivariate regimes in univariate settings. In this paper, we study a multivariate model with a potentially high number of tree-structured thresholds in volatilities and correlations, as well as a feasible estimation strategy that can be applied to estimate the model also in large dimensional applications. The threshold construction is obtained using a binary tree, in which each terminal node defines a local GARCH-type dynamics for volatilities and correlations over a partition cell of the multivariate state space. The estimation is performed by a simple two-step procedure that estimates the number and the structure of the underlying thresholds together with the parameters of the local GARCH dynamics for volatilities and correlations. The optimal threshold structure is identified by solving a high dimensional model selection problem based on the Schwarz Bayesian Information Criterion (BIC).

We estimate our model in three distinct applications to US stock and bond market data and focus on its explanatory power for future conditional correlations with respect to a set of relevant competing models in the literature. These models include Engle (2002) DCC-model, Ledoit et al. (2003) flexible multivariate GARCH model and Pelletier (2006) Regime Switching Dynamic Correlations (RSDC) model. The first and the third of these models can be estimated, like our one, by a two step procedure that separates the estimation of the conditional volatility and correlation dynamics. In order to measure, where possible, the additional forecasting power

for correlations, we estimate these models using a set of univariate tree structured GARCH-dynamics for volatility identical to those in our model. The flexible multivariate GARCH model cannot be estimated by a two step estimation procedure. Therefore, in this case the estimated volatility dynamics are different from those estimated for our model. The major difference between our model and the other ones arises, however, in the way how we specify the correlation dynamics. The DCC and the flexible multivariate GARCH models are single-regime models for correlations. The RSDC model specifies a very simple regime structure for conditional correlations. Our setting can account in a parsimonious way for GARCH-type dynamics and complex threshold structures in conditional correlations, without fixing from the beginning the structure and the number of the thresholds in the model.

Using our tree-structured GARCH-DCC model, we study empirically the relative importance of GARCH and threshold effects in the conditional correlation dynamics of US stock and bond returns. Even if the conditional volatility functions of US stock returns exhibit GARCH and threshold features, we find that their conditional correlations depend on a simple threshold structure with no local GARCH features. A similar result is derived also for the conditional correlations between US government and corporate bond returns. On the contrary, we find rich threshold and GARCH structures in the conditional correlations between stock and government bond returns. In these cases, past equity and bond returns impact on the arising multivariate correlation thresholds. In all applications, the tree-structured partition of the state space improves significantly the in-sample and out-of-sample forecasting power for future conditional correlations with respect to the other multivariate GARCH models analyzed in the paper.

In Section 2 we present our tree-structured GARCH-DCC model and the two-step estimation procedure that can be applied to estimate it. Section 4 presents the empirical findings from our real data applications to the estimation of conditional (volatility and) correlation dynamics for the US stock and bond markets. Section 5 summarizes the main results and concludes.

2 The model

We consider a multivariate stochastic process $(\mathbf{X}_t)_{t \in \mathbb{Z}}$ with values in \mathbb{R}^d :

$$\mathbf{X}_t = D_t \epsilon_t, \tag{2.1}$$

where $D_t := \text{diag}[\sigma_{1,t}, \dots, \sigma_{d,t}]$ and $\sigma_{i,t}$ is the conditional standard deviation of the i -th component of \mathbf{X}_t at time $t - 1$. $(\epsilon_t)_{t \in \mathbb{Z}}$ is a zero-mean process in \mathbb{R}^d with components having a

unit conditional standard deviation by construction.¹ The conditional covariance matrix of ϵ_t at time $t - 1$ is denoted by R_t . Therefore, we obtain the following standard factorization of the conditional covariance matrix of \mathbf{X}_t :

$$\text{Cov}_{t-1}(\mathbf{X}_t) = D_t R_t D_t, \quad (2.2)$$

Our tree-structured DCC-GARCH model parameterizes the conditional volatility matrix D_t and the conditional correlation matrix R_t by means of two parametric threshold functions. Each diagonal element of D_t is modeled as a univariate tree-structured threshold GARCH(1,1)-model, as in Audrino and Bühlmann (2001) and Audrino and Trojani (2006). The conditional correlation matrix R_t is modeled according to a threshold DCC-type model described in more details below.

2.1 Tree-structured model for D_t

Let $X_{t,j}$ be the j -th component of \mathbf{X}_t . In principle, the thresholds in the volatility dynamics of $X_{t,j}$ can depend on all components of \mathbf{X}_{t-1} . For simplicity of exposition, let us assume that they are functions of $(X_{t-1,1}, X_{t-1,j})$. Let $\mathcal{P}_j = \{\mathcal{R}_{1,j}, \dots, \mathcal{R}_{k_j,j}\}$ be a partition of the state space $G := \mathbb{R}^2 \times \mathbb{R}^+$ of $(X_{t-1,1}, X_{t-1,j}, \sigma_{t-1,j}^2)$:

$$\mathcal{P}_j = \{\mathcal{R}_{1,j}, \dots, \mathcal{R}_{k_j,j}\}, \cup_{s=1}^{k_j} \mathcal{R}_{s,j} = G, \mathcal{R}_{i,j} \cap \mathcal{R}_{s,j} = \emptyset \ (i \neq s).$$

Given a partition cell $\mathcal{R}_{i,j}$, we specify the local conditional variance dynamics of $X_{t,j}$ on $\mathcal{R}_{i,j}$ as a GARCH(1,1) model. Therefore, threshold function $\sigma_{t,j}^2$ takes the form

$$\sigma_{t,j}^2 = \sum_{i=1}^{k_j} (\alpha_{ij} + \beta_{ij} X_{t-1,j}^2 + \gamma_{ij} \sigma_{t-1,j}^2) I_{[(X_{t-1,1}, X_{t-1,j}, \sigma_{t-1,j}^2) \in \mathcal{R}_{i,j}]}, \quad (2.3)$$

where $I_{[\cdot]}$ is the indicator function and $\theta_{1,j}$ is the parameter vector:

$$\theta_{1,j} = \{\alpha_{ij}, \beta_{ij}, \gamma_{ij} ; i = 1, \dots, k_j\}.$$

To specify completely the conditional variance function (2.3), we have to define the class of partitions \mathcal{P}_j that are admissible in our tree-structured model. Essentially, the only restriction we impose is that \mathcal{P}_j is composed by rectangular cells $\mathcal{R}_{i,j}$, $i = 1, \dots, k_j$, delimited by a set of multivariate thresholds for $(X_{t-1,1}, X_{t-1,j}, \sigma_{t-1,j}^2)$. For example, in a model with three regimes and two thresholds, the partitioning cells $\mathcal{R}_{1,j}, \mathcal{R}_{2,j}, \mathcal{R}_{3,j}$ could be of the form:

$$\begin{aligned}
\mathcal{R}_{1,j} &= \{X_{t-1,j} \leq d_1\}, \\
\mathcal{R}_{2,j} &= \{X_{t-1,j} > d_1 \text{ and } X_{t-1,1} \leq d_2\}, \\
\mathcal{R}_{3,j} &= \{X_{t-1,j} > d_1 \text{ and } X_{t-1,1} > d_2\},
\end{aligned}$$

where the parameters d_1, d_2 define the two multivariate thresholds in the model. In this case, $\mathcal{R}_{1,j}$ is associated with a regime of low conditioning values $X_{t-1,j}$. $\mathcal{R}_{2,j}$ corresponds to a regime with higher conditioning values of $X_{t-1,j}$, but low values of $X_{t-1,1}$. Finally, $\mathcal{R}_{3,j}$ implies a regime in which both conditioning values $X_{t-1,1}$ and $X_{t-1,j}$ are large. The multivariate threshold function in our model is defined by means of a binary tree \mathcal{T}_j in which every terminal node represents a particular cell $\mathcal{R}_{i,j}$. Details on the construction and the interpretation of binary trees for our applied examples are provided in Audrino and Trojani (2006). For each component $X_{t,j}$, estimation of (2.3) is achieved by a high dimensional model selection problem that determines the optimal number and the structure of the relevant thresholds (and hence the partition cells) in \mathcal{P}_j . Details on this estimation procedure for univariate tree structured GARCH(1,1) models are given in Audrino and Bühlmann (2001) and Audrino and Trojani (2006), Section 2.3.

2.2 Tree structured model for R_t

Let

$$\epsilon_t = D_t(\theta_1)^{-1} \mathbf{X}_t,$$

so that

$$R_t = \text{Corr}_{t-1}(\mathbf{X}_t) = \text{Cov}_{t-1}(\epsilon_t).$$

We model R_t by means of a tree-structured model, in which conditional correlations satisfy an Engle (2002)-type local DCC model across several multivariate thresholds. In order to keep the model tractable, we assume that thresholds in the R_t dynamics depend on ϵ_{t-1} only via the average

$$\bar{\rho}_{t-1} = \frac{1}{d(d-1)} \sum_{u \neq v} \epsilon_{t-1,u} \epsilon_{t-1,v}$$

of the cross products of the component of ϵ_{t-1} . Intuitively, this choice allows us to account for asymmetric effects in conditional correlations, as a function of particular lagged process realizations \mathbf{X}_{t-1} and specific movements in average lagged conditional correlations shocks $\bar{\rho}_{t-1}$.

To define the parametric threshold function R_t in our model, let $\mathcal{P} = \{\tilde{\mathcal{R}}_1, \dots, \tilde{\mathcal{R}}_w\}$ be a partition of the state space $\tilde{G} := \mathbb{R}^{d+1}$ of $(\mathbf{X}_{t-1}, \bar{\rho}_{t-1})$. We consider the following family of functional forms for R_t :

$$R_t = \sum_{i=1}^w c_i R_{it} I_{\{(\mathbf{x}_{t-1}, \bar{\rho}_{t-1}) \in \tilde{\mathcal{R}}_i\}} + \left(1 - \sum_{i=1}^w c_i I_{\{(\mathbf{x}_{t-1}, \bar{\rho}_{t-1}) \in \tilde{\mathcal{R}}_i\}}\right) Id_n \quad (2.4)$$

where $c_1, \dots, c_n \in [0, 1]$, Id_n is the d -dimensional identity matrix and the parametric processes for R_{it} , $i = 1, \dots, w$, are given by:

$$R_{it} = \text{diag}[Q_{it}]^{-1/2} Q_{it} \text{diag}[Q_{it}]^{-1/2} \quad (2.5)$$

with

$$Q_{it} = (1 - \phi_i - \lambda_i) \bar{Q} + \phi_i \epsilon_{t-1} \epsilon'_{t-1} + \lambda_i Q_{it-1}, \quad (2.6)$$

parameters $\phi_i, \lambda_i \geq 0$ such that $\phi_i + \lambda_i < 1$ for all $i = 1, \dots, w$, and \bar{Q} is, like in the classical Engle (2002) DCC model, the unconditional covariance matrix of the residuals ϵ_t . Given a fixed partition \mathcal{P} , the parameter vector

$$\theta_2 = \{c_i, \phi_i, \lambda_i, \text{vech}(\bar{Q})\}; \quad i = 1, \dots, w. \quad (2.7)$$

completely parameterizes the threshold function defining the conditional correlation function (2.4).

Since for any $i = 1, \dots, w$, the local model for Q_{it} satisfies a Engle (2002) DCC-type dynamics, positive definiteness of the resulting threshold model for R_t is easily implied by the model structure under the above conditions on the model parameters. When $\mathcal{P} = \{\tilde{G}\}$, i.e. the partition is trivial, we obtain Engle (2002) DCC model by setting $c_1 = \dots = c_w = 1$. Therefore, this model is nested in our one. Moreover, by setting $\phi_i = \lambda_i = 0$ for $i = 1, \dots, w$, we can write R_t as:

$$R_t = \sum_{i=1}^w c_i \bar{R} I_{\{(\mathbf{x}_{t-1}, \bar{\rho}_{t-1}) \in \tilde{\mathcal{R}}_i\}} + \left(1 - \sum_{i=1}^w c_i I_{\{(\mathbf{x}_{t-1}, \bar{\rho}_{t-1}) \in \tilde{\mathcal{R}}_i\}}\right) Id_n. \quad (2.8)$$

where \bar{R} is a fixed d -dimensional correlation matrix.² In this case, we obtain a piecewise constant correlation matrix defined by a multivariate threshold function over the partition \mathcal{P} . In contrast to the RSDC model in Pelletier (2006), this particular subcase of our model can account for a flexible description of multiple multivariate regimes in correlations, because the number and the structure of the regimes in the estimated model does not have to be fixed from

the beginning. Finally, when $\phi_i > 0$ or $\lambda_i > 0$ for $i = 1, \dots, w$ and \mathcal{P} is not a trivial partition, we obtain by setting $c_1 = \dots = c_w = 1$ a tree-structured DCC-model satisfying locally Engle's DCC-dynamics over the distinct partitioning cells $\tilde{\mathcal{R}}_i$.

As for the univariate tree-structured volatility dynamics of the last section, we need to define the class of admissible partitions \mathcal{P} for our correlation function. Again, the only restriction we put on \mathcal{P} is that it is composed by rectangular partition cells $\tilde{\mathcal{R}}_i$, $i = 1, \dots, w$. Consistently with our assumptions, these partition cells are delimited by a set of multivariate thresholds for $(\mathbf{X}_{t-1}, \bar{\rho}_{t-1})$. In order to construct such rectangular partition cells, we make use of a binary tree, in which every terminal node represents a cell $\tilde{\mathcal{R}}_i$. Estimation of the threshold function in the correlation dynamics (2.4) is achieved by a high dimensional model selection procedure that determines the optimal number and the structure of the relevant thresholds in the underlying partition. This model selection scheme is not computationally feasible if applied directly to the multivariate time series $(\epsilon_t \epsilon_t')_{t \in \mathbb{Z}}$. A natural way to reduce estimation complexity is to remark that the partition \mathcal{P} is identical to the one implied by a corresponding tree-structured univariate model for the time series $(\bar{\rho}_t)_{t \in \mathbb{Z}}$. Indeed, since $R_t = E_{t-1}(\epsilon_t \epsilon_t')$ it follows:

$$\begin{aligned} E_{t-1}(\bar{\rho}_t) &= \frac{1}{d(d-1)} \sum_{u \neq v} R_{it}^{uv} \\ &= \sum_{i=1}^w c_i \left(\frac{1}{d(d-1)} \sum_{u \neq v} R_{it}^{uv} \right) I_{[(\mathbf{X}_{t-1}, \bar{\rho}_{t-1}) \in \tilde{\mathcal{R}}_i]} \end{aligned} \quad (2.9)$$

where R_t^{uv} (R_{it}^{uv}) denotes the uv -th component of the matrix R_t (R_{it}). Therefore, the tree-structured model

$$\bar{\rho}_t = E_{t-1}(\bar{\rho}_t) + \eta_t, \quad (2.10)$$

where $(\eta_t)_{t \in \mathbb{Z}}$ is a martingale difference process and $E_{t-1}(\bar{\rho}_t)$ is given by equation (2.9), defines a univariate tree-structured process for $\bar{\rho}_t$ based on the same partition \mathcal{P} as in the correlation dynamics (2.4). It follows, that we can exploit the univariate model (2.10) to estimate the threshold structure in equation (2.4). In particular, we can develop a model selection procedure for selecting the optimal threshold structure in the correlation dynamics. The most simple dynamics arises in the piecewise constant case:

$$E_{t-1}(\bar{\rho}_t) = \sum_{i=1}^w c_i \left(\frac{1}{d(d-1)} \sum_{u \neq v} \bar{R}^{uv} \right) I_{[(\mathbf{X}_{t-1}, \bar{\rho}_{t-1}) \in \tilde{\mathcal{R}}_i]} \quad (2.11)$$

where \overline{R}^{uv} is the uv -component of the correlation matrix \overline{R} in the piecewise-constant dynamics (2.8). This piecewise constant function is the optimal one that has been estimated in our applications to the US equity market in Section 4.1 and to the US bond market in Section 4.3. More generally, for $c_1 = \dots = c_n = 1$ and $\lambda_i, \phi_i > 0$, $i = 1, \dots, w$, we can also encompass the univariate dynamics of $\overline{\rho}_t$ that are consistent with a tree-structured DCC model of the form (2.4) for correlations. This threshold structure is the one we estimate in our application of Section 4.1 when we model the correlation between Treasury bond and stock returns. Model selection across this class of potential threshold functions for $E_{t-1}[\overline{\rho}_t]$ is performed using the BIC-information criterion. Once the partition \mathcal{P} in (2.10) has been estimated, the parameter (2.7) of the multivariate correlation dynamics can be estimated using a multivariate conditional pseudo likelihood for ϵ_t , in which the selected partition \mathcal{P} is held fixed. The next section provides additional details on the estimation procedure used to estimate our tree structured DCC model.

3 Estimation of the tree-structured DCC model

Estimation of our tree-structured model is achieved in two steps. In the first step, an estimate of the volatility process D_t is obtained by performing d estimations of the univariate tree structured conditional volatility dynamics $\sigma_{t,1}(\theta_{1,1}), \dots, \sigma_{t,d}(\theta_{1,d})$ implied by the specification (2.3). The resulting point estimate $\widehat{D}_t := D_t(\widehat{\theta}_1)$ is used to compute the estimated scaled residuals

$$\widehat{\epsilon}_t := \widehat{D}_t^{-1} \mathbf{X}_t. \quad (3.12)$$

The scaled residuals $\widehat{\epsilon}_t$ are used in the second step of our procedure to estimate the tree-structured conditional correlation dynamics (2.4).

3.1 Estimation of tree-structured univariate GARCH-dynamics

Estimation of the d tree-structured univariate volatility functions (2.3) is achieved by a high dimensional model selection problem, which determines the optimal structure of the relevant thresholds in any partition \mathcal{P}_j of the univariate volatility dynamics (2.3), $j = 1, \dots, d$.

In a first step, a largest univariate tree-structured GARCH model is estimated for any $j = 1, \dots, d$, given a fixed maximal number M_j of possible thresholds in (2.3). This first step delivers a maximal possible partition \mathcal{P}_j^{max} of the relevant state space G in the univariate volatility dynamics (2.3).

In a second step, a tree-structured model selection procedure for non nested models is applied, which selects the optimal subpartition $\mathcal{P}_j \subset \mathcal{P}_j^{max}$ out of the maximal one. Model selection is performed according to the BIC information criterion implied by a conditionally gaussian log likelihood for any process coordinate $X_{t,j}$, $j = 1, \dots, d$. The resulting optimal tree-structured volatility model minimizes the BIC information criterion across all tree-structured sub partitions of \mathcal{P}_j^{max} . The complete algorithm used to estimate univariate tree-structured GARCH(1,1) models is given in Audrino and Bühlmann (2001) and Audrino and Trojani (2006).

The construction of the largest partition \mathcal{P}_j^{max} proceeds as follows. We first fix a maximal number $M_j + 1$ of partition cells in the tree. Because of the tree-structured construction of \mathcal{P}_j^{max} , this first step implies a maximal number $M_j + 1$ of conditional volatility regimes (i.e. the number of terminal nodes in the binary tree). A parsimonious specification of the maximal number M_j of thresholds ensures a statistically and computationally tractable model dimension. Moreover, it avoids (over) fitting a too flexible model dynamics, which would result in a poor out-of-sample forecasting power. For any coordinate axis of the multivariate state space that has to be split we search for multivariate thresholds over grid points that are empirical α -quantiles of the data along the relevant coordinate axis. We fix the empirical quantiles as $\alpha = i/\text{mesh}$, $i = 1, \dots, \text{mesh} - 1$, where mesh determines the fineness of the grid on which we search for multivariate thresholds. Typically, we choose $\text{mesh} = 8$. The partition of the state space $G = \mathbb{R}^d \times \mathbb{R}^+$ into a maximal number of $M_j + 1$ cells is performed as follows. A first threshold $d_1 \in \mathbb{R}$ or \mathbb{R}^+ in one coordinate indexed by a component index $\iota_1 \in \{1, 2, \dots, d + 1\}$ partitions G as

$$G = \mathcal{R}_{left} \cup \mathcal{R}_{right},$$

where $\mathcal{R}_{left} = \{(\mathbf{X}_{t-1}, \sigma_{t-1}^2) \in \mathbb{R}^d \times \mathbb{R}^+; (\mathbf{X}_{t-1}, \sigma_{t-1}^2)_{\iota_1} \leq d_1\}$ and $(\mathbf{X}_{t-1}, \sigma_{t-1}^2)_{\iota_1}$ denotes the ι_1 -component of the tuple $(\mathbf{X}_{t-1}, \sigma_{t-1}^2)$. \mathcal{R}_{right} is defined analogously using the relation ' $>$ ' instead of ' \leq '. In a second step, one of the partition cells \mathcal{R}_{left} , \mathcal{R}_{right} is further partitioned with a second threshold d_2 and a second component index ι_2 , in the same way as above. We then iterate this procedure. For the m -th iteration step, we specify a new pair (d_m, ι_m) , which defines a new threshold d_m for the coordinate indexed by ι_m , and an existing partition cell that is going to be splitted into two subcells. For a new pair $(d, \iota) \in \mathbb{R} \times \{1, \dots, d + 1\}$ refinement of an existing partition $\mathcal{P}^{(old)}$ is obtained by picking $\mathcal{R}_{j^*} \in \mathcal{P}^{(old)}$ and splitting it as

$$\mathcal{R}_{j^*} = \mathcal{R}_{j^*,left} \cup \mathcal{R}_{j^*,right} . \tag{3.13}$$

This procedure produces a new (finer) partition of G , given by

$$\mathcal{P}^{(new)} = \{\mathcal{R}_j, \mathcal{R}_{j^*,left}, \mathcal{R}_{j^*,right}, j \neq j^*\}. \quad (3.14)$$

In this partition, the tuple (d, ι) describes a threshold d and a component index ι such that $\mathcal{R}_{j^*,left} = \{(\mathbf{X}_{t-1}, \sigma_{t-1}^2) \in \mathcal{R}_{j^*}; (\mathbf{X}_{t-1}, \sigma_{t-1}^2)_\iota \leq d\}$. $\mathcal{R}_{j^*,right}$ is defined analogously, with the relation ' $>$ ' instead of ' \leq '. The whole procedure finally determines a partition $\mathcal{P}_j^{max} = \{\mathcal{R}_{1,j}, \dots, \mathcal{R}_{M_j+1,j}\}$. This partition can be represented and summarized by a binary tree in which every terminal node represents a partition cell of \mathcal{P}_j^{max} . To select the specific threshold and component index (d, ι) in each iteration step of the above procedure we optimize the corresponding conditional negative (pseudo) log-likelihood in the model.

3.2 Estimation of tree-structured DCC-dynamics

In the first step, we estimate the optimal partition \mathcal{P} using the tree-structured model (2.10) for $\bar{\rho}_t$ and the scaled estimated residuals $\hat{\epsilon}_t$. In the second step, we fix the partition $\hat{\mathcal{P}}$ - say - estimated for the univariate model (2.10), and estimate the parameter θ_2 in (2.7) by a multivariate pseudo maximum likelihood estimator.

(i) *Estimation of the univariate tree structured model (2.10).* Let

$$\hat{\rho}_t = \sum_{u \neq v} \hat{\epsilon}_{t-1,u} \hat{\epsilon}_{t-1,v} / [d(d-1)]. \quad (3.15)$$

The following tree-structured model for $\hat{\rho}_t$ is estimated (compare with equation (2.10)):

$$\hat{\rho}_t = E_{t-1}(\hat{\rho}_t) + \eta_t, \quad (3.16)$$

where $(\eta_t)_{t \in \mathbb{Z}}$ is a martingale difference sequence and

$$E_{t-1}(\hat{\rho}_t) = \sum_{i=1}^w c_i \left(\frac{1}{d(d-1)} \sum_{u \neq v} \hat{R}_{it}^{uv} \right) I_{[(\mathbf{X}_{t-1}, \hat{\rho}_{t-1}) \in \tilde{\mathcal{R}}_i]} \quad (3.17)$$

In this equation, \hat{R}_{it} denotes for $i = 1, \dots, n$ a constant correlation matrix when the tree-structured model for correlations implies a piecewise constant correlation matrix. It then follows in this case that the conditional mean of $\hat{\rho}_t$ is simply a piecewise constant threshold function. More generally, for $c_1 = \dots = c_n = 1$ and $\phi_i > 0$ or $\lambda_i > 0$, where $i = 1, \dots, w$, the local conditional correlation matrix \hat{R}_{it} is simply defined in the same way as R_{it} in equation (2.5) and (2.6), but with $\hat{\epsilon}_{t-1}$ replacing ϵ_{t-1} in equation (2.6).

We apply the same estimation procedure given in the last section for individual conditional variances to the series $\widehat{\rho}_t$. We first estimate a largest univariate tree structured model for $\widehat{\rho}_t$, given a fixed maximal number M of possible thresholds in (2.10). In all our empirical applications we fix the maximal number of candidate thresholds in model (2.10) at $M = 4$. A tree-structured model selection procedure for non nested models is then applied, which selects the optimal subpartition \mathcal{P} out of the maximal one. Model selection is performed according to the BIC criterion implied by a conditionally gaussian pseudo log likelihood for $\widehat{\rho}_t$.³ In our empirical study, we found that this procedure offers a simple and effective way to reduce the computational costs implied by the estimation of our multivariate tree-structured model. In particular, in the applications of Section 4 a piecewise constant conditional correlation function is estimated for equity and different types of bond returns. However, local DCC-type structures are found to model better the conditional correlations between equity and bond returns.

(ii) *Estimation of the tree-structured conditional correlation function R_t .* In the second step of our estimation procedure, we fix the partition $\widehat{\mathcal{P}}$ estimated in step (i) and we estimate the parameter vector θ_2 in (2.7) by a pseudo maximum likelihood estimator $\widehat{\theta}_2$ for θ_2 , under a Gaussian multivariate conditional pseudo likelihood for $\widehat{\epsilon}_t$. If in step (i) the optimal threshold function does not imply piecewise constant correlations, we estimate the matrix \overline{Q} in the dynamics (2.6) by doing correlation targeting, as proposed by Engle and Sheppard (2001) and Pelletier (2006). If in step (i) a piecewise constant correlation structure has been selected, we estimate in the second step a piecewise constant correlation process of the form (2.8). In such a case, we estimate the constant matrix \overline{R} by doing correlation targeting in a rolling window of one year of data. The piecewise constant correlation structure reduces significantly the number of parameters over which the likelihood function has to be maximized.

3.3 Consistency

Proofs of consistency of our model selection procedure for the case where the true model is in the class of tree-structured models are very difficult to obtain. Analogously to CART, it is possible to prove theorems that study the behavior of the prevailing parameter estimators when growing the tree. However, such results do not imply model selection consistency. Furthermore, it is quite hard to believe that the “correct” generating process in our and similar real data examples is indeed exactly a tree-structured model for volatilities and correlations, respectively. For this reason, it is more important to prove consistency of the estimates in a tree-structured

model under a model misspecification, rather than showing consistency of the model selection strategy under the assumption of a correctly specified tree-structured model. Consistency results can be found in Audrino and Bühlmann (2001). Based on such results, consistency of the two-step estimates $(\widehat{\theta}_1, \widehat{\theta}_2)$ in the tree structured DCC-GARCH model under a possible model misspecification can be derived in the standard way under mild regularity conditions; see, for instance, Newey and McFadden (1994). Moreover, efficient estimates could be obtained by performing a further one step Newton-Raphson estimation of the full likelihood, using as starting values the parameter estimates obtained from the two-step procedure (see Pagan, 1986).

4 Results

In this section, we test the in-sample and out-of-sample explanatory power of our tree-structured GARCH-DCC model in three different applications to the econometric analysis of US stock and bond returns. We compare our model with several multivariate GARCH models that have been recently proposed in the literature. Some of these models are nested in our one and can be estimated, like our one, by a two-step estimation procedure:

- The CCC-GARCH model, as proposed by Bollerslev (1990); this model is nested in our one.
- The DCC-GARCH model, as proposed by Engle (2002); this model is nested in our one.
- The RSDC-GARCH model with switching regimes in conditional correlations, as proposed in Pelletier (2006). This model is not formally nested in our one.

Since the individual volatility processes are estimated separately from the correlation dynamics in these models, we can easily focus in our empirical study on the additional explanatory for conditional correlations, which is the main topic of this paper. To achieve this goal, we estimate for all these models identical volatility processes as in our tree-structured DCC-GARCH model, and specify the individual volatility dynamics as univariate tree-structured GARCH(1,1) processes. In addition, we also study the performance of our model in relation to the one of the flexible multivariate GARCH model in Ledoit et al. (2003). This model does not include thresholds or regimes in volatilities or correlations. However, it is based on a more general correlation dynamics than the one implied by Engle (2002) DCC model. Therefore, it is not nested in our setting. Flexible multivariate GARCH models have been shown by Ledoit et al. (2003)

to describe quite accurately the dynamics of stock returns. Therefore, they are further natural competitors to our approach, especially in applications studying the multivariate dynamics of stock markets, as in our first empirical application.

In all our empirical examples, we start from a maximal number $M + 1$ of partition cells in the trees defining the conditional volatility and correlation dynamics, where $M = 4$. This choice implies a maximal number of 5 regimes for conditional volatilities and/or correlations. For any coordinate axis of the multivariate state space that has to be split to determine the thresholds, we search over grid points that are empirical α -quantiles of the data along the relevant coordinate axis. We fix the empirical quantiles as $\alpha = i/\text{mesh}, i = 1, \dots, \text{mesh} - 1$, where mesh determines the fineness of the grid on which we search for multivariate thresholds. In all our estimations we fixed $\text{mesh} = 8$.

To quantify and compare the in-sample and out-of-sample fit of the different models, we compute several goodness-of-fit statistics for conditional covariances. Since the individual volatility processes are identical for all but the flexible multivariate GARCH model in our study, this comparison allows us to investigate the additional explanatory power of our model in explaining the correlation dynamics. We consider the following goodness of fit measures:

- The multivariate negative log-likelihood statistic (NL),
- The multivariate version of the classical mean absolute error statistic (MAE),
- The multivariate version of the root mean squared error statistic (RMSE),

The last two performance measures require the specification of sensible values for the unknown true conditional covariance matrix. A powerful way of computing good proxies for this matrix is by means of the the so-called realized covariance approach, which is the natural multivariate version of the realized volatility approach proposed, among others, in Andersen et al. (2001, 2003) and Barndorff-Nielsen and Shephard (2001, 2002a). We follow this approach in our first two real data applications, in which we collect tick-by-tick return data to compute the realized covariance between returns with the methodology proposed in Corsi and Audrino (2007). In the third and last application, we do not have tick-by-tick data at our disposal to compute the realized covariances between returns. Therefore, we use products of centered returns as (noisy) proxies for the unknown covariances between returns.⁴

The different statistics used to quantify the in-sample and out-of-sample goodness of fit in our empirical analysis are defined as follows: (IS denotes in-sample and OS denotes out-of-sample):

$$\begin{aligned}
\text{IS-NL: } & -\log\text{-likelihood}(\mathbf{X}_1^n; \hat{\phi}) \\
\text{OS-NL: } & -\log\text{-likelihood}(\mathbf{Y}_1^{n_{\text{out}}}; \hat{\phi}) \\
\text{IS-MAE: } & \frac{1}{d^2} \sum_{i,j=1}^d \frac{1}{n} \sum_{t=1}^n |v_{t,ij} - \hat{v}_{t,ij}| \\
\text{OS-MAE: } & \frac{1}{d^2} \sum_{i,j=1}^d \frac{1}{n_{\text{out}}} \sum_{t=1}^{n_{\text{out}}} |v_{t,ij} - \hat{v}_{t,ij}(\mathbf{Y}_1^{t-1})| \\
\text{IS-RMSE: } & \left(\frac{1}{d^2} \sum_{i,j=1}^d \frac{1}{n} \sum_{t=1}^n |v_{t,ij} - \hat{v}_{t,ij}|^2 \right)^{1/2} \\
\text{OS-RMSE: } & \left(\frac{1}{d^2} \sum_{i,j=1}^d \frac{1}{n_{\text{out}}} \sum_{t=1}^{n_{\text{out}}} |v_{t,ij} - \hat{v}_{t,ij}(\mathbf{Y}_1^{t-1})|^2 \right)^{1/2}
\end{aligned}$$

where in the OS performance measures the expression $\hat{v}_{t,ij}(\mathbf{Y}_1^{t-1})$ is the ij -th covariance prediction implied by our out-of-sample data $\mathbf{Y}_1^{n_{\text{out}}} = \{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_{\text{out}}}\}$ at time t under the parameter estimates obtained from the in-sample data $\mathbf{X}_1^n = \{\mathbf{X}_1, \dots, \mathbf{X}_n\}$. $v_{t,ij}$ is the realized covariance between the return series i and j at time t , in our first two applications, or the product of the centered returns of series i and j at time t , in our last application. We mainly focus on the out-of-sample goodness of fit measures. In all cases, a lower goodness of fit measure indicates a higher forecasting power of a model for conditional correlations.

4.1 First real data application: US equity returns

We consider a multivariate time series of (annualized) daily log-returns for ten US stocks: Alcoa, Citigroup, Hasbro, Harley Davidson, Intel, Microsoft, Nike, Pfizer, Tektronix and Exxon. Data are for the sample period between January 2, 2001 and December 30, 2005, amounting to 1256 trading days. The source of the data is *Tick Data*, a division of Nexa Technologies, Inc. (see the webpage <http://www.tickdata.com>). Using these tick-by-tick data, we construct realized covariances with the method in Corsi and Audrino (2007) and obtain the quantities $v_{t,ij}$ needed to compute our goodness of fit measures.

We split the sample in two subperiods. The first one consists of $n = 752$ trading days, from January 2, 2001 to December 31, 2003. Data from this subperiod are used for in-sample estimation and performance evaluation. The second subperiod consists of the remaining $n_{\text{out}} = 504$ observations, up to December 30, 2005, and is used for out-of-sample performance evaluation.

We focus on differences in goodness of fit implied by the conditional correlation matrix dynamics under the different model settings. We estimate our model in two steps as follows. First, we estimate separately the univariate conditional volatility dynamics for each single return series and include as possible conditioning variables in the threshold definition (i) its estimated conditional volatility and (ii) the first lag of all components in the multivariate return series. This threshold volatility structure has proven to produce good empirical results in applications of tree-structured GARCH models to financial data; see, for example, Audrino and Trojani (2006). This first step of the estimation procedure is kept identical for all models in which volatilities can be estimated separately from correlations: the CCC, the DCC, the RSDCC and our tree-structured DCC model. In this way, we ensure that differences in the goodness of fit of these models with respect to the estimated conditional covariance matrix dynamics are exclusively due to differences in the explanatory power with respect to conditional correlations. In the second step of our estimation procedure, we estimate possible tree-structured thresholds and GARCH-type dynamics in conditional correlations. We include as possible conditioning variables for the definition of the threshold structure of conditional correlations (i) the first lag of the average conditional correlation shocks across returns and (ii) the first lag of all components of our multivariate return series; see also Section 2.2 for details.

4.1.1 Estimation results

The estimation results of our tree-structured DCC-GARCH model for the ten-dimensional time series of US stock returns introduced above are summarized in Table 1.

TABLE 1 ABOUT HERE.

Table 1, Panel A, highlights that at most two regimes are necessary to model accurately the individual conditional variance dynamics. The most important predictor variables impacting on the corresponding threshold structures are the lagged returns of Microsoft and Harley Davidson. The structure of the estimated conditional correlation dynamics in our model is summarized in Panel B of Table 1. Similar to volatilities, the most important and statistically significant predictor variable impacting on the threshold structure of conditional correlations is the lagged return of Harley Davidson. Moreover, the complete estimated threshold structure of conditional correlations is fully characterized using only two further lagged stock returns, the returns of Alcoa and Intel (in descending order of statistical significance), and implies four different correlation

regimes. The first regime is associated with contemporarily low lagged Harley Davidson and Alcoa returns.⁵ The second one arises for lagged large Alcoa returns and low Harley Davidson returns. The third regime is obtained for lagged large Harley Davidson and low Intel returns. Finally, the fourth regime is caused by contemporarily large lagged returns of Harley Davidson and Intel. A striking difference between the estimated volatility and correlation dynamics arises for the local DCC models implied by the estimated threshold structures. In contrast to volatilities, the estimated local correlation dynamics never exhibit GARCH-type effects across the different correlation regimes: Conditional correlations are regime dependent, but piecewise constant. Even if the levels of the correlation across regimes are similar, we find that the BIC criterion increases significantly in all cases when incorporating the additional correlation regimes into the model. This finding is confirmed by our following out-of-sample analysis on the model's forecasting power for future correlations.

4.1.2 Multivariate performance results

We now compare the accuracy of the conditional correlation predictions implied by our model with the one implied by the CCC, the DCC, the RSDC and the flexible multivariate GARCH model.⁶ In Table 2, we present the goodness of fit measures defined in Section 4 for the real data application to our ten dimensional stocks returns series.

TABLE 2 ABOUT HERE.

Our multivariate tree structured model achieves the highest goodness of fit compared to the CCC, the DCC and the RSDC models, which are all based on the same volatility dynamics. Only for the in-sample NL measure the RSDC model obtains a larger goodness of fit. Note, however, that this larger in-sample forecasting power is not surprising at all, since the RSDC model has more than twice the number of parameters of the other models. It follows that our tree-structured setting provides a higher out-of-sample forecasting power for conditional correlations. The improvement goes from 0.5% to 17%, depending on the goodness of fit measure used. When comparing the results of our model with those of the flexible multivariate GARCH model, we see that the out-of-sample forecasting performance of the latter specification is only marginally better in two out of three cases, despite having more than double the number of parameters. The heavy parametrization of the flexible multivariate GARCH model can make it unpracticable for problems with dozens to hundreds individual time series. Our and all other models in this

paper, instead, can be used to estimate the conditional variance-covariance dynamics also for very high-dimensional time series settings.

4.2 Second real data application: US stock index and bond returns

We now consider a two-dimensional time series of (annualized) daily log-returns for the US S&P500 stock index and the US 30-years Treasury bond. The time period under investigation goes from January 3, 1996 to October 30, 2003, and contains 1899 trading days. The data is provided by *Tick Data*. As in the previous section, we exploit the tick-by-tick data to construct the series of realized volatilities and covariances between stock index and bond returns. As before, for forecasting evaluation purposes we split the sample in two sub-periods. The first sub-period consists of $n = 1219$ trading days and goes from January 3, 1996 to December 29, 2000. The second sub-period consists of the last three years of data ($n_{\text{out}} = 680$ observations).

4.2.1 Estimation results

The estimation procedure follows the same steps of the one described in the last paragraph of Section 4.1. The individual variance structures and the correlation threshold functions estimated by our tree structured DCC model are summarized in Table 3.

TABLE 3 ABOUT HERE.

The individual threshold functions estimated for the volatilities of each return series depend only on one lag of each series. As previously reported in the literature, for instance in Audrino and Trojani (2006), the conditional variances of stock index returns are driven by more than two regimes. The estimated conditional variance dynamics of Treasury bond returns is determined by two different regimes.

Two regimes also arise for the estimated conditional correlation function between stock index and Treasury bond returns, which is characterized by one threshold that depends exclusively on the lagged return of the S&P500 index. The first regime is determined by lagged low S&P500 returns and the second regime by lagged large S&P500 returns. The estimated local correlation dynamics across regimes feature GARCH-type effects as in Engle (2002) DCC model, but with regime dependent parameters that are significantly different according to the BIC criterion. Therefore, we obtain conditional correlation dynamics that are very different from those obtained in the previous section for our application to a ten dimensional stock returns series.

The flexibility of our setting enables us to take easily into account the different conditional correlation features prevailing among the returns of different classes of assets.

4.2.2 Multivariate performance results

Table 4 summarizes the goodness of fit measures of Section 4 implied by the models estimated on our stock index and Treasury bond return data.

TABLE 4 ABOUT HERE.

Our multivariate tree-structured model obtains the best goodness of fit results across all in-sample and out-of-sample measures used and with respect to all models considered. The improvement in the out-of-sample performance of our model is in the order of 1%-25%, depending on the goodness of fit measure applied. In this application, a piecewise constant conditional correlation dynamics, as the one implied by the RSDC model, has no additional value for forecasting conditional correlations. However, GARCH-type effects in conditional correlations dramatically improve the model's out-of-sample correlation fit.

4.3 Third real data application: US government and corporate AAA bond returns

We conclude our empirical analysis by studying a three-dimensional time series of daily bond index log-returns for the following bond types: corporate AAA intermediate bonds (5 years maturity), corporate AAA long bonds (10 years maturity) and Treasury long bonds (10 years maturity). The data under investigation are for the time period between April 23, 1996 to December 31, 2002, for a total of 1745 observations. We use the first 1223 observations (until the end of 2000) to estimate the different models. The remaining ones are used for forecasting evaluation purposes. The data is provided by *Lehmann Brothers*.

Since no tick-by-tick or other smaller frequencies data are available in this setting, we have to rely on products of (centered) returns as proxies for the unobserved conditional covariances of returns. The main disadvantage caused by this feature is that the arising IS- and OS- MAE and RMSE measures are now noisily proxied. Therefore, small differences in the true unknown goodness of fit measures can be easily obscured by a low signal to noise ratio. Thus, our IS- and OS- measures can be expected to have discrimination power only between models implying quite large differences in forecasting power.

Moreover, as Patton (2006) recently showed in his study, when using such noisy approximations for conditional covariances, the MAE loss function leads to an optimal forecast that is significantly biased. Therefore we report goodness of fit results only for the “robust” NL and RMSE measures.

4.3.1 Estimation results

The individual variance structures and the correlation threshold functions estimated by our tree structured DCC model for the three dimensional time series of bond returns are summarized in Table 5.

TABLE 5 ABOUT HERE.

The conditional variance dynamics estimated for the long Treasury and corporate bond return series feature one single regime. The one estimated for the intermediate corporate bond return series is characterized by a single threshold that depends exclusively on the lagged return of intermediate corporate bonds. The conditional correlation dynamics features a piecewise constant complex structure that is characterized by three regimes. The estimated threshold function for conditional correlations depends completely on the lagged return of intermediate AAA bonds. The first regime is associated with lagged low returns on intermediate AAA corporate bonds. The second one is due to lagged moderately negative AAA intermediate bond returns. The third regime implies lagged positive AAA intermediate bond returns.

4.3.2 Multivariate performance results

The goodness of fit measures of Section 4, implied by the models estimated using our three-dimensional bond return series, are presented in Table 6.

TABLE 6 ABOUT HERE.

Our multivariate tree structured model achieves the best goodness of fit across all in-sample and out-of-sample measures, with the only exception of the IS-NL statistic. Improvements in the out-of-sample goodness of fit of our model over the competitors are approximately 10%. This result is in line with those found in the previous empirical applications and confirms the higher forecasting power of our tree-structured setting for the conditional correlation of future returns in this application, too.

5 Conclusions

We propose a new multivariate DCC-GARCH model that extends previous models by admitting multivariate thresholds in conditional volatilities and correlations. Thresholds in conditional volatilities and correlations are modeled using a tree-structured partition of the multivariate state space and are estimated from the data, together with all other model parameters. Three distinct real data applications support the good forecasting power of our model for forecasting future correlations of financial returns. Competing multivariate GARCH models, including Bollerslev's CCC model, Engle's DCC model and Pelletier's RSDC model have difficulties in fitting adequately the conditional correlation features of financial data, which are found to be often characterized by both multivariate regimes and local GARCH-type effects. Our model is able to cope in a parsimonious way with these features of the data also in applications with large cross-sections of financial asset returns. An interesting venue for future research is the joint empirical modeling of the dynamic correlation features of the returns of several asset classes, like stocks, government and corporate bonds, nominal and index-linked bonds, and exchange rates, which are likely to exhibit rich threshold and GARCH-type effects that could be parsimoniously taken into account by our setting.

Notes

¹To simplify the notation, conditional means of \mathbf{X}_t have been set to zero in (2.1).

²For this case, the trivial parametrization $\bar{R} = \text{diag}[\bar{Q}]^{-1/2}\bar{Q}\text{diag}[\bar{Q}]^{-1/2}$ is superfluous.

³See again Audrino and Trojani (2006), Section 2.3, for details on this estimation procedure.

⁴All estimated models also include a linear autoregressive conditional mean function modeled by a simple diagonal VAR(1) process.

⁵Note that with the terminology “low” (“large”) returns we mean in fact returns that are below (above) the thresholds in the estimated threshold functions.

⁶Since all models, with the only exception of the flexible multivariate GARCH model, have the same individual volatility dynamics, any difference in the goodness of fit for the forecasts of the returns covariance matrix is due to a difference in the quality of the forecasts for conditional correlations.

References

- Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. 2001. The distribution of exchange rate volatility. *Journal of the American Statistical Association* **96**, 42–55.
- Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. 2003. Modeling and forecasting realized volatility. *Econometrica* **71**(2), 579–625.
- Audrino F. (2006). Tree-structured multiple regimes in interest rates. *Journal of Business & Economic Statistics* **24**, No. 3, 338-353.
- Audrino, F. and Bühlmann, P. (2001). Tree-Structured GARCH Models. *Journal of the Royal Statistical Society, Series B*, **63**, No. 4, 727-744.
- Audrino, F. and Trojani, F. (2006). Estimating and predicting multivariate volatility regimes in global stock markets. *Journal of Applied Econometrics* **21**, No. 3, 345-369.
- Barndorff-Nielsen, O. E. and Shephard, N. 2001. Non-gaussian ornstein-uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B*, **63**, 167–241.
- Barndorff-Nielsen, O. E. and Shephard, N. 2002a. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B*, **64**, 253–280.
- Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *The Review of Economics and Statistics* **72**, 498-505.
- Corsi, F. and Audrino, F. (2007). Realized correlation tick-by-tick. Discussion Paper, Department of Economics, University of St. Gallen.
- Engle, R.F. (2002). Dynamic conditional correlation - a simple class of multivariate GARCH models. *Journal of Business and Economic Statistics* **20**, 339-350.
- Engle, R.F. and Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Mimeo, University of California, San Diego.
- Ledoit, O., Santa-Clara, P. and Wolf, M (2003). Flexible multivariate GARCH modeling with an application to international stock markets. *The Review of Economics and Statistics*, **85**, Issue 3, 735-747.

- Newey, W.K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. *Handbook of Econometrics*, Vol. 4, 2111-2245. North-Holland, Amsterdam.
- Pagan, A. (1986). Two stage and related estimators and their applications. *The Review of Economic Studies* **53**, No. 4, 517-538.
- Patton, A.J. (2006). Volatility forecast evaluation and comparison using imperfect volatility proxies. Working paper, London School of Economics.
- Pelletier, D. (2006). Regime switching for dynamic correlations. *Journal of Econometrics* **131**, 445-473.

Panel A. Individual conditional variance structures.

Series	Regimes	Optimal predictors
Alcoa	1	–
Citigroup	2	Microsoft
Hasbro	2	Harley Davidson
Harley Davidson	2	Harley Davidson
Intel	1	–
Microsoft	1	–
Nike	2	Exxon
Pfizer	2	Microsoft
Tektronix	2	Harley Davidson
Exxon	1	–

Panel B: Conditional correlation structure and parameters.

Cond. corr. structure	Cond. corr. parameters
\mathcal{R}_k	c_k
$X_{t-1,Harley\ Davidson} \leq -19.983$ and $X_{t-1,Alcoa} \leq -3.553$	0.928
$X_{t-1,Harley\ Davidson} \leq -19.983$ and $X_{t-1,Alcoa} > -3.553$	0.897
$X_{t-1,Harley\ Davidson} > -19.983$ and $X_{t-1,Intel} \leq -15.016$	0.962
$X_{t-1,Harley\ Davidson} > -19.983$ and $X_{t-1,Intel} > -15.016$	0.955

Table 1: Estimation results for a multivariate time series of ten daily (annualized) US stock returns (in %). Data are for the in-sample time period between January 2, 2001 and December 31, 2003, consisting of 752 observations. Estimated individual conditional variance structures (Panel A) and estimated conditional correlation structure and parameters (Panel B) are for the tree-structured GARCH-DCC model fit.

US equity returns: Goodness of fit results.

Model	# par.	IS-			OS-		
		NL	MAE	RMSE	NL	MAE	RMSE
CCC-GARCH	80	34959	269.413	360.798	21694	105.470	144.161
DCC-GARCH	82	34942	269.525	362.569	21652	103.259	142.440
RSDC-GARCH	173	34684	276.604	371.628	21634	105.470	149.253
TreeDCC-GARCH	84	34927	241.830	327.791	21594	86.0280	126.659
F-MGARCH	185	34378	188.850	296.328	21588	79.6050	129.462

Table 2: Goodness-of-fit of different models for a multivariate time series of ten daily (annualized) US stock returns (in %). Data are for the time period between January 2, 2001 and December 30, 2005, for a total of 1256 observations. The in-sample estimation period goes from the beginning of the sample to the end of 2003 (752 observations). NL, MAE and RMSE are multivariate versions of the standard univariate negative log-likelihood, the mean absolute error, and the root mean squared error statistics. # par. reports the number of parameters estimated by the different models.

Panel A. Individual conditional variance structures.		
Series	Regimes	Optimal predictors
S&P500	3	S&P500
30-years Treasury bond	2	30-years Treasury bond

Panel B: Conditional correlation structure and parameters.		
Cond. corr. structure	Cond. corr. parameters	
\mathcal{R}_k	ϕ_k	λ_k
$X_{t-1,S\&P500} \leq -3.847098$	0.0490	0.9129
$X_{t-1,S\&P500} > -3.847098$	0.0222	0.9724

Table 3: Estimation results for a two-dimensional time series of daily (annualized) returns (in %) for the US S&P500 index and the US 30-years Treasury bond. Data are for the in-sample time period between January 3, 1996 and December 29, 2000, consisting of 1219 observations. Estimated individual conditional variance structures (Panel A) and estimated conditional correlation structure and parameters (Panel B) are for the tree-structured GARCH-DCC model fit.

US index and bond returns: Goodness of fit results.

Model	# par.	IS-			OS-		
		NL	MAE	RMSE	NL	MAE	RMSE
CCC-GARCH	25	9334.7	63.6310	112.334	5511.5	97.2933	146.769
DCC-GARCH	27	9299.2	59.4418	105.613	5451.7	74.8853	120.934
RSDC-GARCH	30	9334.7	63.6310	112.334	5511.5	97.2933	146.769
TreeDCC-GARCH	29	9290.2	58.7296	105.146	5440.9	70.7387	108.145
F-MGARCH	13	9389.2	65.0761	117.326	5483.4	86.1455	141.501

Table 4: Goodness-of-fit of different models for a two-dimensional time series of daily (annualized) returns (in %) on the US S&P500 index and the 30-years Treasury bond. Data are for the time period between January 3, 1996 and October 30, 2003, for a total of 1899 observations. The in-sample estimation period goes from the beginning of the sample to the end of 2000 (1219 observations). NL, MAE and RMSE are the multivariate versions of the standard univariate negative log-likelihood, the mean absolute error, and the root mean squared error statistics. # par. reports the number of parameters estimated in the different models.

Panel A. Individual conditional variance structures.		
Series	Regimes	Optimal predictors
Corporate AAA intermediate	2	Corporate AAA intermediate
Corporate AAA long	1	–
Government long	1	–

Panel B: Conditional correlation structure and parameters.	
Cond. corr. structure	Cond. corr. parameters
\mathcal{R}_k	c_k
$X_{t-1, \text{Corporate AAA intermediate}} \leq -0.0583$	0.965
$-0.0583 < X_{t-1, \text{Corporate AAA intermediate}} \leq 0$	0.985
$X_{t-1, \text{Corporate AAA intermediate}} > 0$	0.947

Table 5: Estimation results for a three-dimensional time series of daily returns (in %) of the following US bond: 5-years corporate AAA (intermediate), 10-years corporate AAA (long) and 10-years government (long). Data are for the in-sample time period between April 23, 1996 and December 29, 2000, consisting of 1223 observations. Estimated individual conditional variance structures (Panel A) and estimated conditional correlation structure and parameters (Panel B) are for the tree-structured GARCH-DCC model fit.

US corporate and government bond returns: Goodness of fit results.

Model	# par.	IS-		OS-	
		NL	RMSE	NL	RMSE
CCC-GARCH	20	13.4818	0.4142	433.01	0.4620
DCC-GARCH	22	-297.56	0.4145	454.41	0.4624
RSDC-GARCH	29	-336.12	0.4159	418.16	0.4653
TreeDCC-GARCH	23	-100.02	0.3723	377.91	0.4102
F-MGARCH	24	-41.331	0.4294	425.32	0.4841

Table 6: Goodness-of-fit of different models for a three-dimensional time series of daily returns (in %) of US 5-years and 10-years corporate AAA bonds and 10-years Treasury bonds. Data are for the time period between April 23, 1996 and December 31, 2002, for a total of 1745 observations. The in-sample estimation period goes from the beginning of the sample to the end of 2000 (1223 observations). NL and RMSE are multivariate versions of the standard univariate negative log-likelihood and the root mean squared error statistics. # par. reports the number of parameters estimated in the different models.