REFERENCES

Linked references are available on JSTOR for this article:
https://www.jstor.org/stable/25800785?seq=1&cid=pdf-reference#references_tab_contents
You may need to log in to JSTOR to access the linked references.

A General Multivariate Threshold GARCH Model With Dynamic Conditional Correlations

Francesco AUDRINO
Department of Economics, Institute of Mathematics and Statistics, University of St. Gallen,
CH-9000 St. Gallen, Switzerland (*francesco.audrino@unisg.ch*)

Fabio TROJANI
University of Lugano and Swiss Finance Institute, CH-6900 Lugano, Switzerland (*fabio.trojani@usi.ch*)

We introduce a new multivariate GARCH model with multivariate thresholds in conditional correlations and develop a two-step estimation procedure that is feasible in large dimensional applications. Optimal threshold functions are estimated endogenously from the data and the model conditional covariance matrix is ensured to be positive definite. We study the empirical performance of our model in two applications using U.S. stock and bond market data. In both applications our model has, in terms of statistical and economic significance, higher forecasting power than several other multivariate GARCH models for conditional correlations.

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

1. INTRODUCTION

In this article, 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 that are often encountered, for instance, in financial real data applications. As in 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 volatilities and correlations. As shown in Audrino and Trojani (2006) and Audrino (2006), the tree-structured threshold construction can incorporate a potentially large number of multivariate regimes in univariate settings in a parsimonious way. In this article, we study a multivariate model with a potentially large number of tree-structured thresholds in volatilities and correlations and we develop a feasible estimation strategy that can be applied to estimate the model in large dimensional applications as well. The threshold construction is obtained using a binary tree in which each terminal node defines 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 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 two distinct applications to U.S. stock and bond market data and focus on the explanatory power for future conditional correlations, comparing the results with a set of competing models in the literature: Engle's (2002) DCC model, Ledoit, Santa-Clara, and Wolf's (2003) flexible multivariate GARCH model, and Pelletier's (2006) regime switching dynamic correlations (RSDC) model. As in our model, the DCC and the RSDC models can be estimated 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 the one in our model. The flexible multivariate GARCH model cannot be estimated by a two-step estimation procedure. Therefore, the estimated volatility dynamics are different from the dynamics of our model. Our model differs from the others in the way it specifies the correlation dynamics. The DCC and the flexible multivariate GARCH models are single-regime models for correlations and the RSDC model specifies a very simple regime structure for conditional correlations. By contrast, our setting can account parsimoniously for GARCH-type dynamics and multivariate conditional correlation thresholds without ex ante fixing the structure and the number of thresholds.

Using our tree-structured GARCH-DCC model, we empirically study the relative importance of GARCH and threshold effects in the conditional correlation dynamics of U.S. stock and bond returns. The conditional volatility functions of U.S. stock returns exhibit GARCH and threshold features, but conditional correlations depend on a piecewise constant threshold function. By contrast, we find both threshold and GARCH effects for the correlation process of stock and government bond returns, with

138

estimated thresholds that are functions of lagged stock and bond returns. In all applications, the estimated tree-structured partition of the state space improves the forecasting power for conditional correlations relative to the other multivariate GARCH models. Out-of-sample improvements are, in most cases, statistically and economically significant based on different criteria recently proposed in the literature including the test for superior predictive ability (SPA) introduced by Hansen (2005), the model confidence set approach of Hansen, Lunde, and Nason (2003), and the economic measure of volatility and correlation timing ability for portfolio allocation in Engle and Colacito (2006) and Bandi, Russell, and Zhu (2008). These findings highlight the importance of flexible multivariate threshold and GARCH structures for forecasting the conditional correlation of stock and bond markets.

In this article, we develop a multivariate GARCH model for variances and correlations, which has good forecasting power and which can be estimated exclusively from information on multivariate returns. A completely different approach can instead formulate a dynamic model for realized volatilities and correlations, using information from intraday returns when available, in order to produce even more accurate forecasts; see, for example, Andersen et al. (2003) or Audrino and Corsi (2009). This empirical strategy produces very good results in univariate settings. However, its extension to the multivariate context is not straightforward because of the difficulty of accurately estimating realized correlations under nonsynchronous intraday returns. It was shown that this problem can produce high efficiency losses in estimating high-dimensional realized variance covariance matrices; see, for example, Barndorff-Nielsen et al. (2008) or Chiriac and Voev (2009). Solving this important issue is a crucial topic of ongoing research.

Section 2 presents our tree-structured GARCH-DCC model and Section 3 describes the two-step estimation procedure that can be applied to estimate it. Section 4 presents our empirical study of the conditional correlation dynamics of U.S. stock and bond returns. 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 $\mathbb{R}^d$:

$$\mathbf{X}_t = \mathbf{D}_t \boldsymbol{\epsilon}_t, \qquad (2.1)$$

where $\mathbf{D}_t := \mathrm{diag}[\sigma_{1,t}, \ldots, \sigma_{d,t}]$ and $\sigma_{i,t}$ is the conditional standard deviation of the ith component of $\mathbf{X}_t$ at time $t-1$. $(\boldsymbol{\epsilon}_t)_{t \in \mathbb{Z}}$ is a zero-mean process in $\mathbb{R}^d$ with components having a unit conditional standard deviation by construction. To simplify the notation, conditional means of $\mathbf{X}_t$ were set to zero in Equation (2.1). The conditional covariance matrix of $\boldsymbol{\epsilon}_t$ at time $t-1$ is denoted by $\mathbf{R}_t$. Therefore, we obtain the following standard factorization of the conditional covariance matrix of $\mathbf{X}_t$:

$$\mathrm{cov}_{t-1}(\mathbf{X}_t) = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t. \qquad (2.2)$$

Our tree-structured DCC-GARCH model parameterizes the conditional volatility matrix $\mathbf{D}_t$ and the conditional correlation matrix $\mathbf{R}_t$ by means of two parametric threshold functions.

Each diagonal element of $\mathbf{D}_t$ 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 $\mathbf{R}_t$ is modeled according to a threshold DCC-type model described in more detail in the following.

2.1 Tree-Structured Model for $\mathbf{D}_t$

Let $X_{t,j}$ be the jth component of $\mathbf{X}_t$. In principle, the thresholds in the volatility dynamics of $X_{t,j}$ may 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 $\mathcal{P}_j = \{\mathcal{R}_{1,j}, \ldots, \mathcal{R}_{k_j,j}\}$ be a partition of the state space $G := \mathbb{R}^2 \times \mathbb{R}^+$ of $(X_{t-1,1}, X_{t-1,j}, \sigma^2_{t-1,j})$:

$$\mathcal{P}_j = \{\mathcal{R}_{1,j}, \ldots, \mathcal{R}_{k_j,j}\},$$

$$\bigcup_{s=1}^{k_j} \mathcal{R}_{s,j} = G, \qquad \mathcal{R}_{i,j} \cap \mathcal{R}_{s,j} = \emptyset \qquad (i \neq s).$$

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, the threshold function $\sigma^2_{t,j}$ takes the form

$$\sigma^2_{t,j} = \sum_{i=1}^{k_j} (\alpha_{ij} + \beta_{ij} X^2_{t-1,j} + \gamma_{ij} \sigma^2_{t-1,j})$$

$$\times I_{[(X_{t-1,1}, X_{t-1,j}, \sigma^2_{t-1,j}) \in \mathcal{R}_{i,j}]}, \qquad (2.3)$$

where $I_{[\cdot]}$ is the indicator function and $\boldsymbol{\theta}_{1,j}$ is the parameter vector

$$\boldsymbol{\theta}_{1,j} = \{\alpha_{ij}, \beta_{ij}, \gamma_{ij}; i = 1, \ldots, k_j\}.$$

To completely specify the conditional variance function in Equation (2.3), we have to define the class of partitions $\mathcal{P}_j$ that are admissible in our tree-structured model. The first assumption is that $\mathcal{P}_j$ is composed of rectangular cells $\mathcal{R}_{i,j}$, $i = 1, \ldots, k_j$, so that they can be easily parameterized by a set of multivariate thresholds for $(X_{t-1,1}, X_{t-1,j}, \sigma^2_{t-1,j})$. The second assumption is that the potential partition cells satisfy a hierarchical structure that can be mapped one-to-one on a so-called binary tree by means of an iterative statistical procedure. In this way, the multivariate threshold function in the model follows the structure of a binary tree $\mathcal{T}_j$ in which every terminal node represents a particular regime.

In constructing the binary tree, it is first tested whether a threshold d_1 can split the whole state space into two rectangular partition cells representing two different volatility regimes. In this case, d_1 defines the first node on the binary tree and the first two partition cells represent the two stems associated with the first node. This tree-structure is preferred with respect to a model with no variance–covariance thresholds if the implied improvement in goodness of fit is large enough. In the second step, it is checked whether one of the two cells identified from the first step can be further split into two rectangular subcells by an additional threshold d_2. The decision on which subcell can be further split is again based on a comparison of the implied improvement in the implied goodness of fit. Threshold d_2 defines a second node on the binary tree. Two subcells are associated with this node that define two new stems of the tree and two

further volatility subregimes. Such a procedure is iterated until a maximal tree and a maximal number of variance–covariance regimes are obtained. See the following for additional details on the model construction and the estimation procedure. Details on the interpretation of binary trees in the context of volatility models are provided in Audrino and Trojani (2006).

Consider, for example, a model with three regimes, i.e., two thresholds, and partitioning cells $\mathcal{R}_{1,j}, \mathcal{R}_{2,j}, \mathcal{R}_{3,j}$ of the form:

$$\mathcal{R}_{1,j} = \{X_{t-1,j} \leq d_1\},$$

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

$$\mathcal{R}_{3,j} = \{X_{t-1,j} > d_1 \text{ and } X_{t-1,1} > d_2\},$$

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. For each component $X_{t,j}$, estimation of Equation (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 of 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

Given $\theta_1 = \{\theta_{1,j} : j = 1, \dots, d\}$, let

$$\epsilon_t = D_t(\theta_1)^{-1} X_t$$

and

$$R_t = \text{corr}_{t-1}(X_t) = \text{cov}_{t-1}(\epsilon_t).$$

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 X_{t-1} and specific movements in average lagged conditional correlation shocks $\bar{\rho}_{t-1}$.

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

$$R_t = \sum_{i=1}^{w} c_i R_{it} I_{\{(X_{t-1}, \bar{\rho}_{t-1}) \in \widetilde{\mathcal{R}}_i\}}$$

$$+ \left(1 - \sum_{i=1}^{w} c_i I_{[(X_{t-1}, \bar{\rho}_{t-1}) \in \widetilde{\mathcal{R}}_i]}\right) I_d, \quad (2.4)$$

where $c_1, \dots, c_w \in [0, 1]$, I_d 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, as 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

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

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

Since for any $i = 1, \dots, w$, the local model for Q_{it} satisfies an 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 $\mathcal{P} = \{\widetilde{G}\}$, i.e., the partition is trivial, we obtain the Engle (2002) DCC model by setting $c_1 = \cdots = c_w = 1$. Therefore, this model is nested in our model. Moreover, by setting $\phi_i = \lambda_i = 0$ for $i = 1, \dots, w$, we can write R_t as

$$R_t = \sum_{i=1}^{w} c_i \bar{R} I_{[(X_{t-1}, \bar{\rho}_{t-1}) \in \widetilde{\mathcal{R}}_i]}$$

$$+ \left(1 - \sum_{i=1}^{w} c_i I_{[(X_{t-1}, \bar{\rho}_{t-1}) \in \widetilde{\mathcal{R}}_i]}\right) I_d, \quad (2.8)$$

where $\bar{R}$ is a fixed d-dimensional correlation matrix. 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, by setting $c_1 = \cdots = c_w = 1$ we obtain a tree-structured DCC model locally satisfying Engle's DCC dynamics over the distinct partitioning cells $\widetilde{\mathcal{R}}_i$.

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 of rectangular partition cells. Consistent with our assumptions, these partition cells are delimited by a set of multivariate thresholds for $(X_{t-1}, \bar{\rho}_{t-1})$. In order to construct such rectangular partition cells, we make use of a binary tree in which every terminal node represents a cell $\widetilde{\mathcal{R}}_i$. Estimation of the threshold function in the correlation dynamics in Equation (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 notice that the partition $\mathcal{P}$ is identical to the one implied by a corresponding

tree-structured univariate model for the time series $(\overline{\rho}_t)_{t \in \mathbb{Z}}$. Indeed, since $\mathbf{R}_t = E_{t-1}(\epsilon_t \epsilon_t')$ it follows:

$$E_{t-1}(\overline{\rho}_t) = \frac{1}{d(d-1)} \sum_{u \neq v} R_t^{uv}$$

$$= \sum_{i=1}^{w} c_i \left(\frac{1}{d(d-1)} \sum_{u \neq v} R_{it}^{uv} \right) I_{[(\mathbf{X}_{t-1}, \overline{\rho}_{t-1}) \in \widetilde{\mathcal{R}}_i]}, \quad (2.9)$$

where R_t^{uv} (R_{it}^{uv}) denotes the uvth component of the matrix $\mathbf{R}_t$ ($\mathbf{R}_{it}$). Therefore, the tree-structured model

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

where $(\eta_t)_{t \in \mathbb{Z}}$ is a martingale difference process and $E_{t-1}(\overline{\rho}_t)$ is given by Equation (2.9), defines a univariate tree-structured process for $\overline{\rho}_t$ based on the same partition $\mathcal{P}$ as in the correlation dynamics in Equation (2.4). It follows that we can exploit the univariate model in Equation (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 simplest dynamics arise in the piecewise constant case

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

where $\overline{R}^{uv}$ is the uv component of the correlation matrix $\overline{\mathbf{R}}$ in the piecewise-constant dynamics in Equation (2.8). This piecewise-constant function is the optimal one that was estimated in our applications to the U.S. equity market in Section 4.1. More generally, for $c_1 = \cdots = c_w = 1$ and $\lambda_i, \phi_i > 0$, $i = 1, \ldots, w$, we can also encompass the univariate dynamics of $\overline{\rho}_t$ that are consistent with a tree-structured DCC model of the form in Equation (2.4) for correlations. This threshold structure is the one we estimate in our application of Section 4.2 where 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 Equation (2.10) is estimated the parameter in Equation (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 accomplished in two steps. In the first step, an estimate of the volatility process $\mathbf{D}_t$ is obtained by performing d estimations of the univariate tree-structured conditional volatility dynamics $\sigma_{t,1}(\boldsymbol{\theta}_{1,1}), \ldots, \sigma_{t,d}(\boldsymbol{\theta}_{1,d})$ implied by the specification in Equation (2.3). The resulting point estimate $\widehat{\mathbf{D}}_t := \mathbf{D}_t(\widehat{\boldsymbol{\theta}}_1)$ is used to compute the estimated scaled residuals

$$\widehat{\epsilon}_t := \widehat{\mathbf{D}}_t^{-1} \mathbf{X}_t. \quad (3.1)$$

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

3.1 Estimation of Tree-Structured Univariate GARCH-Dynamics

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

In the first step, the largest univariate tree-structured GARCH model is estimated for any $j = 1, \ldots, d$, given a fixed maximal number M_j of possible thresholds in Equation (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 in Equation (2.3).

In the second step, a tree-structured model selection procedure for nonnested models is applied that 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 $X_{t,j}, j = 1, \ldots, d$. The resulting optimal tree-structured volatility model minimizes the BIC information criterion across all tree-structured subpartitions 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 an overly flexible model dynamics, which will 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 $\alpha = i/\text{mesh}$, $i = 1, \ldots, \text{mesh} - 1$, where mesh determines the fineness of the grid on which we search for multivariate thresholds. We choose mesh $= 8$ because, as was shown in different empirical studies in the literature, this value leads to reliable forecasting results. 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 $\iota_1 \in \{1, 2, \ldots, d+1\}$ partitions G as

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

where $\mathcal{R}_{\text{left}} = \{(\mathbf{X}_{t-1}, \sigma_{t-1}^2) \in \mathbb{R}^d \times \mathbb{R}^+; (\mathbf{X}_{t-1}, \sigma_{t-1}^2)_{\iota_1} \leq d_1\}$ and $(\mathbf{X}_{t-1}, \sigma_{t-1}^2)_{\iota_1}$ denotes the ι_1 component of the tuple $(\mathbf{X}_{t-1}, \sigma_{t-1}^2)$. $\mathcal{R}_{\text{right}}$ is defined analogously using the relation ">" instead of "$\leq$." In the second step, one of the partition cells $\mathcal{R}_{\text{left}}, \mathcal{R}_{\text{right}}$ is further partitioned with a second threshold d_2 and a second component index ι_2, in the same way as previously. We then iterate this procedure. For the mth 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 split into two subcells. For a

new pair $(d, \iota) \in \mathbb{R} \times \{1, \ldots, d + 1\}$ refinement of an existing partition $\mathcal{P}^{(\text{old})}$ is obtained by picking $\mathcal{R}_{j*} \in \mathcal{P}^{(\text{old})}$ and splitting it as

$$\mathcal{R}_{j*} = \mathcal{R}_{j*,\text{left}} \cup \mathcal{R}_{j*,\text{right}}. \qquad (3.2)$$

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

$$\mathcal{P}^{(\text{new})} = \{\mathcal{R}_j, \mathcal{R}_{j*,\text{left}}, \mathcal{R}_{j*,\text{right}}, j \neq j^*\}. \qquad (3.3)$$

In this partition, the tuple (d, ι) describes a threshold d and a component index ι such that $\mathcal{R}_{j*,\text{left}} = \{(\mathbf{X}_{t-1}, \sigma_{t-1}^2) \in \mathcal{R}_{j*}; (\mathbf{X}_{t-1}, \sigma_{t-1}^2)_\iota \leq d\}$. $\mathcal{R}_{j*,\text{right}}$ is defined analogously, with the relation ">" instead of "$\leq$." The whole procedure finally determines a partition $\mathcal{P}_j^{\max} = \{\mathcal{R}_{1,j}, \ldots, \mathcal{R}_{M_j+1,j}\}$. 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 in Equation (2.10) for $\bar{\rho}_t$ and the scaled estimated residuals $\widehat{\epsilon}_t$. In the second step, we fix the partition $\widehat{\mathcal{P}}$ estimated for the univariate model in Equation (2.10), and estimate the parameter θ_2 in Equation (2.7) by a multivariate pseudo-maximum likelihood estimator.

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

$$\widehat{\bar{\rho}}_t = \sum_{u \neq v} \widehat{\epsilon}_{t-1,u} \widehat{\epsilon}_{t-1,v} / [d(d-1)]. \qquad (3.4)$$

The following tree–structured model for $\widehat{\bar{\rho}}_t$ is estimated; compare with Equation (2.10)

$$\widehat{\bar{\rho}}_t = E_{t-1}(\widehat{\bar{\rho}}_t) + \eta_t, \qquad (3.5)$$

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

$$E_{t-1}(\widehat{\bar{\rho}}_t) = \sum_{i=1}^{w} c_i \left(\frac{1}{d(d-1)} \sum_{u \neq v} \hat{R}_{it}^{uv} \right) I_{[(\mathbf{X}_{t-1}, \widehat{\bar{\rho}}_{t-1}) \in \widetilde{\mathcal{R}}_i]}. \qquad (3.6)$$

In this equation, $\hat{\mathbf{R}}_{it}$ denotes for $i = 1, \ldots, w$ 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 $\widehat{\bar{\rho}}_t$ is simply a piecewise-constant threshold function. More generally, for $c_1 = \cdots = c_w = 1$ and $\phi_i > 0$ or $\lambda_i > 0$, where $i = 1, \ldots, w$, the local conditional correlation matrix $\hat{\mathbf{R}}_{it}$ is simply defined in the same way as $\mathbf{R}_{it}$ in Equations (2.5) and (2.6), but with $\hat{\epsilon}_{t-1}$ replacing ϵ_{t-1} in Equation (2.6).

We apply to the series $\widehat{\bar{\rho}}_t$ the same estimation procedure given in the last section for individual conditional variances. First, we estimate a largest univariate tree-structured model for $\widehat{\bar{\rho}}_t$, given a fixed maximal number M of possible thresholds in Equation (2.10). In all our empirical applications, we fix the maximal number of candidate thresholds in Equation (2.10) at $M = 4$. A tree-structured model selection procedure for

nonnested models is then applied that 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 $\widehat{\bar{\rho}}_t$; see again Audrino and Trojani (2006, section 2.3), for details on this estimation procedure. In our empirical study, we find 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 returns. However, local DCC-type structures are found to better model the conditional correlations between equity and bond returns.

(ii) *Estimation of the tree-structured conditional correlation function* $\mathbf{R}_t$. 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 Equation (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 $\overline{\mathbf{Q}}$ in the dynamics in Equation (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 is selected, in the second step we estimate a piecewise constant correlation process of the form in Equation (2.8). In such a case, we estimate the constant matrix $\overline{\mathbf{R}}$ by doing correlation targeting in a rolling window of one year of data. The piecewise-constant correlation structure significantly reduces the number of parameters over which the likelihood function has to be maximized.

3.3 Consistency

Proofs of the consistency of our model selection procedure are very difficult to obtain for the case where the true model is in the class of tree-structured GARCH-DCC models. Analogously to the standard classification and regression trees (CART) procedure introduced by Breiman et al. (1984), 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 unlikely that the "correct" generating process in our example and other similar ones with real data is indeed exactly a tree-structured model for volatilities and correlations. For this reason, it is more important to prove consistency of the estimators for the parameters of a tree-structured model under a model misspecification than it is to prove consistency of the model selection strategy under the assumption of a correctly specified, tree-structured model. Such consistency results can be found in Audrino and Bühlmann (2001). Based on these 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 can 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, e.g., Pagan 1986).

4. RESULTS

In this section we test the in-sample and out-of-sample explanatory power of our tree-structured GARCH-DCC (TreeDCC) model in two different applications involving the econometric analysis of U.S. stock and bond returns. We compare our model with several multivariate GARCH models that were recently proposed in the literature. Some of these models are nested in ours and can be estimated by a two-step estimation procedure:

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

Since the individual volatility processes are estimated separately from the correlation dynamics in these models, in our empirical study we can easily focus on the additional explanatory power for conditional correlations, which is the main topic of this article. To achieve this goal, we estimate volatility processes identical to those in our tree-structured DCC-GARCH setting. The volatility processes are all specified as univariate tree-structured GARCH(1, 1) processes. We also study the performance of our model relative to the flexible multivariate GARCH setting in Ledoit, Santa-Clara, and Wolf (2003), which does not include thresholds in volatilities or correlations and is based on a more general correlation dynamics than the one implied by the Engle (2002) DCC model. Therefore, this model is not nested in our setting. Flexible multivariate GARCH models were shown by Ledoit, Santa-Clara, and Wolf (2003) to describe the dynamics of stock returns quite accurately. Therefore, they are further natural competitors of our approach, especially in applications that study the multivariate dynamics of stock markets, as in our first empirical example.

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, this comparison allows us to investigate the additional explanatory power of our model for 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), and
- The multivariate version of the classical mean squared error statistic (MSE).

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 so-called realized covariance approach, which is the natural multivariate version of the realized volatility approach proposed, among others, by Andersen et al. (2001, 2003) and Barndorff-Nielsen and Shephard (2001, 2002). We follow this approach in our two real data applications, in which we collect tick-by-tick return data to compute the realized covariance between returns using the methodology proposed in Corsi and Audrino (2007). Using such an accurate proxy for the unobservable conditional covariance matrix allows us to avoid possible misleading results implied by an unfortunate choice of the loss function used to quantify the goodness of fit, and therefore, a wrong ranking of the different models under investigation; see Laurent, Rombouts, and Violante (2009) for more details. All estimated models also include a linear autoregressive conditional mean function modeled by a simple diagonal VAR(1) process.

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):

$$\text{IS-NL:} \quad -\text{log-likelihood}(\mathbf{X}_1^n; \widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2),$$

$$\text{OS-NL:} \quad -\text{log-likelihood}(\mathbf{Y}_1^{n_{\text{out}}}; \widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2),$$

$$\text{IS-MAE:} \quad \frac{1}{d^2} \sum_{i,j=1}^{d} \frac{1}{n} \sum_{t=1}^{n} |v_{t,ij} - \widehat{v}_{t,ij}|,$$

$$\text{OS-MAE:} \quad \frac{1}{d^2} \sum_{i,j=1}^{d} \frac{1}{n_{\text{out}}} \sum_{t=1}^{n_{\text{out}}} |v_{t,ij} - \widehat{v}_{t,ij}(\mathbf{Y}_1^{t-1})|,$$

$$\text{IS-MSE:} \quad \frac{1}{d^2} \sum_{i,j=1}^{d} \frac{1}{n} \sum_{t=1}^{n} |v_{t,ij} - \widehat{v}_{t,ij}|^2,$$

$$\text{OS-MSE:} \quad \frac{1}{d^2} \sum_{i,j=1}^{d} \frac{1}{n_{\text{out}}} \sum_{t=1}^{n_{\text{out}}} |v_{t,ij} - \widehat{v}_{t,ij}(\mathbf{Y}_1^{t-1})|^2,$$

where in the OS performance measures the expression $\widehat{v}_{t,ij}(\mathbf{Y}_1^{t-1})$ is the ijth covariance prediction implied by our out-of-sample data $\mathbf{Y}_1^{n_{\text{out}}} = \{\mathbf{Y}_1, \ldots, \mathbf{Y}_{n_{\text{out}}}\}$ at time t under the parameter estimates obtained from the in-sample data $\mathbf{X}_1^n = \{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$. $v_{t,ij}$ is the realized covariance between the return series i and j at time t. In all cases, a lower goodness-of-fit measure indicates a higher forecasting power of a model for conditional correlations.

To evaluate whether differences in performance among the models considered are statistically and/or economically significant, we perform a series of tests recently proposed in the literature. We investigate the statistical relevance of the improvements in forecasting accuracy of the TreeDCC model relative to the other approaches:

- We perform the superior predictive ability (SPA) test introduced by Hansen (2005). This allows us to verify whether each of the models considered is significantly outperformed by one (or more) of the alternatives.
- We construct a model confidence set (MCS) at the 5% and 10% confidence levels as proposed by Hansen, Lunde, and Nason (2003). We introduce the MCS approach with the goal of characterizing the best subset of models out of a set of competing ones; see the Appendix for more details about the MCS construction.

We also evaluate the alternative model specifications using an economic criterion. We apply the methodology suggested by West, Edison, and Cho (1993) and Fleming, Kirby, and Ostdiek

(2001, 2003) to quantify the economic benefit of different correlation forecasts in the context of a portfolio strategy based on volatility timing. Similarly to Engle and Colacito (2006) and Bandi, Russell, and Zhu (2008), we employ the variance component of an investor's long-run mean–variance utility as a metric to quantify the economic differences between the alternative correlation (covariance) forecasts

$$AU = \frac{\lambda}{2} \frac{1}{n_{\text{out}}} \sum_{t=1}^{n_{\text{out}}} (R_t^p - \overline{R}^p)^2,$$

where the portfolio return at time t is given by

$$R_t^p = R^f + \mathbf{w}_{t-1}'(\mathbf{Y}_t - R^f \mathbf{I}_d), \qquad t = 1, \ldots, n_{\text{out}}.$$

$\overline{R}^p$ is the sample mean of the portfolio returns across the out-of-sample period and λ is a coefficient of risk aversion. As in Bandi, Russell, and Zhu (2008), we use three values of $\lambda = 2, 7, 10$. In the computation of the portfolio returns, R^f is the risk-free rate, $\mathbf{I}_d$ is a $d \times 1$ unit vector, and $\mathbf{w}_{t-1}$ is a d vector of portfolio weights obtained by solving the classical mean–variance optimization problem at time $t - 1$, given a fixed target return on the portfolio. In our real data applications, we set R^f equal to the average value of the U.S. three-month rate in the out-of-sample period. In the mean–variance optimization, we use the one-step-ahead conditional covariance forecasts $\hat{v}_{t,ij}(\mathbf{Y}_1^{t-1})$ obtained from the alternative multivariate GARCH models under investigation. We can interpret the difference between the quantity AU implied by the TreeDCC model and each other model as the fee that an investor will be willing to pay to switch from correlation forecasts generated by each alternative model to those of the TreeDCC model. We also use the generalization of the Diebold–Mariano test for pairwise equal predictive ability (EPA) and the joint test introduced in Engle and Colacito (2006) to study the statistical significance of the differences in estimated economic gains. For details, see again Engle and Colacito (2006) and Bandi, Russell, and Zhu (2008).

4.1 First Real Data Application: U.S. Equity Returns

We consider a multivariate time series of (annualized) daily log-returns for 10 U.S. 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 using 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 into 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 separately estimate 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 proved 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 RSDC, 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 again Section 2.2 for details.

4.1.1 Estimation Results.
The estimation results of our TreeDCC-GARCH model for the 10-dimensional time series of U.S. stock returns introduced earlier are summarized in Table 1.

Table 1, Panel A, highlights that, at most, two regimes are necessary to model the individual conditional variance dynamics accurately. The most important predictor variables impacting on the corresponding threshold structures are the lagged returns of Microsoft and Harley Davidson. Microsoft and Harley Davidson are the largest stocks in our empirical example. Thus, the apparent influence of their lagged returns on estimated variance covariance regimes is likely to proxy for the latent impact of the aggregate market return on the multivariate variance covariance dynamics. 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 threshold structure of conditional correlations is characterized using only two further lagged stock returns, the returns of Alcoa and Intel (in descending order of statistical significance), implying four correlation regimes overall. The first regime is associated with simultaneously low (i.e., under the threshold values) lagged Harley Davidson and Alcoa returns. The second one arises for lagged, low Harley Davidson returns and for large Alcoa returns. The third regime is obtained for lagged low Intel returns and large Harley Davidson returns. Finally, the fourth regime is caused by contemporarily somewhat higher returns of Harley Davidson and Intel. An important difference between the estimated volatility and correlation dynamics is that the local correlation dynamics never exhibit GARCH-type effects across the different correlation regimes. In other words, conditional correlations are regime dependent but piecewise constant. The local average correlation levels in the different regimes are similar and vary from 0.897 to 0.962. However, we find that the BIC criterion increases significantly in all cases when incorporating

Table 1. Estimation results for a multivariate time series of 10 daily (annualized) U.S. stock returns (in %)

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 $\mathcal{R}_k$	Cond. corr. parameters c_k
$X_{t-1,\text{Harley Davidson}} \leq -19.983$ and $X_{t-1,\text{Alcoa}} \leq -3.553$	0.918 (0.029)
$X_{t-1,\text{Harley Davidson}} \leq -19.983$ and $X_{t-1,\text{Alcoa}} > -3.553$	0.897 (0.056)
$X_{t-1,\text{Harley Davidson}} > -19.983$ and $X_{t-1,\text{Intel}} \leq -15.016$	0.962 (0.016)
$X_{t-1,\text{Harley Davidson}} > -19.983$ and $X_{t-1,\text{Intel}} > -15.016$	0.935 (0.023)

NOTE: 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. Standard errors computed using 1000 model-based bootstrap replications are given in parentheses.

additional correlation regimes into the model. This finding is also supported by our out-of-sample tests on the model's forecasting power for correlations, which further indicate a clear superiority of our model over a CCC model with constant correlations.

4.1.2 Multivariate Performance Results.
We now compare the accuracy of the conditional correlation predictions

implied by our model with those implied by the CCC, the DCC, the RSDC, and the flexible multivariate GARCH model. Since all models have the same individual volatility dynamics, with the sole exception being the flexible multivariate GARCH model, any difference in the goodness of fit for the forecasts of the return covariance matrix is due to a difference in the quality of the forecasts for conditional correlations. In Table 2, we present the goodness-of-fit measures defined in Section 4 for the real data application to our 10-dimensional stock return series.

Our multivariate tree-structured model achieves the highest goodness of fit overall across the models with identical volatility dynamics (TreeDCC, CCC, DCC, and RSDC models). The RSDC model achieves a better in-sample log-likelihood criterion, but it implies worse MAE and MSE criteria both in-sample and out-of-sample. Since the RSDC model has more than twice the number of parameters of the other models, we interpret this finding as evidence of over-fitting for this particular model. The TreeDCC model's better performance relative to CCC, DCC, and RSDC ranges between 0.5% and 17% depending on the goodness-of-fit measure applied. In comparison with the flexible multivariate GARCH model, we find that the out-of-sample forecasting performance of the latter specification is slightly better in two out of three cases, despite this model's having more than double the number of parameters of ours. The heavy parameterization of the flexible multivariate GARCH model can make it impracticable for settings that include dozens to hundreds of individual time series. Our TreeDCC model can produce flexible conditional covariance specifications combined with a sufficient degree of parsimony. This is why the TreeDCC model, like the CCC, DCC, and RSDC models, can be used to estimate the conditional variance–covariance dynamics of very high-dimensional time-series settings.

4.1.3 Statistical and Economic Significance of the Improvements.
In this section, we provide additional evidence of the statistical and economic significance of the goodness-of-fit improvements provided by our model. We focus first on the statistical significance of improvements in conditional variance–covariance forecasts. To this end, we first compute Hansen

Table 2. Goodness of fit of different models for a multivariate time series of 10 daily (annualized) U.S. stock returns (in %)

U.S. equity returns: Goodness-of-fit results

Model	# par.	NL	IS-MAE	MSE	NL	OS-MAE	MSE
CCC-GARCH	80	34,959	269.413	176,147	21,694 (0.0002)	105.470 (0)	73,205.9 (0)
DCC-GARCH	82	34,942	269.525	177,848	21,652 (0.008)	103.259 (0)	72,731.0 (0)
RSDC-GARCH	173	34,684	276.604	182,999	21,634 (0.2203)	105.470 (0)	74,520.1 (0)
TreeDCC-GARCH	84	34,927	241.830	152,813	21,594 (0.857)	86.0280 (0.124)	68,831.5 (0.453)
F-MGARCH	185	34,378	188.850	135,608	21,588 (0.603)	79.6050 (0.509)	69,832.7 (0.0826)

NOTE: 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 MSE are multivariate versions of the standard univariate negative log-likelihood, the mean absolute error, and the mean squared error statistics. # par. reports the number of parameters estimated by the different models. For the out-of-sample performance measures, p-values of superior predictive ability (SPA) tests are reported in parentheses.

(2005) tests of superior predictive ability. In Table 2, p-values of these tests are reported for each model and each out-of-sample performance measure under investigation.

Consistent with our previous results, the TreeDCC and flexible multivariate GARCH models are the only ones not significantly outperformed by any other model at the 5% significance level based on the MAE and MSE statistics. With respect to the negative log-likelihood statistic, the RDSC model is also not significantly outperformed by any other model at the 20% significance level. In this case, however, the p-values implied by TreeDCC and flexible GARCH models are even higher (above 85% and 60%, respectively). To formally characterize the set of models not significantly outperformed by other ones, we also computed Hansen, Lunde, and Nason (2003) model confidence sets (MCS). Consistent with the previous findings for SPA tests, we find that the 5% MCS based on the OS-MAE and OS-MSE criteria using the range and semiquadratic statistics consist solely of the TreeDCC and flexible MGARCH models. Ten percent MCS based on the OS-NL criterion include the TreeDCC, flexible MGARCH and RSDC models.

We now economically quantify the improvements implied by the different variance–covariance forecasts and compute the difference between the variance component of the average utility AU implied by TreeDCC model and the one implied by each of the other models. Table 3 presents results for three levels of risk aversion parameters $\lambda = 2, 7, 10$ and three target expected returns (6%, 10%, and 15%) for the mean variance optimal portfolio. The resulting difference between average utilities is shown as an annualized fee and the (annualized) risk-free rate is 3%.

The economic gains relative to CCC and DCC models range from about 1 to 180 basis points. The economically most important gains arise for highly risk-averse investors ($\lambda = 10$) and high target expected returns of 15%. Economically, very

Table 3. U.S. equity returns: Annualized fees (in basis points)

Target	Alternative model	$\lambda = 2$	$\lambda = 7$	$\lambda = 10$
6%	CCC-GARCH	2.24	7.87	11.24
	DCC-GARCH	1.97	6.88	9.83
	RSDC-GARCH	0.66	2.31	3.31
	F-MGARCH	0.94	3.29	4.73
10%	CCC-GARCH	12.24	42.85	61.22
	DCC-GARCH	10.70	37.46	53.52
	RSDC-GARCH	3.60	12.59	17.99
	F-MGARCH	4.37	15.29	21.85
15%	CCC-GARCH	35.98	125.92	179.88
	DCC-GARCH	31.46	110.09	157.28
	RSDC-GARCH	10.58	37.02	52.88
	F-MGARCH	16.17	56.60	80.85

NOTE: The table contains the annualized fees (in basis points) that a conditional mean-variance investor with absolute risk-aversion parameter $\lambda = 2, 7,$ and 10 will be willing to pay to perform volatility timing using the one-step-ahead conditional correlation and covariance forecasts from the TreeDCC model (benchmark model) versus those obtained using the CCC, DCC, RSDC, and the flexible MGARCH models. The portfolio weights are obtained by minimizing the one-step-ahead conditional variance forecast of a portfolio containing 10 U.S. stocks and the risk-free asset for a given target expected return on the portfolio. The annual risk-free rate is set to be equal to 3%. Data are for the out-of-sample time period beginning in January 2004 and ending in December 2005, for a total of 504 daily observations.

small gains arise for low risk aversion ($\lambda = 2$) and a target expected return of 6%. Consistent with our previous findings, economic gains are clearly smaller relative to RSDC and flexible MGARCH models. In this case, for a high risk aversion $\lambda = 10$ and a target expected return of 15%, they are 52 and 81 basis points, respectively. We can use a Diebold–Mariano-type test and the joint test proposed in Engle and Colacito (2006) to test the statistical significance of these differences. Similarly to Bandi, Russell, and Zhu (2008), we find that these differences are not significant at the 5% significance level.

Overall, these results provide evidence that economically and statistically significant differences in variance–covariance forecasts of our TreeDCC model are likely to arise with respect to the CCC and DCC models, whereas with respect to the RSDC and flexible GARCH models, such differences are less likely.

4.1.4 Sensitivity Analysis. To investigate the robustness of our findings, it is useful to study the sensitivity of results to moderate changes in the structure of the estimated TreeDCC model. We perform this task along several dimensions.

First, we investigate whether moderate changes in the estimated location and threshold parameters of the conditional correlation process imply significantly different out-of-sample performances given a fixed tree-structure. Overall, the resulting effect on the out-of-sample performance relative to the estimated TreeDCC model is economically small. The largest impact arises by modifying the location parameter c_4 in the fourth regime (of ± 1 standard error) or the first threshold parameter d_1 (to the nearest possible threshold value in the positive and negative directions). However, the changes in forecasting accuracy are typically smaller than the changes in MAE and MSE observed for the CCC, DCC, RSDC, and flexible GARCH models.

Second, we investigate whether moderate changes in the tree-structure of the estimated TreeDCC model imply significantly different out-of-sample performances. We consider the second, third, and fourth best (in-sample) TreeDCC models estimated according to the procedure described in Section 3. Once again, the good forecasting performance of our TreeDCC modeling approach relative to other models seems to be quite robust across various similar choices of the parameters and threshold structure in the model. More detailed results of this sensitivity exercise are available from the authors upon request.

4.2 Second Real Data Application: U.S. Stock Index and Bond Returns

We consider a two-dimensional time series of (annualized) daily log-returns for the U.S. S&P500 stock index and the U.S. 30-year Treasury bond. The time period under investigation goes from January 3, 1996 to October 30, 2003 and contains 1899 trading days. The data are 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 subperiods. The first subperiod consists of $n = 1219$ trading days and goes from January 3, 1996 to December 29, 2000. The second subperiod consists of the last three years of data ($n_{\text{out}} = 680$ observations).

Table 4. Estimation results for a two-dimensional time series of daily (annualized) returns (in %) for the U.S. S&P500 index and the U.S. 30-year Treasury bond

Panel A: Individual conditional variance structures

Series	Regimes	Optimal predictors
S&P500	3	S&P500
30-year Treasury bond	2	30-year Treasury bond

Panel B: Conditional correlation structure and parameters

Cond. corr. structure $\mathcal{R}_k$	Cond. corr. parameters	
	ϕ_k	λ_k
$X_{t-1,S\&P500} \leq -3.847098$	0.0490 (0.001)	0.9129 (0.024)
$X_{t-1,S\&P500} > -3.847098$	0.0222 (0.002)	0.9724 (0.019)

NOTE: 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. Standard errors computed using 1000 model-based bootstrap replications are given in parentheses.

4.2.1 Estimation Results.

The estimation procedure follows the same steps as the one described in Section 4.1. The individual variance structures and correlation threshold functions estimated for our TreeDCC model are summarized in Table 4.

The estimated threshold functions for volatility each depend only on one lagged return of each time series of returns. Similarly to previous findings in the literature, e.g., Audrino and Trojani (2006), we find that the conditional variance of stock index returns implies more than two regimes. The estimated conditional variance of Treasury bond returns implies only two regimes.

We estimate only two regimes as well for the estimated conditional correlation function of stock index and Treasury bond returns. These regimes depend on the lagged return of the S&P500 index only. Each regime features local GARCH-type DCC effects as in Engle (2002), in which the regime-dependent parameters ϕ_k and λ_k imply a different persistence of correlation shocks: Here the lagged S&P500 index return is below or

above the threshold value $d_1 = -3.847$. This threshold corresponds to the 37.5% quantile of the distribution of S&P500 index returns. The differences in the estimated parameters of the local DCC dynamics are statistically significant. Interestingly, we find that correlation shocks are quite substantially more persistent, conditional on a sufficiently negative past stock index return ($X_{t-1,S\&P500} \leq -3.847$). This might be interpreted as an indication that correlation shocks between bond and stock returns are likely to last longer conditional on flight-to-quality effects caused by a drop in the stock market. It is also interesting to note that the correlation dynamics estimated for stock index and bond returns are quite different from those estimated earlier in our application to a 10-dimensional stock returns time series. The flexibility of our TreeDCC setting is crucial for allowing us to take these different dynamic correlation features of some asset returns adequately into account.

4.2.2 Multivariate Performance Results.

Table 5 presents results for the goodness-of-fit measures in Section 4, estimated using stock index and Treasury bond return data.

The TreeDCC model clearly has the best goodness-of-fit results across all in-sample and out-of-sample measures used. The improvements in out-of-sample performance range from 1% to 25%, depending on the goodness-of-fit measure applied. Constant or piecewise-constant conditional correlation dynamics (like the ones implied by the CCC and RSDC models) are largely rejected by the data and lead to very inaccurate correlation forecasts. In contrast to the previous application to single stock returns, GARCH-type DCC effects are now crucial in order to improve the model's out-of-sample forecasting power for correlations.

4.2.3 Statistical and Economic Significance of the Improvements.

In Table 5, p-values of Hansen (2005) SPA tests are reported. They show that the TreeDCC model is the only one not significantly outperformed by any other model at the 10% confidence level and for all performance criteria used.

The Engle-DCC model yields quite good results and is not outperformed by any model at the 5% confidence level. All other models are clearly dominated at standard significance levels. MCS results consistently support these findings. Ten percent MCS using the range and the semiquadratic statistics consist only of the TreeDCC model for all performance criteria used. Additionally, 5% MCS include the DCC model.

Table 5. U.S. index and bond returns: Goodness-of-fit results

Model	# par.	NL	IS-MAE	MSE	NL	OS-MAE	MSE
CCC-GARCH	25	9334.7	63.6310	21,197.8	5511.5 (0.0003)	97.2933 (0)	27,592.1 (0)
DCC-GARCH	27	9299.2	59.4418	20,458.1	5451.7 (0.4469)	74.8853 (0.0868)	22,357.5 (0.3319)
RSDC-GARCH	30	9334.7	63.6310	21,197.8	5511.5 (0.0014)	97.2933 (0)	27,592.0 (0)
TreeDCC-GARCH	29	9290.2	58.7296	20,413.5	5440.9 (0.8435)	70.7387 (0.2711)	20,544.6 (0.2246)
F-MGARCH	13	9389.2	65.0761	22,752.7	5483.4 (0.0012)	86.1455 (0.0007)	28,271.6 (0.0652)

NOTE: Goodness of fit of different models for a two-dimensional time series of daily (annualized) returns (in %) on the U.S. S&P500 index and the 30-year 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 MSE are the multivariate versions of the standard univariate negative log-likelihood, the mean absolute error, and the mean squared error statistics. # par. reports the number of parameters estimated in the different models. For the out-of-sample performance measures, p-values of SPA tests are reported in parentheses.

Table 6. U.S. index and bond returns: Annualized fees
(in basis points)

Target	Alternative model	$\lambda = 2$	$\lambda = 7$	$\lambda = 10$
6%	CCC-GARCH	5.54	19.38	27.68
	DCC-GARCH	0.54	1.87	2.68
	RSDC-GARCH	5.54	19.38	27.68
	F-MGARCH	3.68	12.89	18.41
10%	CCC-GARCH	30.14	105.50	150.72
	DCC-GARCH	2.92	10.21	14.59
	RSDC-GARCH	30.14	105.50	150.72
	F-MGARCH	20.05	70.18	100.26
15%	CCC-GARCH	88.58	310.05	442.92
	DCC-GARCH	8.57	29.99	42.85
	RSDC-GARCH	88.58	310.05	442.92
	F-MGARCH	58.93	206.25	294.64

NOTE: The table contains the annualized fees (in basis points) that a conditional mean–variance investor with absolute risk-aversion parameter $\lambda = 2$, 7, and 10 will be willing to pay to perform volatility timing using the one-step-ahead conditional correlation and covariance forecasts from the TreeDCC model (benchmark model) versus those obtained using the CCC, DCC, RSDC, and the flexible MGARCH models. The portfolio weights are obtained by minimizing the one-step-ahead conditional variance forecast of a portfolio containing the U.S. S&P500 index, the 30-year U.S. Treasury bond, and the risk-free asset for a given target expected return on the portfolio. The annual risk-free rate is set to be equal to 3%. Data are for the out-of-sample time period beginning in January, 2001 and ending in October, 2003, for a total of 680 daily observations.

We conclude this section by economically quantifying the forecast improvements of our TreeDCC model relative to the competing ones. Table 6 presents the annualized fees (in basis points) that a conditional mean–variance investor will be ready to pay in order to forecast future variance–covariance matrices with the TreeDCC instead of the other models considered. The (annualized) risk-free rate is equal to 3%.

The largest economic gains arise with respect to the CCC and RSDC models: For risk aversion parameters $\lambda \geq 7$ and target expected returns larger than 10%, they range from approximately 100 to 440 basis points per year. Sizable economic gains between 70 and 294 basis points for risk aversion parameters $\lambda \geq 7$ and target expected returns larger than 10% also arise with respect to the flexible MGARCH model. Economic gains relative to the DCC model are small in almost all cases and never exceed 43 basis points per year. Using Diebold–Mariano-type tests we find that estimated economic gains of the TreeDCC are significant at the 1% significance level relative to the CCC and RSDC model. Estimated economic gains relative to the flexible MGARCH model are significant at the 5% significance level. Improvements relative to the classical DCC model are not statistically significant. Overall, these results provide evidence that economically and statistically significant differences in variance–covariance forecasts of our TreeDCC model are likely to arise with respect to the CCC, RSDC, and flexible GARCH models, and less likely relative to the DCC model, which was clearly dominated by the TreeDCC model in the previous empirical application.

5. CONCLUSION

We propose a new multivariate DCC-GARCH model that extends previous models by admitting multivariate thresholds in conditional volatilities and correlations. The thresholds are modeled by a tree-structured partition of the multivariate state space and are estimated with all other model parameters. Two real data applications support the overall higher forecasting power of the TreeDCC model for return correlations, relative to Bollerslev's CCC model, Engle's DCC model, Pelletier's RSDC model, and the flexible MGARCH model. We find that the conditional correlations of financial data are often characterized by multivariate thresholds and local GARCH-type structures and that the forecast improvements of our TreeDCC model are often economically relevant. Our model can cope in a parsimonious way with such features of the data even in applications with large cross sections of financial assets. An interesting avenue for future research is the joint empirical modeling of the dynamic correlation of the returns of several asset classes, which are likely to exhibit rich threshold and GARCH-type effects that can be parsimoniously taken into account by our model.

APPENDIX: TESTING STATISTICAL RELEVANCE: THE MODEL CONFIDENCE SET

We formally test for differences in the forecasting power of the competing models in order to select, if possible, a best one (or a best subset of models) that significantly dominates the others in our real data application. To this end, we apply the model confidence set (MCS) method proposed by Hansen, Lunde, and Nason (2003).

Without loss of generality, let us denote by $\widehat{D}_{t,wk}$ the differences of each term in the OS-MSE statistic

$$\widehat{D}_{t,wk} = \widetilde{U}_{t;\text{model}_w} - \widetilde{U}_{t;\text{model}_k},$$

$$t = 1, \ldots, n_{\text{out}}, w, k = 1, \ldots, 9, w < k,$$

where

$$\sum_{t=1}^{n_{\text{out}}} \widetilde{U}_{t;\text{model}} = \text{OS-MSE}.$$

Statistics based on time averages $\overline{D}_{wk}$ of $\widehat{D}_{t,wk}$ allow us to investigate whether there is a systematic difference in the out-of-sample forecasting power between the different models. Tests based on $\widehat{D}_{t,wk}$ are t-type tests. In a similar way, one can proceed by using a different out-of-sample goodness-of-fit statistic. In our application, we compute tests based on OS-MSE, OS-MAE, and OS-NL.

The MCS is defined as the smallest set of models, which at a given confidence level α, cannot be significantly distinguished based on forecasting power. The MCS is determined after sequentially trimming the set of candidate models, which in our application consists of the five multivariate GARCH specifications introduced earlier. At each step of such a trimming procedure, the null-hypothesis of equal predictive ability (EPA) $\mathcal{H}_0 : \mathbb{E}[D_{t,wk}] = 0, \forall w, k \in \mathcal{M}$, is tested for the relevant set of models $\mathcal{M}$ at a confidence level α. In the first step, $\mathcal{M}$ consists of all models under investigation. If, in the first step, $\mathcal{H}_0$ is rejected, then the worst-performing model according to the relevant criterion is eliminated. The test procedure is then repeated for the new set $\mathcal{M}$ of surviving models and it is iterated until the first nonrejection of the EPA hypothesis occurs. The set of resulting models is called the model confidence set $\widehat{\mathcal{M}}_\alpha$ at

the given confidence level α. In our application, we work with $\alpha = 0.05, 0.10$.

Our tests of EPA are based on the range statistic T_R and the less conservative semiquadratic statistic T_{SQ}

$$T_R = \max_{k,w \in \mathcal{M}} \frac{|\overline{D}_{kw}|}{\sqrt{\widehat{\text{var}}(\overline{D}_{kw})}} \quad \text{and} \quad T_{SQ} = \sum_{k<w} \frac{\overline{D}_{kw}^2}{\widehat{\text{var}}(\overline{D}_{kw})},$$

where the sum in T_{SQ} is taken over the models in $\mathcal{M}$, $\overline{D}_{kw} = n_{out}^{-1} \sum_{t=1}^{n_{out}} \widehat{D}_{t,kw}$, and $\widehat{\text{var}}(\overline{D}_{kw})$ is an estimate of $\text{var}(\overline{D}_{kw})$ obtained from a block-bootstrap of the series $\widehat{D}_{t,kw}$, $t = 1, \ldots, n_{out}$. Using statistics T_R or T_{SQ}, we test the null hypothesis EPA at confidence level α for model set $\mathcal{M}$. If the hypothesis EPA is rejected for model set $\mathcal{M}$, we compute a worst-performing index in order to trim the worst-performing model from $\mathcal{M}$.

The worst-performing index for Model$_k$ is computed as the mean across models $w \neq k$ of statistic $\overline{D}_{wk}$. More specifically, it is defined as $\overline{D}_k/\sqrt{\text{var}(\overline{D}_k)}$, where $\overline{D}_k = \text{mean}_{w \neq k \in \mathcal{M}} \overline{D}_{kw}$. As earlier, our estimate of $\text{var}(\overline{D}_k)$ is based on a block-bootstrap. The model with the highest worst-performing index is finally trimmed from $\mathcal{M}$. Consistency of estimates of the (asymptotic) distributions of T_R and T_{SQ} can be proved under mild regularity conditions on the bootstrap. For details, see Hansen, Lunde, and Nason (2003).

ACKNOWLEDGMENTS

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. [143]

—— (2003), "Modeling and Forecasting Realized Volatility," *Econometrica*, 71 (2), 579–625. [139,143]

Audrino, F. (2006), "Tree-Structured Multiple Regimes in Interest Rates," *Journal of Business & Economic Statistics*, 24 (3), 338–353. [138]

Audrino, F., and Bühlmann, P. (2001), "Tree-Structured GARCH Models," *Journal of the Royal Statistical Society, Ser. B*, 63 (4), 727–744. [138-142]

Audrino, F., and Corsi, F. (2009), "Modeling Tick-by-Tick Realized Correlations," working paper, University of St. Gallen. [139]

Audrino, F., and Trojani, F. (2006), "Estimating and Predicting Multivariate Volatility Regimes in Global Stock Markets," *Journal of Applied Econometrics*, 21 (3), 345–369. [138-142,144,147]

Bandi, F., Russell, J. R., and Zhu, Y. (2008), "Using High-Frequency Data in Dynamic Portfolio Choice," *Econometric Reviews*, 27, 163–198. [139,144, 146]

Barndorff-Nielsen, O. E., and Shephard, N. (2001), "Non-Gaussian Ornstein–Uhlembech-Based Models and Some of Their Uses in Financial Economics," *Journal of the Royal Statistical Society, Ser. B*, 63, 167–241. [143]

—— (2002), "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, Ser. B*, 64, 253–280. [143]

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008), "Multivariate Realised Kernels: Consistent Positive Semi-Definite Estimators of the Covariation of Equity Prices With Noise and Non-Synchronous Trading," working paper, Stanford University. [139]

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. [143]

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984), *Classification and Regression Trees*, Belmont, CA: Wadsworth. [142]

Chiriac, R., and Voev, V. (2009), "Modelling and Forecasting Multivariate Realized Volatility," working paper, Aarhus University. [139]

Corsi, F., and Audrino, F. (2007), "Realized Correlation Tick-by-Tick," discussion paper, University of St. Gallen, Dept. of Economics. [143,144]

Engle, R. F. (2002), "Dynamic Conditional Correlation—A Simple Class of Multivariate GARCH Models," *Journal of Business & Economic Statistics*, 20, 339–350. [138,140,143,147]

Engle, R., and Colacito, R. (2006), "Testing and Valuing Dynamic Correlations for Asset Allocation," *Journal of Business & Economic Statistics*, 24, 238–253. [139,144,146]

Engle, R. F., and Sheppard, K. (2001), "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH," mimeo, University of California, San Diego. [142]

Fleming, J., Kirby, C., and Ostdiek, B. (2001), "The Economic Value of Volatility Timing," *Journal of Finance*, 56, 329–352. [143,144]

—— (2003), "The Economic Value of Volatility Timing Using Realized Volatility," *Journal of Financial Economics*, 67, 473–509. [144]

Hansen, P. R. (2005), "A Test for Superior Predictive Ability," *Journal of Business & Economic Statistics*, 23, 365–380. [139,143,145-147]

Hansen, P. R., Lunde, A., and Nason, J. M. (2003), "Choosing the Best Volatility Models: The Model Confidence Set Approach," *Oxford Bulletin of Economics and Statistics*, 65, 839–861. [139,143,146,148,149]

Laurent, S., Rombouts, J. V. K., and Violante, F. (2009), "Consistent Ranking of Multivariate Volatility Models," discussion paper, CORE. [143]

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 (3), 735–747. [138,143]

Newey, W. K., and McFadden, D. (1994), "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics*, Vol. 4, Amsterdam: North-Holland, pp. 2111–2245. [142]

Pagan, A. (1986), "Two Stage and Related Estimators and Their Applications," *The Review of Economic Studies*, 53 (4), 517–538. [142]

Pelletier, D. (2006), "Regime Switching for Dynamic Correlations," *Journal of Econometrics*, 131, 445–473. [138,140,142,143]

West, K. D., Edison, H. J., and Cho, D. (1993), "A Utility-Based Comparison of Some Models of Exchange Rate Volatility," *Journal of International Economics*, 35, 23–45. [143]