Markus Haas[1] / Ji-Chun Liu[2]

A multivariate regime-switching GARCH model with an application to global stock market and real estate equity returns

[1] University of Kiel, Institute for Quantitative Business and Economics Research (QBER), Kiel, Germany, E-mail: Markus.haas@qber.uni-Kiel.de

[2] Xiamen University, School of Mathematical Sciences, Xiamen, China

Abstract:

We consider a multivariate Markov-switching GARCH model which allows for regime-specific volatility dynamics, leverage effects, and correlation structures. Conditions for stationarity and expressions for the moments of the process are derived. A Lagrange Multiplier test against misspecification of the within-regime correlation dynamics is proposed, and a simple recursion for multi-step-ahead conditional covariance matrices is deduced. We use this methodology to model the dynamics of the joint distribution of global stock market and real estate equity returns. The empirical analysis highlights the importance of the conditional distribution in Markov-switching time series models. Specifications with Student's t innovations dominate their Gaussian counterparts both in- and out-of-sample. The dominating specification appears to be a two-regime Student's t process with correlations which are higher in the turbulent (high-volatility) regime.

Keywords: conditional volatility, covariance forecasts, Markov-switching, multivariate GARCH

DOI: 10.1515/snde-2016-0019

1 Introduction

Asset return distributions are typically characterized by fat tails, conditional heteroskedasticity, and nonlinear dependence. Regarding the latter, a frequent concern is that the dependence between assets increases in periods of market turbulence. This has important implications for portfolio and risk management, because it means that "the benefits of diversification are partly eroded when they are needed most" (Campbell, Koedijk & Kofman, 2002). An overview over the extensive literature studying this phenomenon and further evidence is provided, e.g. by Kasch and Caporin (2013), Mittnik (2014), and Gribisch (2016). For example, Kasch and Caporin (2013) develop a multivariate GARCH model with dynamic correlations being allowed to depend on conditional volatility and, for major stock markets, find that "turbulent periods coincide with an increase in cross-market comovement."

Markov-switching models (MSMs) are able to capture all of the aforementioned *stylized facts* of asset return distributions, and their use is very popular in financial modeling because, in addition to their flexibility, "the idea of regime changes is natural and intuitive" (Ang & Timmermann, 2012). For example, in bearish markets, expected returns, conditional volatilities and their dynamics, and correlations can differ from their respective counterparts in more normal or bullish market periods. Regime-specific dynamics may also be related to various types of trading patterns, as represented by "information" and "feedback" traders (Dean & Faff, 2008). See Guidolin (2011) and Ang and Timmermann (2012) for an overview over the many applications of MSMs.

In this paper, we investigate the properties of a multivariate extension of the Markov-switching (MS) GARCH model of Haas, Mittnik, and Paolella (2004b), allowing for regime-specific volatility dynamics, leverage effects, and correlation structures. We derive conditions for strict and weak stationarity and provide expressions for the unconditional overall and regime-specific covariance and correlation matrices, and for the dynamic correlation structure of the absolute values of the process. The model we consider is an extension of Pelletier's (2006) regime-switching model for dynamic correlations (RSDC) in that it combines constant conditional within-regime correlations with regime-dependent conditional variance dynamics. Among other convenient features, this property allows the derivation of a simple recursion for multi-step-ahead conditional covariance matrices for, e.g. mean-variance portfolio allocation. Alternative extensions of Pelletier (2006) have been proposed by Billio and Caporin (2005) and Otranto (2010), who allow for within-regime correlation dynamics

---

Markus Haas is the corresponding author.

à la Engle (2002), but do not allow for switching variance dynamics. To test whether the assumption of constant conditional within-regime correlations is justified empirically, we construct a test against within-regime conditional correlation dynamics, adopting Hamilton's (1996) Lagrange Multiplier (LM) framework along with Tse's (2000) test for constant conditional correlations in a single-regime GARCH model. The methodology is applied to global stock market and real estate equity returns. The empirical analysis highlights the importance of the conditional distribution in MS time series models. Namely, since the conditional distribution in MSMs with Gaussian regimes is a discrete mixture of normals and thus already thick-tailed, one might guess that use of a more flexible within-regime distribution is unnecessary in this framework. Indeed, as observed by Guidolin (2011), "it seems that most authors are still finding that traditional Gaussian mixture models are generally sufficient to the task assigned to MSMs." However, in our application, specifications with Student's t innovations dominate their Gaussian counterparts both in– and out-of-sample. In particular, as discussed in Section 4, the Gaussian specification turns out to suffer from its inability to correctly track the regime-switching process. The dominating specification appears to be a two-regime Student's t process with correlations which are higher in the turbulent (high-volatility) regime.

The structure of the paper is as follows. In Section 2, we define the model and discuss its relation to the literature. Statistical properties are presented in Section 3. Section 4 provides an application to financial data, and Section 5 concludes. Proofs of theorems, calculations of (conditional and unconditional) moments, and details of the LM test against misspecification of conditional correlations are gathered in Appendices A, B, and C, respectively. Finally, Appendix D provides a brief discussion of the asymmetric multivariate normal mixture GARCH model of Bauwens, Hafner, and Rombouts (2007) and Haas, Mittnik, and Paolella (2009), which for comparison purposes has been included in the empirical application in Section 4.3.

2   Definition of the process

The multivariate Markov-switching (MS) GARCH process introduced in this section generalizes the univariate model proposed in Haas, Mittnik, and Paolella (2004b). For alternative approaches to MS GARCH, see, e.g. Gray (1996), Dueker (1997), Klaassen (2002), Augustyniak (2014), and Reher and Wilfling (2016), as well as the review in Haas and Paolella (2012). Liu (2007) extended the model of Haas, Mittnik, and Paolella (2004b) to allow for an asymmetric response of volatility to positive and negative shocks, which is also incorporated in the model discussed herein.

Let the M–dimensional time series $\{\boldsymbol{\epsilon}_t\}$ satisfy

$$\boldsymbol{\epsilon}_t = \boldsymbol{D}_{\Delta_t,t} \cdot \boldsymbol{z}_t, \tag{1}$$

where $\{\Delta_t\}$ is a Markov chain with finite state space $\mathcal{E} = \{1, ..., k\}$ and irreducible and aperiodic transition matrix $\boldsymbol{P}$,

$$\boldsymbol{P} = \begin{pmatrix} p_{11} & \cdots & p_{k1} \\ \vdots & \cdots & \vdots \\ p_{k1} & \cdots & p_{kk} \end{pmatrix}, \tag{2}$$

where the transition probabilities $p_{ij} = p(\Delta_t = j \,|\, \Delta_{t-1} = i)$, $i, j \in \mathcal{E}$, and the stationary distribution of the chain is denoted by $\boldsymbol{\pi}_{\infty} = (\pi_{1,\infty}, \pi_{2,\infty}, ..., \pi_{k,\infty})'$. Matrix $\boldsymbol{D}_{\Delta_t,t} = \mathrm{diag}(\boldsymbol{\sigma}_{\Delta_t,t})$, where $\boldsymbol{\sigma}_{jt} = (\sigma_{1jt}, ..., \sigma_{Mjt})' \in \mathbb{R}^M$, $j \in \mathcal{E}$, contains the regime-specific conditional standard deviations of the elements of $\boldsymbol{\epsilon}_t$. Moreover,

$$\boldsymbol{z}_t = \boldsymbol{R}_{\Delta_t}^{1/2} \boldsymbol{\zeta}_t, \tag{3}$$

where $\boldsymbol{R}_j = (\rho_{\ell m,j})_{\ell,m=1,...,M}$, $j = 1, ..., k$, is a (regime-specific) correlation matrix, and $\{\boldsymbol{\zeta}_t\}$ is a sequence of iid random vectors with zero mean and identity covariance matrix. In the applications below, we assume that $\boldsymbol{\xi}_t$ has a Student's t distribution with $\nu > 2$ degrees of freedom, with density given by (29) in Appendix C, i.e.

$$\boldsymbol{\zeta_t} \overset{iid}{\sim} t(\mathbf{0}, \boldsymbol{I}_M, \nu), \tag{4}$$

which includes normality as a limiting case ($\nu \to \infty$). It is assumed that $\{\Delta_t\}$ and $\{\boldsymbol{\zeta}_t\}$ are independent.

The regime-specific conditional standard deviations follow simultaneous asymmetric absolute value GARCH(1,1) (AGARCH) processes, i.e.

$$\begin{aligned}
\boldsymbol{\sigma}_{jt} &= \boldsymbol{\omega}_j + \boldsymbol{A}_j|\boldsymbol{\epsilon}_{t-1}| - (\boldsymbol{A}_j \odot \boldsymbol{\Gamma}_j)\boldsymbol{\epsilon}_{t-1} + \boldsymbol{B}_j\boldsymbol{\sigma}_{j,t-1} \\
&= \boldsymbol{\omega}_j + (\boldsymbol{A}_j|\boldsymbol{Z}_t| - (\boldsymbol{A}_j \odot \boldsymbol{\Gamma}_j)\boldsymbol{Z}_t)\boldsymbol{\sigma}_{\Delta_{t-1},t-1} + \boldsymbol{B}_j\boldsymbol{\sigma}_{j,t-1} \quad j \in \mathcal{E},
\end{aligned} \tag{5}$$

where $\boldsymbol{Z}_t = \mathrm{diag}(\boldsymbol{z_t})$, a matrix in absolute value bars means that the absolute value of each element is taken, $\boldsymbol{\omega}_j = (\omega_{1j}, ..., \omega_{Mj})'$, and

$$\boldsymbol{A}_j = [a_{\ell m,j}]_{\ell,m=1,...,M}, \quad \boldsymbol{\Gamma}_j = [\gamma_{\ell m,j}]_{\ell,m=1,...,M}, \quad \boldsymbol{B}_j = [b_{\ell m,j}]_{\ell,m=1,...,M}, \quad j \in \mathcal{E}. \tag{6}$$

Parameters $\gamma_{\ell m,j} \in [-1, 1]$, $\ell, m = 1, ..., M, j \in \mathcal{E}$, allow the conditional standard deviations to react asymmetrically to positive and negative news of the same magnitude as in Ding, Granger, and Engle (1993). Equivalently, as in McAleer, Hoti, and Chan (2009) and Francq and Zakoïan (2012), we could write the asymmetric volatility process à la Glosten, Jagannathan, and Runkle (1993) as

$$\boldsymbol{\sigma}_{jt} = \boldsymbol{\omega}_j + \boldsymbol{A}_j^+\boldsymbol{\epsilon}_{t-1}^+ + \boldsymbol{A}_j^-|\boldsymbol{\epsilon}_{t-1}^-| + \boldsymbol{B}_j\boldsymbol{\sigma}_{j,t-1}, \quad j \in \mathcal{E}, \tag{7}$$

where $x^+ = \max\{x, 0\}$, $x^- = \min\{0, x\}$, and

$$\boldsymbol{\epsilon}_t^+ = (\epsilon_{1t}^+, ..., \epsilon_{Mt}^+)', \quad \boldsymbol{\epsilon}_t^- = (\epsilon_{1t}^-, ..., \epsilon_{Mt}^-)'.$$

Denoting the typical elements of matrices $\boldsymbol{A}_j^+$ and $\boldsymbol{A}_j^-$ in (20) by $a_{\ell m,j}^+$ and $a_{\ell m,j}^-$, respectively, the relation between the parameters in (18) and (20) is

$$a_{\ell m,j}^+ = a_{\ell m,j}(1 - \gamma_{\ell m,j}), \quad a_{\ell m,j}^- = a_{\ell m,j}(1 + \gamma_{\ell m,j}).$$

The model defined by (14)–(19) will be referred to as a k–component Markov-switching constant conditional correlation GARCH process, or, in short, MS(k) CCC-GARCH. It is an asymmetric multi-regime version of the extended CCC (ECCC) model studied by Jeantheau (1998), which itself generalizes the CCC of Bollerslev (1990) by allowing for volatility interactions, which are often of interest in finance and macroeconomics (e.g. Nakatani & Teräsvirta, 2009; and Conrad and Karanasos 2010, 2015). In many applications the diagonal model, with all $\boldsymbol{A}_j, \boldsymbol{B}_j$, and $\boldsymbol{\Gamma}_j$ being diagonal matrices, will be preferred for reasons of parsimony; an ARCH version of such a model was used by Ramchand and Susmel (1998).

The specification of the volatility dynamics (18) in terms of the conditional standard deviation instead of the conditional variance, as originally proposed by Taylor (1986), serves two purposes: First, empirically, it appears that it typically improves the fit as compared to the formulation in terms of the conditional variance, and is very often close to the MLE when the "power parameter" (as in Ding, Granger & Engle, 1993) is freely estimated from the data (e.g. Giot & Laurent, 2003; Lejeune, 2009; and Broda et al., 2013). Second, as noted by Pelletier (2006), this specification allows for closed-form calculation of multi-step-ahead conditional covariance matrices, as required, e.g. for mean-variance portfolio optimization over horizons longer than one period. The model suggested by Pelletier (2006), referred to as the regime-switching dynamic correlation (RSDC) model, is nested in (14)–(19) when only the conditional correlation matrices are subject to regime-switching, i.e. in (18), $\boldsymbol{\omega}_1 = \cdots = \boldsymbol{\omega}_k$, $A_1 = \cdots = A_k$, and $\boldsymbol{B_1} = \cdots = \boldsymbol{B_k}$. Covariance matrix forecasts for the RSDC are considered in Pelletier (2006) and Haas (2010), whereas a convenient scheme for forecasting the general model (14)–(19) is derived in Appendix B.3.

In (18), conditions have to be imposed to make sure that all elements of $\boldsymbol{\sigma}_{jt}$ remain positive with probability 1, $j = 1, ..., k$. As observed by He and Teräsvirta (2004), an obvious set of sufficient conditions is that

$$\boldsymbol{\omega}_j > 0 \quad \text{and} \quad \boldsymbol{A}_j, \boldsymbol{B}_j \geq 0 \quad \text{elementwise,} \tag{8}$$

but these are not necessary when diagonality is not imposed (Nakatani & Teräsvirta, 2008; Conrad & Karanasos, 2010). For the diagonal model, which is of particular importance in the applications, conditions (21) are necessary, however.

Regarding the distribution of the innovations $\{\boldsymbol{\xi}_t\}$, note that (17) includes Gaussian innovations as a limiting case, when $\nu \to \infty$. Though normality is still the dominant distributional assumption in regime-switching models (cf. Guidolin, 2011), allowing for fat-tailed innovations can improve both in-sample fit and out-of-sample forecasting performance of MS GARCH models, as pointed out, e.g. by Klaassen (2002), Ardia (2009), and Shi and Feng (2016). Due to the dependence structure implied by the multivariate t distribution, this also holds for Pelletier's (2006) model where volatility is regime-independent; see Section 4 for a detailed discussion and illustration.

3 Properties of the model: stationarity and moment structure

In this section, we discuss the statistical properties of the MS(k) CCC-GARCH process. In particular, we present conditions for strict stationarity and ergodicity and the existence of unconditional moments, with proofs of the theorems given in Appendix A. Explicit formulas for moments of frequent interest are provided (and derived) in Appendix B, namely the unconditional covariance matrix and the autocorrelations of the absolute values, which can be used to characterize the joint volatility dynamics. Moreover, a simple recursive scheme for obtaining multi-step-ahead covariance matrices is derived in Appendix B.3, fostering applications to mean-variance portfolio selection in environments with changing volatilities and correlations.

To set out the properties of the MS(k) CCC-GARCH process, we define the matrices

$$\boldsymbol{X}_t = \begin{pmatrix} \boldsymbol{\sigma}_{1t} \\ \vdots \\ \boldsymbol{\sigma}_{kt} \end{pmatrix}, \quad \boldsymbol{\omega} = \begin{pmatrix} \boldsymbol{\omega}_1 \\ \vdots \\ \boldsymbol{\omega}_k \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 \\ \vdots \\ \boldsymbol{A}_k \end{pmatrix}, \quad \tilde{\boldsymbol{A}} = \begin{pmatrix} \boldsymbol{A}_1 \odot \boldsymbol{\Gamma}_1 \\ \vdots \\ \boldsymbol{A}_k \odot \boldsymbol{\Gamma}_k \end{pmatrix},$$

and $\boldsymbol{B} = \text{blockdiag}(\boldsymbol{B}_1, ..., \boldsymbol{B}_k)$. This gives rise to the representation

$$\boldsymbol{X}_t = \boldsymbol{\omega} + \boldsymbol{C}_{\Delta_{t-1}, t-1} \boldsymbol{X}_{t-1}, \tag{9}$$

where

$$\boldsymbol{C}_{\Delta_t, t} = (\boldsymbol{A}|\boldsymbol{Z}_t| - \tilde{\boldsymbol{A}}\boldsymbol{Z}_t)(\boldsymbol{e}'_{\Delta_t} \otimes \boldsymbol{I}_M) + \boldsymbol{B}, \tag{10}$$

and $\boldsymbol{e}_j$ is the jth unit vector in $\mathbb{R}^k$, $j = 1, ..., k$.

We first provide a necessary and sufficient condition for the existence of a strictly stationary solution of the MS(k) CCC-GARCH process with the nonnegativity conditions (21) imposed. Theorem 1 generalizes results for the univariate MS GARCH process in Liu (2006 and 2007).[1]

Theorem 1

The MS(k) CCC-GARCH(1,1) process defined by (14)–(19) has a unique strictly stationary and ergodic solution if and only if the top Lyapunov exponent γ_C associated to the random matrices $\{\boldsymbol{C}_{\Delta_t, t}\}$ is strictly negative. Moreover, this stationary solution is explicitly expressed as

$$\boldsymbol{\epsilon}_t = \left[\text{diag}\left((\boldsymbol{e}'_{\Delta_t} \otimes \boldsymbol{I}_M)\left(\boldsymbol{\omega} + \sum_{n=1}^{\infty} \boldsymbol{C}_{\Delta_{t-1}, t-1} \boldsymbol{C}_{\Delta_{t-2}, t-2} ... \boldsymbol{C}_{\Delta_{t-n+1}, t-n+1} \boldsymbol{\omega} \right) \right) \right]^{1/2} \boldsymbol{R}_{\Delta_t}^{1/2} \boldsymbol{\xi}_t.$$

The condition in Theorem 1 may be inconvenient to check in practice. Theorem 2 offers an alternative criterion which is easier to handle and provides additional information about the moment structure of the process. To state this criterion, we define the matrices

$$\begin{aligned} \boldsymbol{C}_1(j) &= \text{E}(\boldsymbol{C}_{jt}|\Delta_t = j), \quad \boldsymbol{C}_2(j) = \text{E}(\boldsymbol{C}_{jt} \otimes \boldsymbol{C}_{jt}|\Delta_t = j), \quad ..., \\ \boldsymbol{C}_l(j) &= \text{E}(\boldsymbol{C}_{jt}^{\otimes l}|\Delta_t = j), \quad j \in \mathscr{E}, \quad l \in \mathbb{N}. \end{aligned} \tag{11}$$

Furthermore, we adopt the following notation from Francq and Zakoïan (2005): For any function $f: \mathscr{E} \mapsto M_{n \times n'}(\mathbb{R})$, where $M_{n \times n'}(\mathbb{R})$ is the space of real $n \times n'$ matrices, and $\mathscr{E} = \{1, ..., k\}$ is the state space of $\{\Delta_t\}$, define the matrix

$$\mathbb{P}_f = \begin{pmatrix} p_{11}f(1) & \cdots & p_{k1}f(1) \\ \vdots & \cdots & \vdots \\ p_{1k}f(k) & \cdots & p_{kk}f(k) \end{pmatrix}. \tag{12}$$

Theorem 2

Suppose that the l-th moments of (ξ_t) are finite and

$$\lambda(\mathbb{P}_{C_l}) < 1,$$

where $\lambda(\mathbb{P}_{C_l})$ denotes the spectral radius of $\mathbb{P}_{C_l}$ defined in (25), and l is a strictly positive integer. Then (14)–(19) has a unique strictly stationary and ergodic solution (ϵ_t), and the l-th absolute moments of (ϵ_t) are finite.

For example, the matrices required by Theorem 2 to check for the first moment are given by

$$\boldsymbol{C}_1(j) = \kappa_1 \boldsymbol{A}(\boldsymbol{e}_j' \otimes \boldsymbol{I}_M) + \boldsymbol{B},$$

where

$$\kappa_1 = \mathrm{E}(|z_{it}|) = \begin{cases} \sqrt{\frac{2}{\pi}} & \text{if } z_{it} \sim \mathrm{N}(0,1) \\ \frac{\sqrt{\nu-2}\,\Gamma\left(\frac{\nu-1}{2}\right)}{\sqrt{\pi}\,\Gamma(\nu/2)} & \text{if } z_{it} \sim t_\nu(0,1). \end{cases}$$

To check the condition for covariance stationarity, we need the (regime-specific) second moment matrices of the absolute innovations, i.e.

$$\tilde{\boldsymbol{R}}_j := \mathrm{E}(|\mathbf{z}_t \mathbf{z}_t'| \,|\Delta_t = j), \quad j \in \mathscr{E},$$

the elements of which are provided by the result of Nabeya (1951) that, for bivariate standard normal x and y with correlation ρ, we have

$$\mathrm{E}(|xy|) = \frac{2}{\pi}(\sqrt{1-\rho^2} + \rho \arcsin \rho). \tag{13}$$

Equation (26) continues to hold for a unit-variance bivariate Student's t distribution, as detailed in the appendix of Haas (2010). Moreover, let

$$\begin{aligned} \boldsymbol{\Omega}(j) &= \mathrm{E}(\boldsymbol{Z}_t \otimes \boldsymbol{Z}_t | \Delta_t = j) = \mathrm{diag}(\mathrm{vec}(\boldsymbol{R}_j)), \\ \tilde{\boldsymbol{\Omega}}(j) &= \mathrm{E}(|\boldsymbol{Z}_t| \otimes |\boldsymbol{Z}_t| | \Delta_t = j) = \mathrm{diag}(\mathrm{vec}(\tilde{\boldsymbol{R}}_j)), \quad j \in \mathscr{E}. \end{aligned}$$

Then matrices $\boldsymbol{C}_2(j)$, $j \in \mathscr{E}$, are given by

$$\begin{aligned} \boldsymbol{C}_2(j) &= \big((\boldsymbol{A} \otimes \boldsymbol{A})\tilde{\boldsymbol{\Omega}}(j) + (\tilde{\boldsymbol{A}} \otimes \tilde{\boldsymbol{A}})\boldsymbol{\Omega}(j)\big)(\boldsymbol{e}_j' \otimes \boldsymbol{I}_M \otimes \boldsymbol{e}_j' \otimes \boldsymbol{I}_M) \\ &\quad + \kappa_1\big(\boldsymbol{e}_j' \otimes \boldsymbol{A} \otimes \boldsymbol{B} + \boldsymbol{B} \otimes \boldsymbol{e}_j' \otimes \boldsymbol{A}\big) + \boldsymbol{B} \otimes \boldsymbol{B}. \end{aligned}$$

E.g. in the practically important case of two regimes, checking for covariance stationarity involves inspecting the eigenvalues of the matrix

$$\mathbb{P}_{C_2} = \begin{pmatrix} p_{11}\boldsymbol{C}_2(1) & p_{21}\boldsymbol{C}_2(1) \\ p_{12}\boldsymbol{C}_2(2) & p_{22}\boldsymbol{C}_2(2) \end{pmatrix}.$$

4 Application to financial data

We consider volatility and correlation dynamics of global stock market and real estate equity returns, using dollar-denominated weekly (Wednesday-to-Wednesday) returns of the MSCI world and the FTSE EPRA/NAREIT global indices from January 1990 to October 2011 ($T = 1137$ observations). The latter index represents the evolution of real estate equities.[2] The analysis is based on continuously compounded percentage returns, i.e. $r_{it} = 100 \times \log(I_{it}/I_{i,t-1})$, $i = 1, 2$, where I_{1t} and I_{2t} are the MSCI and the FTSE EPRA/NAREIT index levels, respectively. Both the index levels and the returns are shown in the top and middle panels of Figure 1, reflecting the turbulent development of markets particularly since the beginning of the current millennium. Sample moments of the returns are reported in Table 1, along with the Jarque-Bera (JB) test for normality and Engle's (1982) Lagrange Multiplier (LM) test for conditional heteroskedasticity.
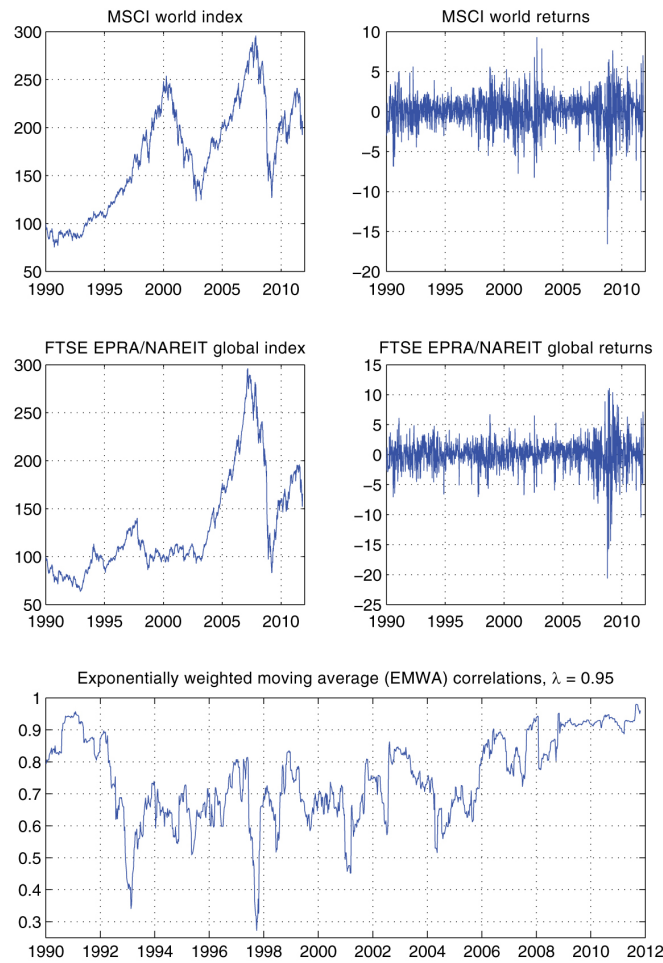
Figure 1: The top panel shows the weekly index levels (left plot) and percentage log returns (right plot) of the MSCI world stock market index from January 1990 to October 2011. The middle panel is similar, but for the FTSE EPRA/NAREIT global index reflecting the evolution of real estate equities. The bottom panel shows conditional correlations implied by an exponentially weighted moving average (EWMA) covariance matrix estimator H_t with smoothing constant $\lambda = 0.95$, i.e. $H_t = (1-\lambda)r_{t-1}r'_{t-1} + \lambda H_{t-1}$, where the initial matrix H_1 is set equal to the sample covariance matrix.

Table 1: Properties of weekly global stock market and real estate equity returns.

	Mean	**Covariance/Correlation matrix**		**Skewness**	**Kurtosis**	**JB**	**ARCH(10)**
MSCI	0.064	5.077	0.795	−0.762	7.49	1062.7***	175.6***
EPRA/NAREIT	0.044	4.639	6.714	−1.034	10.81	3091.6***	290.0***

The top right entry of the "covariance/correlation matrix" is the correlation coefficient, and the bottom left entry is the covariance. The return vector at time t is $r_t = (r_{1t}, r_{2t})'$, where r_{1t} and r_{2t} are the MSCI and FTSE EPRA/NAREIT returns, respectively, i.e. the first asset is the MSCI world stock index. JB is the Jarque-Bera test for normality, and ARCH(10) is the LM test for ARCH effects with 10 lags (cf. Engle, 1982). Asterisks *** indicate significance at the 1% level.

From Table 1, we note that the return series exhibit a considerable correlation of 0.795, which reflects the common finding that real estate equities display much more similarity to the general stock market than direct real estate investments (e.g. Morawski, Rehkugler & Füss, 2008; Heaney & Sriananthakumar, 2012). Moreover, the bottom panel of Figure 1 shows conditional correlations implied by an exponentially weighted moving average (EWMA) estimator (cf. Alexander, 2008, Ch. 3.8), which hints at time-varying conditional correlations with a particularly strong degree of comovement both at the beginning and the end of the sample, with the latter being also characterized by an outburst of unprecedented volatility.[3] Results for versions of Tse's (2000) Lagrange Multiplier (LM) test for constant conditional correlations in a multivariate GARCH model are reported in Table 2 and also provide support for time-varying conditional correlations.

Table 2: Tse's (2000) test for constant conditional correlations.

	Gaussian innovations	Student's t innovations
test statistic	7.16[***]	9.84[***]

Reported are the results of Tse's (2000) Lagrange Multiplier (LM) test for constant conditional correlations under the assumption of both Gaussian (left column) and Student's t (right column) innovations. Under the null hypothesis, returns are generated by a CCC-AGARCH process as in (25), with $k = 1$ and (27) imposed in (26). Under the null, the LM test statistic given by (38) has a limiting $\chi^2(1)$ distribution. Asterisks [***] indicate significance at the 1% level.

4.1 Fitting MS CCC-GARCH(1,1) processes

The evidence for time-varying correlations coupled with periods of low and high volatility makes the MS CCC-GARCH model defined in Section 2 a candidate for modeling these series. We fit the model with $k = 1$, 2, and 3 regimes,[4] where we confine ourselves to the *diagonal* model, with A_j, B_j, and Γ_j in (18) being diagonal matrices, $j = 1, ..., k$. In addition, we restrict the asymmetry parameters to be regime-independent, i.e.

$$\Gamma_1 = \Gamma_2 = \cdots = \Gamma_k =: \Gamma.$$

There are no clear-cut signs of conditional mean dynamics in the data, and thus we specify the model for return vector r_t as

$$r_t = \mu + \epsilon_t,$$

where μ is the constant conditional mean, and ϵ_t is generated by an MS(k) CCC-GARCH process as described in Section 2. We compare the fit of models with different k by means of the Bayesian information criterion (BIC) of Schwarz (1978), which, from results of Keribin (2000) and Francq, Roussignol, and Zakoïan (2001), can be expected to have favorable properties for this purpose. Results are reported in Table 3 for both Gaussian and Student's t innovations $\{\xi_t\}$ in (16). In both cases, models with two components are preferred, as is a conditional t distribution. Thus we focus on two-component models in the following discussion. Both normal and Student's t innovations are considered in order to highlight the role of the conditional distribution.

Table 3: Likelihood-based goodness-of-fit of MS(k) CCC-GARCH models.

	Gaussian innovations			Student's t innovations		
	$k = 1$	$k = 2$	$k = 3$	$k = 1$	$k = 2$	$k = 3$
K	11	20	31	12	21	32
$\log L$	−4371.5	−4288.0	−4261.0	−4323.7	−4260.0	−4247.6
BIC	8820.4	8716.6	8740.1	8731.7	8667.7	8720.3

Reported are likelihood-based goodness-of-fit measures for *diagonal* MS(k) CCC-GARCH models fitted to the MSCI world and FTSE EPRA/NAREIT global returns. The number of regimes is denoted by k, and $k = 1$ corresponds to the single-regime CCC of Bollerslev (1990). In all models, the asymmetry parameters are restricted to be constant across regimes, i.e. in (18), $\Gamma_1 = \cdots = \Gamma_k$. K is the number of parameters of a model, $\log L$ is the value of the maximized log-likelihood, and BIC is the Bayesian information criterion of Schwarz (1978), i.e. BIC = $−2 \times \log L + K \log T$, where T is the sample size. Smaller values of BIC are preferred.

The diagonal MS(k) CCC-GARCH model without further restrictions is rather flexible in that it allows the variances as well as the correlations to be regime-dependent. The contribution of both of these features to the overall improvement over the single-regime specification documented in Table 3 is not clear *a priori*. It is thus of interest to test various restricted models against the unrestricted specification. Specifically, we consider Pelletier's (2006) RSDC model where the switching applies to the conditional correlation matrix only, i.e. conditional volatilities are constant across regimes. The second constrained specification represents the opposite of Pelletier's (2006) model, namely the case where volatility can switch but $R_1 = R_2$ in (16). The results reported in Table 4 show that, although both restrictions are rejected against the full model by means of likelihood ratio tests, allowance for regime-specific correlations appears to be more important than switching in the univariate GARCH dynamics, and particularly so for the (generally preferred) models with Student's t innovations.

Table 4: Likelihood ratio tests (LRT) of restricted MS(2) CCC-GARCH specifications against the full (diagonal) model

	Gaussian innovations			Student's t innovations		
	Full	Pelletier (RSDC)	$R_1 = R_2$	Full	Pelletier (RSDC)	$R_1 = R_2$
K	20	14	19	21	15	20
$\log L$	−4288.0	−4311.7	−4313.0	−4260.0	−4272.6	−4310.9
LRT	–	47.5[***]	50.1[***]	–	25.2[***]	101.9[***]

The table reports likelihood ratio tests (LRT) for restricted versions of the two-component diagonal MS(2) CCC-GARCH model. The unrestricted specification, denoted as "full", is the model introduced in Section 2 with $k = 2$, and where the matrices A_j, B_j, $j = 1, 2$, and $\mathbf{\Gamma}_1 = \mathbf{\Gamma}_2$ in (18) are diagonal. "Pelletier" refers to Pelletier's (2006) regime-switching dynamic correlation (RSDC) model where only the correlation matrix is subject to regime-switching, i.e. the additional restrictions $\boldsymbol{\omega}_1 = \boldsymbol{\omega}_2$, $A_1 = A_2$, and $B_1 = B_2$ are imposed in (18). The third model restricts the correlation to be the same in both regimes, i.e. $R_1 = R_2$ in (16). $\log L$ is the value of the maximized log-likelihood, and K is the number of parameters of a model. The associated likelihood ratio test statistics, denoted as LRT, have 6 and 1 degrees of freedom, respectively. Asterisks [***] indicate significance at the 1% level.

Several characteristics of the estimated MS(2) CCC-GARCH models with Gaussian and Student's t innovations are reported in Table 5, where the single regime CCC-GARCH models are included for comparison. In Table 5, the regimes are ordered such that $\pi_{1,\infty} > \pi_{2,\infty}$. Both two-regime models have in common that Regime 1 is a low-volatility regime with moderate correlation (relative to the unconditional correlation), and Regime 2 is a high-volatility regime with rather high correlation, i.e. the diversification potential deteriorates in turbulent market periods. As reported in the bottom part of Table 5, the unconditional moments implied by the single-regime models are close to those of the two-regime specifications and are in between their regime-specific counterparts documented in the top and middle parts of the table for Regimes 1 and 2, respectively. All estimated models are covariance stationary, since $\lambda(\mathbb{P}_{C_2}) < 1$ for all estimated specifications (cf. Theorem 2).

Table 5: Characteristics of estimated MS CCC-GARCH(1,1) models.

Estimated characteristic	Gaussian innovations		Student's t innovations	
	$k = 1$	$k = 2$	$k = 1$	$k = 2$
$\rho_{12,1}$	0.766	0.636	0.769	0.655
	(0.013)	(0.025)	(0.014)	(0.024)
$E(\epsilon_{1t}^2\|\Delta_t = 1)$	4.265	3.691	3.847	3.161
$E(\epsilon_{2t}^2\|\Delta_t = 1)$	5.528	3.540	4.713	3.318
$\text{Corr}(\epsilon_{1t}, \epsilon_{2t}\|\Delta_t = 1)$	0.754	0.617	0.756	0.636
p_{11}	1	0.985	1	0.996
		(0.007)		(0.003)
$\pi_{1,\infty}$	1	0.610	1	0.563
		(0.151)		(0.109)
$(1 - p_{11})^{-1}$	∞	66.54[a]	∞	233.8[a]
$\rho_{12,2}$	–	0.929	–	0.921
		(0.009)		(0.009)
$E(\epsilon_{1t}^2\|\Delta_t = 2)$	–	6.080	–	5.846
$E(\epsilon_{2t}^2\|\Delta_t = 2)$	–	7.867	–	7.785
$\text{Corr}(\epsilon_{1t}, \epsilon_{2t}\|\Delta_t = 2)$	–	0.916	–	0.912
p_{22}	0	0.976	0	0.994
		(0.012)		(0.005)
$\pi_{2,\infty}$	0	0.390	0	0.437
		(0.151)		(0.109)
$(1 - p_{22})^{-1}$	–	42.49[a]	–	181.8[a]
$E(\epsilon_{1t}^2)$	4.265	4.622	3.847	4.336
$E(\epsilon_{2t}^2)$	5.528	5.226	4.713	5.272
$\text{Corr}(\epsilon_{1t}, \epsilon_{2t})$	0.754	0.779	0.756	0.805
$\delta = p_{11} + p_{22} - 1$	–	0.961	–	0.990
		(0.017)		(0.006)
ν	–	–	7.368	8.584
			(1.025)	(1.370)
γ_{11}	0.680	0.738	0.575	0.627
	(0.157)	(0.168)	(0.181)	(0.180)
γ_{22}	0.401	0.635	0.273	0.472
	(0.105)	(0.147)	(0.116)	(0.152)
$\lambda(\mathbb{P}_{C_2})$	0.903	0.933	0.921	0.953

Standard errors are given in parentheses. $\rho_{12,j}$ is the constant *conditional* correlation in Regime j; $\pi_{j,\infty}$ is the stationary probability and $(1-p_{jj})^{-1}$ is the expected duration of the jth regime, $j=1,2$; $\delta = p_{11} + p_{22} - 1$ is a measure for the persistence of the regime process; γ_{11} and γ_{22} are the asymmetry parameters in the volatility equation (18) for the MSCI and the FTSE/NAREIT, respectively; $\lambda(\mathbb{P}_{C_2})$ is the largest eigenvalue of matrix $\mathbb{P}_{C_2}$ defined in (24) and (25) with $l=2$, and with $\lambda(\mathbb{P}_{C_2}) < 1$ being the condition for covariance stationarity (cf. Theorem 2).

[a] Asymptotic standard errors for the expected duration of regime j, $(1-p_{jj})^{-1}$, $j=1,2$, could, at least in principle, also be calculated via the delta method. However, with $\hat{p}_{jj}$ rather close to unity, as in the case under consideration, the normal approximation would be basically useless and thus we abstain from reporting them. For example, with the estimates in the table above, the asymptotic standard error of $(1-\hat{p}_{11})^{-1}$ in the Student's t switching model would be estimated as $\sqrt{\widehat{\mathrm{Var}}(\hat{p}_{11})}/(1-\hat{p}_{11})^2 = 0.0028/(1-0.9957)^2 = 151.4$.

Comparing the regime-switching models with Gaussian and Student's t innovations, we observe that both models are characterized by fairly persistent regimes, but the persistence is more pronounced with Student's t innovations, where both "staying probabilities" p_{11} and p_{22} are rather close to unity. To illustrate the differences in estimated persistence, expected regime durations implied by estimated parameters, given by $(1-\hat{p}_{jj})^{-1}$, $j=1,2$, are also reported in Table 5. With Gaussian regimes, expected duration of the low (high)–volatility regime is slightly longer (shorter) than 1 year, whereas it is almost five (four) years with Student's t regimes.[5] This pattern, which is also discussed in Bulla (2011), Haas (2010), and Haas and Paolella (2012), is due to the tendency of a model with Gaussian regime densities to signal a regime shift whenever an untypically large (small) observation occurs within an otherwise calm (turbulent) regime.[6] Such untypical observations are easier accommodated within a given regime when the regime densities are allowed to be leptokurtic, i.e. display fatter tails and higher peaks than the normal. The same logic applies to Pelletier's model where only the correlations are subject to regime-switching, since, for fixed correlation, simultaneous extreme realizations of both variables are more likely with Student's t innovations.[7]

The upper panel of Figure 2 illustrates the models' inferred switching activity by means of the smoothed regime probabilities of the high-volatility / correlation regime under both types of innovation distribution. Both models indicate a switch to the high-correlation regime at the end of the sample beginning with the financial turmoil in the wake of the burst of the housing bubble. Implications for forecasting are depicted in the lower panel of Figure 2. The left plot of the lower panel of Figure 2 shows conditional correlations as implied by a Gaussian regime-switching model for the two situations where we know for certain that at the forecast origin we are either in the first or second regime.[8] As a function of the forecast horizon, the conditional correlation of the Gaussian model rapidly converges to its unconditional value, whereas forecasts are much more persistently affected by the current state of the world in the Student's t model, as shown in the bottom right graph of Figure 2.
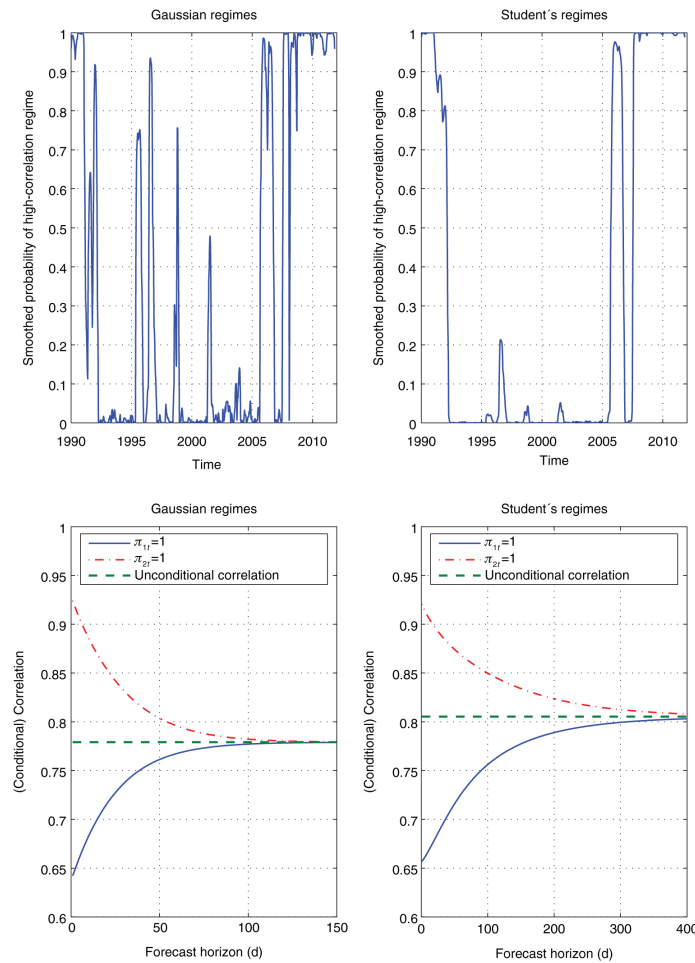
Figure 2: The upper panel shows the smoothed probabilities of Regime 2 (high-volatility/correlation) implied by the MS(2) CCC-GARCH process with Gaussian (left plot) and Student's *t* innovations (right plot). The lower panel shows conditional correlations under the assumption that we either start in the low– or high-volatility/correlation regime, as represented by the solid and dash-dotted lines, respectively. Conditional standard deviations have been initialized with appropriate unconditional expectations (cf. Footnote 5). As in the upper panel, the left and right graphs are for Gaussian and Student's *t* innovations, respectively.

4.2 Testing for within-regime correlation dynamics and comparison with other models

One of the most popular approaches to time-varying conditional correlations is the dynamic conditional correlation (DCC) model of Engle (2002). In the DCC, conditional correlations are driven by standardized shocks rather than by discrete regime shifts as in the Markov-switching processes studied herein. Both models can be combined to produce an even more flexible structure which allows the conditional correlations in each regime to be driven by DCC-type dynamics (e.g. Billio & Caporin, 2005; Otranto, 2010). However, the MS CCC-GARCH model has several advantages over its DCC-type generalization, since it is easier to estimate and admits the computation of multi-step-ahead conditional covariance matrices. In view of these advantages, it is desirable to have at one's disposal a simple test of the regime-switching CCC against the alternative of within-regime correlation dynamics. To this end, we extend Tse's (2000) Lagrange Multiplier (LM) test for constant conditional correlations to the multi-regime framework and allowing for fat-tailed (Student's *t*) innovations.[9] The details of this test, which fits into the general framework described by Hamilton (1996), are developed in Appendix C. Results are reported in Table 6 for two conditional volatility specifications under the null hypothesis, that is, both the "full" model from Table 4 as well as Pelletier's model, and both specifications are considered with Gaussian and Student's *t* innovations. Under the alternative, within-regime correlation dynamics are either regime-dependent (case (a) in Table 6) or regime-independent (case (b)). Overall, the results in Table 6 show no clear-cut sign of within-regime correlation dynamics, i.e. the switching between low– and high-correlation periods appears to capture most of the time-variation in conditional correlations.

Table 6: Lagrange Multiplier (LM) tests for constant within-regime correlations.

	Specification of alternative hypothesis	
	Case (a)	**Case (b)**
Gaussian regime densities		
Pelletier (RSDC)	4.73[*]	0.32
MS CCC	0.46	0.33
Student's t regime densities		
Pelletier (RSDC)	1.46	0.12
MS CCC	4.48	0.00

Reported are Lagrange Multiplier (LM) tests for constant conditional *within-regime* correlations in two-regime MS CCC-GARCH models, as described in Appendix C. "Pelletier" is Pelletier's (2006) regime-switching dynamic correlation (RSDC) model where only the correlation is subject to regime-switching. Note that the tests reported here are based on symmetric volatility processes, i.e. $\Gamma_1 = \Gamma_1 = \mathbf{0}$ in (18). Cases (a) and (b) are distinguished by means of the alternative hypothesis as described in Appendix C:

- Case (a) refers to the situation where, under the alternative, the correlation dynamics may be different in both regimes, i.e. in (26), both $\delta_{12,j}$, $j = 1, 2$, may be nonzero and different.
- In case (b), it is assumed under the alternative that correlation dynamics are regime-independent, i.e. $\delta_{12,1} = \delta_{12,2}$ in (26).

Under the null hypothesis of constant conditional correlations in both regimes, the test statistic for case (a) is asymptotically distributed as $\chi^2(2)$, whereas in case (b) the limiting distribution is $\chi^2(1)$. Asterisk [*] denotes significance at the 10% level (the p–value is 0.094).

In view of these results, we compare conditional correlations implied by the MS CCC and the DCC model of Engle (2002).[10] These correlations are shown in Figure 3 for both Gaussian (top panel) and Student's t innovations (bottom panel). Comparing the upper with the lower panel of Figure 3, MS CCC-implied correlations are smoother with Student's t than with Gaussian innovations. The DCC-implied correlations depend much less on the innovation distribution and, as already observed by Pelletier (2006), are less smooth than their regime-switching CCC counterparts.[11] Roughly, however, both types of models contain similar information about low– and high-correlation periods in the data. In particular, they agree with regard to the jump in correlation at the onset of the recent financial crisis.
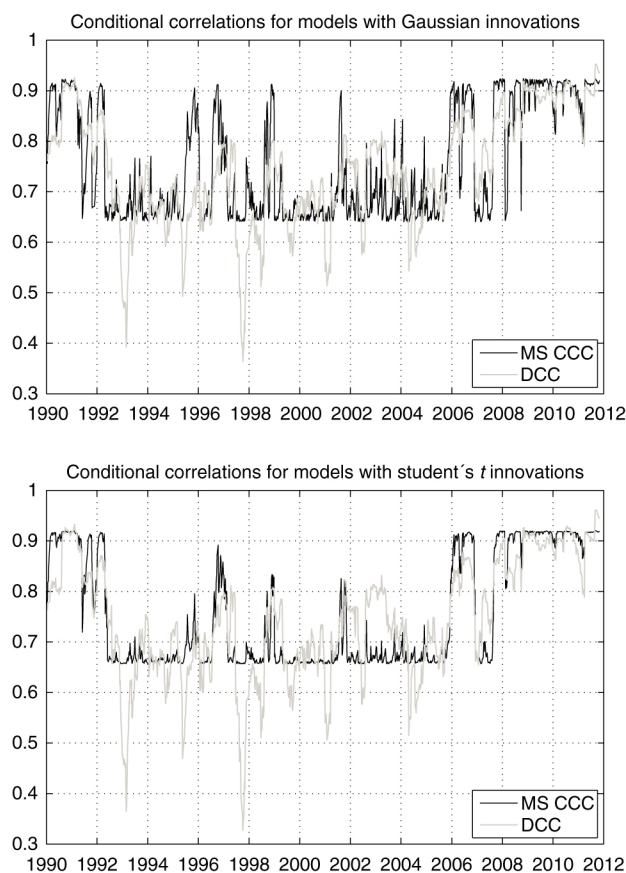

Conditional correlations for models with Gaussian innovations

Conditional correlations for models with student´s t innovations

Figure 3: The upper panel shows conditional correlations as implied by the MS CCC-GARCH model and the DCC process with Gaussian innovations. The lower panel repeats this, but for models with Student's *t* innovations. The one-step-ahead conditional correlations of the MS CCC-GARCH models are extracted from the conditional covariance matrix (23) for *d* = 1.

Multi-step ahead conditional correlations for both types of models are illustrated in Figure 4, which resembles the lower part of Figure 3 but additionally includes conditional correlations implied by the Student's *t* DCC model, as calculated by simulation.[12] Initial values for the conditional correlation matrix and the conditional standard deviations in the DCC were selected such that they match those of the Student's *t* MS CCC in the respective regimes. The long-run correlation of the DCC model is a bit lower than those implied by estimated MS CCC processes. However, with regard to the persistence of multi-step correlations, the DCC is more like the Student's *t* rather than the Gaussian MS CCC-GARCH. This is in line with DCC parameter estimates, as reported in Table 7. Interpreting $\hat{a} + \hat{b}$ as an estimate of the persistence of conditional correlations, i.e. the equivalent of δ in Table 5, then the persistence in conditional correlations implied by estimated DCC models is close to the value of $\hat{\delta} = 0.99$ of the Student's *t* MS CCC model.
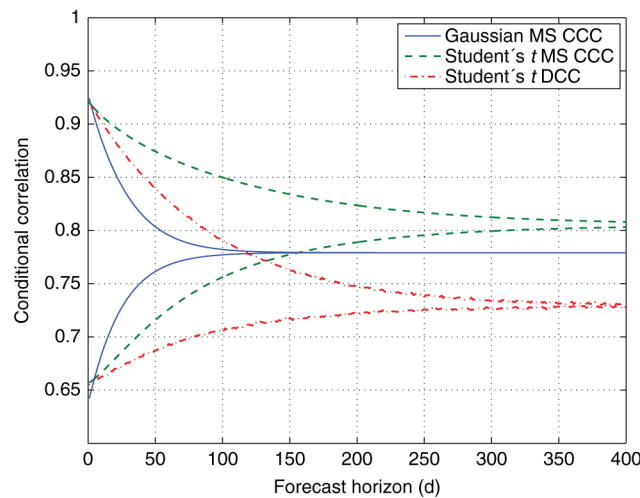

Figure 4: Shown are conditional correlations similar to the lower part of and as explained in the legend of Figure 3. The curves for the MS CCC-GARCH models reproduce those in the bottom part of Figure 3. Initial values for the conditional correlation matrix and the conditional standard deviations in the DCC model were determined such that they match those of the Student's *t* MS CCC in the respective (low– and high correlation/variance) regimes.

Table 7: Parameter estimates for correlation dynamics in DCC models.

Innovations	$\hat{a}$	$\hat{b}$	$\hat{a} + \hat{b}$
Gaussian	0.044	0.942	0.986
	(0.014)	(0.022)	(0.009)
Student's *t*	0.047	0.941	0.988
	(0.015)	(0.021)	(0.008)

Shown are estimates of the parameters driving the correlation dynamics in Gaussian and Student's *t* DCC models (Engle, 2002), with standard errors given in parentheses. The evolution of the conditional correlation matrix R_t is described by

$$Q_t = (1 - a - b)S + a z_{t-1} z'_{t-1} + b Q_{t-1}$$

$$R_t = (I \odot Q_t)^{-1/2} Q_t (I \odot Q_t)^{-1/2},$$

where the z_t are the standardized ("degarched") residuals, and S is estimated via their sample correlation matrix.

4.3 Application to portfolio selection

We finally compare the models' performance in an out-of-sample portfolio application. To do so, we first reestimate all models using roughly the first 10 years of data (the first 500 observations) and then update the estimates every 4 weeks, using an expanding window of observations. Estimated models are used to construct

ex-ante global minimum variance portfolios (GMVP) over holding periods up to 24 weeks (ca. 6 months).[13] Using non-overlapping holding periods, we thus have, e.g. 637 and 318 out-of-sample realized GMVP returns for the 1– and 2–week holding periods, respectively. Closed-form conditional covariances as derived in Section B.3 are used for all CCC-type models, whereas the DCC-implied conditional covariance matrices are estimated from 10,000 simulated sample paths.

For a broader perspective, we also include three multivariate GARCH models outside of the CCC or DCC families. Namely, we consider the BEKK model of Engle and Kroner (1995) which Gaussian and Student's *t* innovations, as well as the multivariate asymmetric mixed normal BEKK-GARCH (MixN BEKK) process of Haas, Mittnik, and Paolella (2009). As detailed in Appendix D, the latter specification combines a conditional mixture distribution with constant mixing weights with conditional regime-specific correlations which are time-varying according to a multivariate BEKK process. Thus the MixN BEKK model can be viewed as the photographic negative of the MS CCC in that the latter's characteristics of time-varying mixing weights and constant regime-specific correlations have been reversed.

Results are reported in Table 8. For the most basic model, i.e. the single-regime Gaussian CCC, we report the standard deviation of the realized returns, whereas for all other models their respective standard deviation divided by that of the Gaussian CCC is shown. The results in Table 8 show that using a Student's *t* rather than a Gaussian distribution improves the results somewhat for all models and forecast horizons. However, the improvements tend to be minor except for the MS CCC-GARCH, for which they are quite substantial, and even more so for longer forecast horizons. At first, it may appear surprising that the MS CCC with Student's *t* innovations displays the best results for all forecast horizons, whereas the performance of its Gaussian cousin is rather disappointing. However, this becomes plausible in view of the discussion in Section 4.1. Namely, in the Gaussian model, with volatilities being allowed to switch, the high-volatility regime will tend to latch onto a few "outliers", which hampers the ability of the model to identify the smooth and long-lived low– and high-correlation regimes. Pelletier's RSDC model, with switching correlations only, suffers less severely from this problem, and thus its performance is more even across distributional assumptions. Still, however, the results for the MS CCC with *t* errors in Table 8 suggest that there may be additional benefits from allowing both correlations and volatilities being regime-dependent, provided the conditional distribution is flexible enough to cope with isolated untypical observations within a given regime. Both MS CCC as well as RSDC consistently outperform DCC at longer forecast horizons, which may indicate that the regime-switching models are better suited than the DCC to capture relatively long-lived persistent correlation regimes. On the other hand, however, both the single-regime as well as the MixN BEKK models also improve upon DCC forecasts at longer horizons, suggesting that regime-switching and suitably specified GARCH-type conditional correlations are to some extent substitutes to each other.

Table 8: Realized standard deviations of out-of-sample global minimum variance portfolio (GMVP) returns.

Horizon (*D*)	1	2	3	4	8	12	16	20	24
# returns	637	318	212	159	79	53	39	31	26
Conditional correlation models with Gaussian innovations									
CCC	2.425	3.324	4.472	4.926	7.192	9.823	10.791	15.506	17.439
DCC	0.954	0.986	0.960	0.979	0.970	0.961	0.992	0.986	0.978
RSDC	0.971	0.973	0.965	0.962	0.952	0.941	0.966	0.942	0.926
MS CCC	0.982	0.989	0.986	0.995	1.010	0.984	1.046	1.040	1.001
Conditional correlation models with Student's *t* innovations									
CCC	0.996	0.995	0.993	0.991	0.991	0.979	0.992	0.990	0.981
DCC	0.946	0.981	0.945	0.968	0.956	0.928	0.978	0.963	0.947
RSDC	0.961	0.968	0.951	0.953	0.935	0.901	0.952	0.924	0.895
MS CCC	0.936	0.965	0.924	0.938	0.902	0.862	0.894	0.843	0.821
BEKK-type models									
Gaussian BEKK	0.958	1.008	0.950	0.974	0.939	0.889	0.945	0.903	0.863
Student's *t* BEKK	0.953	0.977	0.939	0.931	0.919	0.879	0.930	0.898	0.882
MixN BEKK	0.960	0.964	0.948	0.941	0.943	0.907	0.938	0.889	0.867

Reported are the results of constructing *ex-ante* global minimum variance portfolios (GMVP) implied by different GARCH models and for different forecast horizons, D (weeks). Calculations refer to multi-period cumulative returns, i.e. if r_{t+d} is the single-period return vector at time (week) $t + d$, then the D–period ahead cumulative return vector at forecast origin t is $\sum_{d=1}^{D} r_{t+d}$, and the multi-period ahead covariance matrices are calculated accordingly (assuming returns are not autocorrelated). The row labeled "# returns" reports the number of (non-overlapping) holding periods used to produce the results for each respective forecast horizon.
CCC and DCC are Bollerslev's (1990) constant and Engle's (2002) dynamic conditional correlation models, respectively. RSDC is Pelletier's (2006) model, and MS CCC is the Markov-switching GARCH process defined in Section 2. In all CCC-type and DCC models, volatilities are driven by absolute value asymmetric GARCH processes, cf. Equation (18). BEKK and MixN BEKK are the standard and (two-component) mixed normal BEKK processes, respectively, with asymmetric volatility response as described in Appendix D.
For the CCC with normal innovations, the table reports the standard deviation of the *ex-post* (realized) portfolio returns of the *ex-ante* GMVP for each forecast horizon, D. For all other models, their respective standard deviation divided by that of the Gaussian CCC is shown.

5 Concluding remarks

We conclude with a remark referring to the frequently contemplated "curse of dimensionality" problem. An advantage of the (diagonal) CCC, DCC, and Pelletier's (2006) RSDC models is that, via two-step estimation, application to high-dimensional time series is feasible.[14] This property is not shared by the model studied herein with both regime-specific correlations and variance dynamics. We do not deem this to be a disadvantage, since there are many applications where the advantage of a more flexible dynamic structure may very well outweigh the benefits of parsimony as long as the dimension of the problem is sufficiently low. Studies of the dynamics of broadly defined asset classes, as illustrated in Section 4, are a typical example, where a richer specification can lead to a better understanding and potentially improved forecasts of the joint process under study. As another recent example from the literature somewhat related to the application in Section 4, Case, Guidolin, and Yildrim (2014), using monthly data from 1972 to 2009, find that even a four-regime MS model is required to appropriately describe the evolution of the joint conditional distribution of REIT, stock, and bond returns, since in particular the bond market regimes fail to be synchronized with those of the other two markets. For higher-dimensional systems, Pelletier's (2006) model, which is nested in the general specification of this paper, appears to provide a reasonable balance between flexibility on the one hand and parsimony and tractability on the other. This is suggested in particular since the results in Section 4 (cf. Table 4) revealed that allowing for regime-switching correlations is of greater value than doing the same for the dynamics of individual volatilities.

Acknowledgements

Appendix

A Proofs of the theorems

Proof of Theorem 1. The 'if' part follows from Brandt (1986) or Bougerol and Picard (1992).

Conversely, assume that there exists a strictly stationary solution (ϵ_t) of the MS(k)-CCC-GARCH process defined by (14)–(19). Iterating (22), we have, for any $m > 0$, that

$$X_0 = \omega + \sum_{n=1}^{m} C_{\Delta_{-1},-1} C_{\Delta_{-2},-2} \dots C_{\Delta_{-n},-n} \omega + C_{\Delta_{-1},-1} C_{\Delta_{-2},-2} \dots C_{\Delta_{-m-1},-m-1} X_{-m-1}.$$

From all entries of X_t, $C_{\Delta_t,t}$, and ω being nonnegative, we know, for any $m > 0$,

$$\sum_{n=1}^{m} C_{\Delta_{-1},-1} C_{\Delta_{-2},-2} \dots C_{\Delta_{-n},-n} \boldsymbol{\omega} \le \boldsymbol{X}_0, \ a.s.$$

Therefore, $\sum_{n=1}^{m} C_{\Delta_{-1},-1} C_{\Delta_{-2},-2} \dots C_{\Delta_{-n},-n} \boldsymbol{\omega}$ converges a.s. Thus, we have that

$$\lim_{n\to\infty} C_{\Delta_{-1},-1} C_{\Delta_{-2},-2} \dots C_{\Delta_{-n},-n} \boldsymbol{\omega} = 0, \ a.s.$$

By $\boldsymbol{\omega} > 0$, this implies

$$\lim_{n\to\infty} \left\| C_{\Delta_{-1},-1} C_{\Delta_{-2},-2} \dots C_{\Delta_{-n},-n} \right\| = 0, \ a.s.$$

Hence, by Lemma 3.4 in Bougerol and Picard (1992), we know that the top Lyapunov exponent associated with the matrices $(C_{\Delta_t,t})$ is strictly negative. This completes the proof of the theorem.

Proof of Theorem 2. Write

$$\boldsymbol{X}_{t,m} = C_{\Delta_{t-1},t-1} C_{\Delta_{t-2},t-2} \dots C_{\Delta_{t-m},t-m} \boldsymbol{\omega}, \quad m \ge 1,$$

and $\boldsymbol{X}_{t,0} = \boldsymbol{\omega}$. For any vector X such that AX is well defined, we have $(AX)^{\otimes l} = A^{l} X^{l}$. It follows that

$$\boldsymbol{X}_{t,m}^{\otimes l} = C_{\Delta_{t-1},t-1}^{\otimes l} C_{\Delta_{t-2},t-2}^{\otimes l} \dots C_{\Delta_{t-m},t-m}^{\otimes l} \boldsymbol{\omega}^{\otimes l}.$$

By Lemma 1 in Francq and Zakoïan (2005), we have

$$\begin{aligned}
E(\boldsymbol{X}_{t,m}^{\otimes l}) &= E\{E(C_{\Delta_{t-1},t-1}^{\otimes l} C_{\Delta_{t-2},t-2}^{\otimes l} \dots C_{\Delta_{t-m},t-m}^{\otimes l} \boldsymbol{\omega}^{\otimes l} | \Delta_{t-1}, \dots, \Delta_{t-m})\} \\
&= E(C_l(\Delta_{t-1}) C_l(\Delta_{t-2}) \dots C_l(\Delta_{t-m})) \boldsymbol{\omega}^{\otimes l} \\
&= \mathbb{I}(\mathbb{P}_{C_l})^m \boldsymbol{\pi}_{\boldsymbol{\omega}^{\otimes l}},
\end{aligned}$$

where $\mathbb{I} = (\boldsymbol{I}_{(kM)^l}, \dots, \boldsymbol{I}_{(kM)^l})$ is a $(kM)^l \times k(kM)^l$ matrix, and $\boldsymbol{\pi}_{\boldsymbol{\omega}^{\otimes l}} = (\boldsymbol{\pi}_\infty \otimes \mathbf{1}_{k^{l-1}M^l}) \odot \boldsymbol{\omega}^{\otimes l}$. Thus, by $\|A\|\|B\| = \|A \otimes B\| = \|B \otimes A\|$, we have

$$(E\|\boldsymbol{X}_{t,m}\|^l)^{1/l} = (\|E(\boldsymbol{X}_{t,m}^{\otimes l})\|)^{1/l} = \|\mathbb{I}(\mathbb{P}_{C_l})^m \boldsymbol{\pi}_{\boldsymbol{\omega}^{\otimes l}}\|^{1/l} \le \|\mathbb{I}\|^{1/l} \|(\mathbb{P}_{C_l})^m\|^{1/l} \|\boldsymbol{\pi}_{\boldsymbol{\omega}^{\otimes l}}\|^{1/l} \to 0$$

at an exponential rate as $m \to \infty$, because $\lambda(\mathbb{P}_{C_l}) < 1$. This shows that

$$\lim_{n\to\infty} \sum_{m=0}^{n} \boldsymbol{X}_{t,m} = \boldsymbol{X}_t^* = \boldsymbol{\omega} + \sum_{n=1}^{\infty} C_{\Delta_{t-1},t-1} C_{\Delta_{t-2},t-2} \dots C_{\Delta_{t-n},t-n} \boldsymbol{\omega},$$

both in L^l and almost surely. It is obvious that $\boldsymbol{X}_t^*$ satisfies (22) and is strictly stationary and ergodic. This completes the proof of the theorem.

B Calculation of the moments

We use the notation introduced in Section 3. For later reference, we also define, for $j \in \mathcal{E}$,

$$Y(j) = E\left(\boldsymbol{I}_k \otimes \boldsymbol{Z}_t \otimes \boldsymbol{I}_k \otimes \boldsymbol{Z}_t | \Delta_t = j\right) = \text{diag}[\text{vec}((\mathbf{1}_k \mathbf{1}_k') \otimes \boldsymbol{R}_j)], \tag{14}$$

$$\tilde{Y}(j) = E\left(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_t| \otimes \boldsymbol{I}_k \otimes |Z_t| | \Delta_t = j\right) = \text{diag}[\text{vec}((\mathbf{1}_k \mathbf{1}_k') \otimes \tilde{\boldsymbol{R}}_j)], \tag{15}$$

where $\mathbf{1}_k$ is a k–dimensional column of ones.

B.1 The unconditional covariance matrix

For calculating the moments of the MS CCC-GARCH process, we use the following basic result.

Lemma 1

(Francq and Zakoïan (2005), *Lemma 3*) *For $\ell \geq 1$, if the variable $\boldsymbol{Y}_{t-\ell}$ belongs to the information set generated by $\{\epsilon_s : s \leq t - \ell\}$, then*

$$\pi_{j,\infty}\mathrm{E}(\boldsymbol{Y}_{t-\ell}|\Delta_t = j) = \sum_{i=1}^k \pi_{i,\infty} p_{ij}^{(\ell)} \mathrm{E}(\boldsymbol{Y}_{t-\ell}|\Delta_{t-\ell} = i),$$

where the $p_{ij}^{(\ell)} := p(\Delta_t = j|\Delta_{t-\ell} = i)$, $i, j \in \mathscr{E}$, denote the ℓ–step transition probabilities, as given by the elements of $\boldsymbol{P}^\ell$.

Using Lemma 1, we have

$$
\begin{aligned}
\pi_{j,\infty}\mathrm{E}(\boldsymbol{X}_t|\Delta_{t-1} = j) &= \pi_{j,\infty}\boldsymbol{\omega} + \pi_{j,\infty}\boldsymbol{C}_1(j)\mathrm{E}(\boldsymbol{X}_{t-1}|\Delta_{t-1} = j) \\
&= \pi_{j,\infty}\boldsymbol{\omega} + \sum_{i=1}^k p_{ij}\boldsymbol{C}_1(j)\pi_{i,\infty}\mathrm{E}(\boldsymbol{X}_{t-1}|\Delta_{t-2} = i), \quad j = 1, \dots, k.
\end{aligned}
\tag{16}
$$

Equation (16) implies

$$\boldsymbol{V}_1 = \boldsymbol{\pi}_\infty \otimes \boldsymbol{\omega} + \mathbb{P}_{C_1}\boldsymbol{V}_1,$$

where

$$
\boldsymbol{V}_1 = \begin{pmatrix} \pi_{1,\infty}\mathrm{E}(\boldsymbol{X}_t|\Delta_{t-1} = 1) \\ \pi_{2,\infty}\mathrm{E}(\boldsymbol{X}_t|\Delta_{t-1} = 2) \\ \vdots \\ \pi_{k,\infty}\mathrm{E}(\boldsymbol{X}_t|\Delta_{t-1} = k) \end{pmatrix}.
\tag{17}
$$

Thus the first absolute moments are

$$
\begin{aligned}
\mathrm{E}(|\boldsymbol{\epsilon}_t|) &= \sum_{j=1}^k \pi_{j,\infty}\mathrm{E}(|\boldsymbol{\epsilon}_t||\Delta_t = j) \\
&= \kappa_1 \sum_{j=1}^k \sum_{i=1}^k \pi_{j,\infty} p(\Delta_{t-1} = i|\Delta_t = j)\mathrm{E}(\boldsymbol{\sigma}_{jt}|\Delta_{t-1} = i) \\
&= \kappa_1 \sum_{j=1}^k \sum_{i=1}^k p_{ij}\pi_{i,\infty}\mathrm{E}(\boldsymbol{\sigma}_{jt}|\Delta_{t-1} = i) \\
&= \kappa_1 (\mathrm{vec}(\boldsymbol{P})' \otimes \boldsymbol{I}_M)\boldsymbol{V}_1.
\end{aligned}
$$

For the covariance matrix, proceeding similarly,

$$
\begin{aligned}
\pi_{j,\infty}\mathrm{E}[\mathrm{vec}(\boldsymbol{X}_t\boldsymbol{X}_t')|\Delta_{t-1} = j] &= \pi_{j,\infty}(\boldsymbol{\omega} \otimes \boldsymbol{\omega}) + \sum_{i=1}^k p_{ij}\boldsymbol{C}_{21}(j)\mathrm{E}(\boldsymbol{X}_{t-1}|\Delta_{t-2} = i) \\
&\quad + \sum_{i=1}^k p_{ij}\boldsymbol{C}_2(j)\mathrm{E}[\mathrm{vec}(\boldsymbol{X}_{t-1}\boldsymbol{X}_{t-1}')|\Delta_{t-2} = i],
\end{aligned}
\tag{18}
$$

where $\boldsymbol{C}_{21}(j) = \boldsymbol{\omega} \otimes \boldsymbol{C}_1(j) + \boldsymbol{C}_1(j) \otimes \boldsymbol{\omega}$, $j = 1, \dots, k$. Equation (18) implies

$$\boldsymbol{V}_2 = \boldsymbol{\pi}_\infty \otimes \boldsymbol{\omega} \otimes \boldsymbol{\omega} + \mathbb{P}_{C_{21}}\boldsymbol{V}_1 + \mathbb{P}_{C_2}\boldsymbol{V}_2,$$

where $\boldsymbol{V}_2$ is as $\boldsymbol{V}_1$ in (17) but with $\pi_{j,\infty}\,\mathrm{E}(\boldsymbol{X}_t|\boldsymbol{\Delta}_{t-1} = j)$ replaced by $\pi_{j,\infty}\mathrm{E}[\mathrm{vec}(\boldsymbol{X}_t\boldsymbol{X}_t')|\Delta_{t-1} = j]$, $j = 1, \dots, k$. Thus the unconditional covariance matrix of $\{\boldsymbol{\epsilon}_t\}$ is

$$
\begin{aligned}
\mathrm{E}[\mathrm{vec}(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t')] &= \sum_{j=1}^k \pi_{j,\infty}\mathrm{E}[\mathrm{vec}(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t')|\Delta_t = j] \\
&= \sum_{j=1}^k \pi_{j,\infty}\mathrm{E}\left\{\mathrm{vec}[(\boldsymbol{e}_j \otimes \boldsymbol{I}_M)'(\boldsymbol{I}_k \otimes \boldsymbol{Z}_t)\boldsymbol{X}_t\boldsymbol{X}_t'(\boldsymbol{I}_k \otimes \boldsymbol{Z}_t)(\boldsymbol{e}_j \otimes \boldsymbol{I}_M)|\Delta_t = j]\right\} \\
&= \sum_{j=1}^k (\boldsymbol{e}_j \otimes \boldsymbol{I}_M \otimes \boldsymbol{e}_j \otimes \boldsymbol{I}_M)' \sum_{i=1}^k p_{ij}\Upsilon(j)\pi_{i,\infty}\mathrm{E}[\mathrm{vec}(\boldsymbol{X}_t\boldsymbol{X}_t')|\Delta_{t-1} = i] \\
&= \sum_{j=1}^k \boldsymbol{e}_j' \otimes (\boldsymbol{e}_j \otimes \boldsymbol{I}_M \otimes \boldsymbol{e}_j \otimes \boldsymbol{I}_M)'\mathbb{P}_\Upsilon \boldsymbol{V}_2,
\end{aligned}
\tag{19}
$$

where definitions (25) and (14) were used. The *regime-specific* unconditional covariance matrices are also of interest and given by

$$\mathrm{E}[\mathrm{vec}(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t')|\Delta_t = j] = \pi_{j,\infty}^{-1}\boldsymbol{e}_j' \otimes (\boldsymbol{e}_j \otimes \boldsymbol{I}_M \otimes \boldsymbol{e}_j \otimes \boldsymbol{I}_M)'\mathbb{P}_\Upsilon\boldsymbol{V}_2, \quad j = 1, \dots, k.$$

B.2 Autocorrelations of the absolute process

To calculate the autocorrelation function of the absolute process, $\mathrm{E}[\mathrm{vec}(|\boldsymbol{\epsilon}_t||\boldsymbol{\epsilon}_t|')]$ is required, which directly follows from (15) and (19) as

$$\mathrm{E}[\mathrm{vec}(|\boldsymbol{\epsilon}_t||\boldsymbol{\epsilon}_t|')] = \sum_{j=1}^{k} \boldsymbol{e}_j' \otimes (\boldsymbol{e}_j \otimes \boldsymbol{I}_M \otimes \boldsymbol{e}_j \otimes \boldsymbol{I}_M)' \mathbb{P}_{\bar{Y}} \boldsymbol{V}_2.$$

The cross moment matrices are obtained via

$$\begin{aligned}
\mathrm{E}(|\boldsymbol{\epsilon}_t||\boldsymbol{\epsilon}_{t-\tau}|') &= \mathrm{E}\left\{(\boldsymbol{e}_{\Delta_t} \otimes \boldsymbol{I}_M)'(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_t|)\boldsymbol{X}_t\boldsymbol{X}_{t-\tau}'(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-\tau}|)(\boldsymbol{e}_{\Delta_{t-\tau}} \otimes \boldsymbol{I}_M)\right\} \\
&= \kappa_1 \sum_{i=1}^{k}\sum_{j=1}^{k}(\boldsymbol{e}_j \otimes \boldsymbol{I}_M)' p(\Delta_{t-\tau}=i \cap \Delta_t=j) \\
&\quad \times \mathrm{E}\left\{\boldsymbol{X}_t\boldsymbol{X}_{t-\tau}'(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-\tau}|)|\Delta_{t-\tau}=i \cap \Delta_t=j\right\}(\boldsymbol{e}_i \otimes \boldsymbol{I}_M) \\
&= \kappa_1 \sum_{i=1}^{k}\sum_{j=1}^{k}(\boldsymbol{e}_j \otimes \boldsymbol{I}_M)' \boldsymbol{S}_{ij}(\tau)(\boldsymbol{e}_i \otimes \boldsymbol{I}_M),
\end{aligned}$$

where, for $i,j=1,\dots,k$, and with $p(\Delta_{t-\tau}=i \cap \Delta_t=j)=\pi_{i,\infty}p_{ij}^{(\tau)}$,

$$\begin{aligned}
\boldsymbol{S}_{ij}(\tau) &= \pi_{i,\infty}p_{ij}^{(\tau)}\mathrm{E}\left\{\boldsymbol{X}_t\boldsymbol{X}_{t-\tau}'(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-\tau}|)|\Delta_{t-\tau}=i \cap \Delta_t=j\right\} \\
&= \pi_{i,\infty}p_{ij}^{(\tau)}\mathrm{E}\left\{(\boldsymbol{\omega}+\boldsymbol{C}_{\Delta_{t-1},t-1}\boldsymbol{X}_{t-1})\boldsymbol{X}_{t-\tau}'(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-\tau}|)|\Delta_{t-\tau}=i \cap \Delta_t=j\right\} \\
&= \pi_{i,\infty}p_{ij}^{(\tau)}\boldsymbol{\omega}\kappa_1\mathrm{E}(\boldsymbol{X}_{t-\tau}'|\Delta_{t-\tau}=i) \\
&\quad +\pi_{i,\infty}p_{ij}^{(\tau)}\mathrm{E}\left\{\boldsymbol{C}_{\Delta_{t-1},t-1}\boldsymbol{X}_{t-1}\boldsymbol{X}_{t-\tau}'(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-\tau}|)|\Delta_{t-\tau}=i \cap \Delta_t=j\right\} \\
&= \pi_{i,\infty}p_{ij}^{(\tau)}\boldsymbol{\omega}\kappa_1\mathrm{E}(\boldsymbol{X}_{t-\tau}'|\Delta_{t-\tau}=i) \\
&\quad +\sum_{\ell=1}^{k}\pi_{i,\infty}p_{i\ell}^{(\tau-1)}p_{\ell j} \\
&\quad\quad \times\mathrm{E}\left\{\boldsymbol{C}_{\Delta_{t-1},t-1}\boldsymbol{X}_{t-1}\boldsymbol{X}_{t-\tau}'(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-\tau}|)|\Delta_{t-\tau}=i \cap \Delta_{t-1}=\ell \cap \Delta_t=j\right\} \\
&= \pi_{i,\infty}p_{ij}^{(\tau)}\boldsymbol{\omega}\kappa_1\mathrm{E}(\boldsymbol{X}_{t-\tau}'|\Delta_{t-\tau}=i) \\
&\quad +\sum_{\ell=1}^{k}\pi_{i,\infty}p_{i\ell}^{(\tau-1)}p_{\ell j}\boldsymbol{C}_1(\ell)\mathrm{E}\left\{\boldsymbol{X}_{t-1}\boldsymbol{X}_{t-\tau}'(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-\tau}|)|\Delta_{t-\tau}=i \cap \Delta_{t-1}=\ell\right\} \\
&= \pi_{i,\infty}p_{ij}^{(\tau)}\boldsymbol{\omega}\kappa_1\mathrm{E}(\boldsymbol{X}_{t-\tau}'|\Delta_{t-\tau}=i)+\sum_{\ell=1}^{k}p_{\ell j}\boldsymbol{C}_1(\ell)\boldsymbol{S}_{i\ell}(\tau-1), \quad \tau \geq 2,
\end{aligned}$$

that is,

$$\boldsymbol{S}(\tau) = \kappa_1(\boldsymbol{P}^\tau \otimes \boldsymbol{\omega})\tilde{\boldsymbol{V}}_1 + \tilde{\mathbb{P}}_{C_1}\boldsymbol{S}(\tau-1), \quad \tau \geq 2.$$

where

$$\boldsymbol{S}(\tau) = \begin{pmatrix} \boldsymbol{S}_{11}(\tau) & \cdots & \boldsymbol{S}_{k1}(\tau) \\ \vdots & \ddots & \vdots \\ \boldsymbol{S}_{1k}(\tau) & \cdots & \boldsymbol{S}_{kk}(\tau) \end{pmatrix}, \quad \tilde{\boldsymbol{V}}_1 = \begin{pmatrix} \pi_{1,\infty}\mathrm{E}(\boldsymbol{X}_t'|\Delta_t=1) & \cdots & \boldsymbol{0}_{1 \times kM} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0}_{1 \times kM} & \cdots & \pi_{k,\infty}\mathrm{E}(\boldsymbol{X}_t'|\Delta_t=k) \end{pmatrix},$$

the diagonal blocks of $\tilde{\boldsymbol{V}}_1$ can be extracted from the vector $(\boldsymbol{P} \otimes \boldsymbol{I}_{kM})\boldsymbol{V}_1 = (\pi_{1,\infty}\mathrm{E}(\boldsymbol{X}_t|\Delta_t=1)',\dots,\pi_{k,\infty}\mathrm{E}(\boldsymbol{X}_t'|\Delta_t=k)')'$, and, similar to (25),

$$\tilde{\mathbb{P}}_{C_1} = \begin{pmatrix} p_{11}\boldsymbol{C}_1(1) & \cdots & p_{k1}\boldsymbol{C}_1(k) \\ \vdots & \ddots & \vdots \\ p_{1k}\boldsymbol{C}_1(1) & \cdots & p_{kk}\boldsymbol{C}_1(k) \end{pmatrix}. \tag{20}$$

For $\tau=1$, we compute

$$\begin{aligned}
\boldsymbol{S}_{ij}(1) &= \pi_{i,\infty}p_{ij}\mathrm{E}[\boldsymbol{X}_t\boldsymbol{X}_{t-1}'(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-1}|)|\Delta_{t-1}=i \cap \Delta_t=j] \\
&= \pi_{i,\infty}p_{ij}\mathrm{E}[(\boldsymbol{\omega}+\boldsymbol{C}_{\Delta_{t-1},t-1}\boldsymbol{X}_{t-1})\boldsymbol{X}_{t-1}'(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-1}|)|\Delta_{t-1}=i] \\
&= \kappa_1 p_{ij}\boldsymbol{\omega}\pi_{i,\infty}\mathrm{E}(\boldsymbol{X}_t'|\Delta_t=i)+p_{ij}\pi_{i,\infty}\mathrm{E}[\boldsymbol{C}_{\Delta_{t-1},t-1}\boldsymbol{X}_{t-1}\boldsymbol{X}_{t-1}'(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-1}|)|\Delta_{t-1}=i].
\end{aligned}$$

Hence

$$S(1) = \kappa_1 (\boldsymbol{P} \otimes \boldsymbol{\omega}) \tilde{\boldsymbol{V}}_1 + \tilde{\mathbb{P}}_{\check{C}},$$

where $\tilde{\mathbb{P}}_{\check{C}}$ is as in (20) with

$$\check{\boldsymbol{C}}(i) = \pi_{i,\infty} \mathrm{E}[\boldsymbol{C}_{\Delta_{t-1},t-1} \boldsymbol{X}_{t-1} \boldsymbol{X}'_{t-1} (\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-1}|) | \Delta_{t-1} = i], \quad i = 1, \ldots, k. \tag{21}$$

The expectation in (21) is

$$
\begin{aligned}
&\mathrm{E}(\mathrm{vec}(\boldsymbol{C}_{\Delta_{t-1},t-1} \boldsymbol{X}_{t-1} \boldsymbol{X}'_{t-1} (\boldsymbol{I}_M \otimes |\boldsymbol{Z}_{t-1}|)) | \Delta_{t-1} = i) \\
&\quad = \pi_{i,\infty} \mathrm{E}((\boldsymbol{I}_M \otimes |\boldsymbol{Z}_{t-1}| \otimes \boldsymbol{C}_{\Delta_{t-1},t-1}) \mathrm{vec}(\boldsymbol{X}_{t-1} \boldsymbol{X}'_{t-1}) | \Delta_{t-1} = i) \\
&\quad = \mathrm{E}(\boldsymbol{I}_M \otimes |\boldsymbol{Z}_{t-1}| \otimes \boldsymbol{C}_{\Delta_{t-1},t-1}) | \Delta_{t-1} = i) \pi_{i,\infty} \mathrm{E}(\mathrm{vec}(\boldsymbol{X}_t \boldsymbol{X}'_t) | \Delta_t = i),
\end{aligned}
$$

where $\pi_{i,\infty} \mathrm{E}(\mathrm{vec}(\boldsymbol{X}_t \boldsymbol{X}'_t) | \Delta_t = i)$, $i = 1, \ldots, k$, can be extracted from

$$(\boldsymbol{P} \otimes \boldsymbol{I}_{k^2 M^2}) \boldsymbol{V}_2 = (\pi_{1,\infty} \mathrm{E}(\mathrm{vec}(\boldsymbol{X}_t \boldsymbol{X}'_t) | \Delta_t = 1)', \ldots, \pi_{k,\infty} \mathrm{E}(\mathrm{vec}(\boldsymbol{X}_t \boldsymbol{X}'_t) | \Delta_t = k)')',$$

and

$$
\begin{aligned}
&\mathrm{E}(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-1}| \otimes \boldsymbol{C}_{\Delta_{t-1},t-1}) | \Delta_{t-1} = i) \\
&\quad = \mathrm{E}(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-1}| \otimes [(\boldsymbol{A}|\boldsymbol{Z}_{t-1}| - \tilde{\boldsymbol{A}} \boldsymbol{Z}_{t-1})(\boldsymbol{e}'_{\Delta_{t-1}} \otimes \boldsymbol{I}_M) + \boldsymbol{B}] | \Delta_{t-1} = i) \\
&\quad = \mathrm{E}(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-1}| \otimes (\boldsymbol{A}|\boldsymbol{Z}_{t-1}|(\boldsymbol{e}'_{\Delta_{t-1}} \otimes \boldsymbol{I}_M) + \boldsymbol{B}) | \Delta_{t-1} = i) \\
&\quad = (\boldsymbol{I}_{kM} \otimes \boldsymbol{e}'_i \otimes \boldsymbol{A}) \mathrm{E}(\boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-1}| \otimes \boldsymbol{I}_k \otimes |\boldsymbol{Z}_{t-1}| | \Delta_{t-1} = i) + \kappa_1 (\boldsymbol{I}_{kM} \otimes \boldsymbol{B}) \\
&\quad = (\boldsymbol{I}_{kM} \otimes \boldsymbol{e}'_i \otimes \boldsymbol{A}) \tilde{\mathrm{Y}}(i) + \kappa_1 (\boldsymbol{I}_{kM} \otimes \boldsymbol{B}).
\end{aligned}
$$

Finally,

$$\mathrm{E}(|\boldsymbol{\epsilon}_t| |\boldsymbol{\epsilon}_{t-\tau}|') = \kappa_1 (\mathrm{vec}(\boldsymbol{I}_k)' \otimes \boldsymbol{I}_M) \boldsymbol{S}(\tau) (\mathrm{vec}(\boldsymbol{I}_k) \otimes \boldsymbol{I}_M),$$

and the autocorrelation function can be computed.

B.3 Covariance matrix forecasts

Let $\boldsymbol{\pi}_t = (p_t(\Delta_t = 1), \ldots, p_t(\Delta_t = k))'$ denote the *filtered* regime probabilities at time t, i.e. the probability distribution of the chain at time t *conditional* on the history of the process up to time t,[15] and suppose we are given an initial vector $\boldsymbol{X}_{t+1}$ (which is known at time t).

Define

$$\boldsymbol{Y}_t = \begin{pmatrix} \boldsymbol{X}_t \\ \mathrm{vec}(\boldsymbol{X}_t \boldsymbol{X}'_t) \end{pmatrix}, \quad \tilde{\boldsymbol{\omega}} = \begin{pmatrix} \boldsymbol{\omega} \\ \boldsymbol{\omega} \otimes \boldsymbol{\omega} \end{pmatrix},$$

dollar

$$\tilde{\boldsymbol{C}}_{\Delta_t,t} = \begin{pmatrix} \boldsymbol{C}_{\Delta_t,t} & \boldsymbol{0}_{kM \times k^2 M^2} \\ \boldsymbol{C}_{\Delta_t,t} \otimes \boldsymbol{\omega} + \boldsymbol{\omega} \otimes \boldsymbol{C}_{\Delta_t,t} & \boldsymbol{C}_{\Delta_t,t} \otimes \boldsymbol{C}_{\Delta_t,t} \end{pmatrix},$$

so that

$$\boldsymbol{Y}_t = \tilde{\boldsymbol{\omega}} + \tilde{\boldsymbol{C}}_{\Delta_{t-1},t-1} \boldsymbol{Y}_{t-1}. \tag{22}$$

Upon repeated substitution in (22), we can write

$$\boldsymbol{Y}_{t+d} = \sum_{\ell=1}^{d-1} \left\{ \prod_{i=1}^{\ell-1} \tilde{\boldsymbol{C}}_{\Delta_t+d-i,t+d-i} \right\} \tilde{\boldsymbol{\omega}} + \left\{ \prod_{i=1}^{d-1} \tilde{\boldsymbol{C}}_{\Delta_t+d-i,t+d-i} \right\} \boldsymbol{Y}_{t+1}.$$

Let $\underline{\Delta}_t = \{\Delta_s : s \leq t\}$. Then we have, taking expectations with respect to $\{\boldsymbol{\xi}_t\}$,

$$\mathrm{E}_t(\boldsymbol{Y}_{t+d}|\underline{\Delta}_{t+d-1}) = \sum_{\ell=1}^{d-1} \left\{\prod_{i=1}^{\ell-1} \tilde{\boldsymbol{C}}(\Delta_{t+d-i})\right\} \tilde{\boldsymbol{\omega}} + \left\{\prod_{i=1}^{d-1} \tilde{\boldsymbol{C}}(\Delta_{t+d-i})\right\} \boldsymbol{Y}_{t+1},$$

where, as in (24), $\tilde{\boldsymbol{C}}(j) = \mathrm{E}(\tilde{\boldsymbol{C}}_{jt}|\Delta_t = j)$. From Lemma 1 in Francq and Zakoïan (2005), we have

$$
\begin{aligned}
\tilde{\boldsymbol{Y}}_t(d) &:= \begin{pmatrix} p_t(\Delta_{t+d-1}=1)\mathrm{E}_t(\boldsymbol{Y}_{t+d}|\Delta_{t+d-1}=1) \\ p_t(\Delta_{t+d-1}=2)\mathrm{E}_t(\boldsymbol{Y}_{t+d}|\Delta_{t+d-1}=2) \\ \vdots \\ p_t(\Delta_{t+d-1}=k)\mathrm{E}_t(\boldsymbol{Y}_{t+d}|\Delta_{t+d-1}=k) \end{pmatrix} \\
&= \sum_{\ell=1}^{d-1} \mathbb{P}_{\tilde{C}}^{\ell-1}(\boldsymbol{\pi}_{t+d-\ell} \otimes \tilde{\boldsymbol{\omega}}) + \mathbb{P}_{\tilde{C}}^{d-1}(\boldsymbol{\pi}_t \otimes \boldsymbol{Y}_{t+1}),
\end{aligned}
$$

where $\boldsymbol{\pi}_{t+d-\ell} = \boldsymbol{P}^{d-\ell}\,\boldsymbol{\pi}_t$. Define the matrix

$$\mathbb{I} = \boldsymbol{I}_k \otimes (\boldsymbol{0}_{k^2M^2 \times kM}, \boldsymbol{I}_{k^2M^2}).$$

Then the d–step-ahead covariance matrix forecast is given by

$$
\begin{aligned}
\mathrm{E}_t\left(\mathrm{vec}\left(\boldsymbol{\epsilon}_{t+d}\boldsymbol{\epsilon}_{t+d}'\right)\right) &= \sum_{j=1}^k p_t(\Delta_{t+d}=j)\mathrm{E}_t(\mathrm{vec}(\boldsymbol{\epsilon}_{t+d}\boldsymbol{\epsilon}_{t+d}')|\Delta_{t+d}=j) \\
&= \sum_{j=1}^k p_t(\Delta_{t+d}=j) \\
&\times \mathrm{E}_t\left\{\mathrm{vec}\left[(\boldsymbol{e}_j \otimes \boldsymbol{I}_M)'(\boldsymbol{I}_k \otimes \boldsymbol{Z}_{t+d})\boldsymbol{X}_{t+d}\boldsymbol{X}_{t+d}'(\boldsymbol{I}_k \otimes \boldsymbol{Z}_{t+d})(\boldsymbol{e}_j \otimes \boldsymbol{I}_M)\right]|\Delta_{t+d}=j\right\} \\
&= \sum_{j=1}^k p_t(\Delta_{t+d}=j)\left(\boldsymbol{e}_j \otimes \boldsymbol{I}_M \otimes \boldsymbol{e}_j \otimes \boldsymbol{I}_M\right)' \\
&\times \mathrm{E}\left(\boldsymbol{I}_k \otimes \boldsymbol{Z}_{t+d} \otimes \boldsymbol{I}_k \otimes \boldsymbol{Z}_{t+d}|\Delta_{t+d}=j\right)\mathrm{E}_t[\mathrm{vec}(\boldsymbol{X}_{t+d}\boldsymbol{X}_{t+d}')|\Delta_{t+d}=j] \\
&= \sum_{j=1}^k \left(\boldsymbol{e}_j \otimes \boldsymbol{I}_M \otimes \boldsymbol{e}_j \otimes \boldsymbol{I}_M\right)' \\
&\times \sum_{i=1}^k p_{ij}\mathrm{Y}(j)p_t(\Delta_{t+d-1}=i)\mathrm{E}_t[\mathrm{vec}(\boldsymbol{X}_{t+d}\boldsymbol{X}_{t+d}')|\Delta_{t+d-1}=i] \\
&= \left\{\sum_{j=1}^k [\boldsymbol{e}_j' \otimes \left(\boldsymbol{e}_j \otimes \boldsymbol{I}_M \otimes \boldsymbol{e}_j \otimes \boldsymbol{I}_M\right)']\right\} \mathbb{P}_{\mathrm{Y}}\mathbb{I}\tilde{\boldsymbol{Y}}_t(d).
\end{aligned}
\tag{23}
$$

Vector $\tilde{\boldsymbol{Y}}_t(d)$ in (23) can be calculated recursively, with starting value $\tilde{\boldsymbol{Y}}_t(1) = \boldsymbol{\pi}_t \otimes \boldsymbol{Y}_{t+1}$. Namely, for $d \geq 2$,

$$
\begin{aligned}
\tilde{\boldsymbol{Y}}_t(d) &= \sum_{\ell=1}^{d-1} \mathbb{P}_{\tilde{C}}^{\ell-1}(\boldsymbol{\pi}_{t+d-\ell} \otimes \tilde{\boldsymbol{\omega}}) + \mathbb{P}_{\tilde{C}}^{d-1}(\boldsymbol{\pi}_t \otimes \boldsymbol{Y}_{t+1}) \\
&= \boldsymbol{\pi}_{t+d-1} \otimes \tilde{\boldsymbol{\omega}} + \mathbb{P}_{\tilde{C}}\left\{\sum_{\ell=1}^{d-2} \mathbb{P}_{\tilde{C}}^{\ell-1}(\boldsymbol{\pi}_{t+(d-1)-\ell} \otimes \tilde{\boldsymbol{\omega}}) + \mathbb{P}_{\tilde{C}}^{d-2}(\boldsymbol{\pi}_t \otimes \boldsymbol{Y}_{t+1})\right\} \\
&= \boldsymbol{\pi}_{t+d-1} \otimes \tilde{\boldsymbol{\omega}} + \mathbb{P}_{\tilde{C}}\tilde{\boldsymbol{Y}}_t(d-1).
\end{aligned}
\tag{24}
$$

Equations (23) and (24) provide a convenient scheme for calculation of covariance matrix forecasts.

C Lagrange multiplier (LM) test for constant within-regime correlations

The MS-GARCH model with constant within-regime correlations is attractive since it is analytically tractable. E.g. straightforward-to-check conditions for stationarity have been obtained, as well as a simple recursion for calculating multi-step conditional covariance matrices, which is crucial for mean-variance portfolio optimization. Moreover, in some simple cases, such as Pelletier's (2006) model with regime-independent GARCH dynamics, estimation in high dimensions is feasible via a two-step procedure with an embedded EM algorithm. Despite its convenience, it is still desirable to test whether the assumption of constant within-regime correlation matrices is tenable, since otherwise further improvement of out-of-sample portfolio selection might be feasible by extending the model to allow for within-regime correlation dynamics as in Billio and Caporin (2005) and Otranto (2010). Using results of Hamilton (1996), we extend the Lagrange Multiplier (LM) test devised by Tse (2000) for constant conditional correlations in multivariate GARCH models. In Tse (2000) the LM test is derived under normality of the innovations, but he reports simulations indicating it being quite robust against non-normality. However, in view of the discussion in Section 4.1, this cannot be expected to hold for MS-GARCH processes, and thus we derive the test allowing for Student's t errors. The test under normality is then straightforwardly obtained if the degrees of freedom $\nu \to \infty$.

For the volatility dynamics, we assume that the conditional standard deviation of asset i in regime j, σ_{ijt}, is described by a standard (symmetric) AGARCH(1,1) process, i.e. [16]

$$\sigma_{ijt} = \omega_{ij} + a_{ij}|\epsilon_{i,t-1}| + b_{ij}\sigma_{ij,t-1}, \quad i = 1, \dots, M, \quad j = 1, \dots, k. \tag{25}$$

The conditional correlation matrix in regime j is

$$\boldsymbol{R}_{jt} = (\rho_{i\ell,jt})_{i,\ell=1,\dots,M}, \quad j = 1, \dots, k.$$

where, as in Tse (2000), correlations evolve according to[17]

$$\rho_{i\ell,jt} = \bar{\rho}_{i\ell,j} + \delta_{i\ell,j}\epsilon_{i,t-1}\epsilon_{\ell,t-1}, \quad i = 1, \dots, M-1, \quad \ell = i+1, \dots, M, \quad j = 1, \dots, k. \tag{26}$$

The null hypothesis of constant conditional within-regime correlations corresponds to

$$\mathrm{H}_0 : \delta_{i\ell,j} = 0, \quad i = 1, \dots, M-1, \quad \ell = i+1, \dots, M, \quad j = 1, \dots, k. \tag{27}$$

We distinguish between the following cases, which differ in the specification of the alternative hypothesis:

a. The conditional correlation dynamics are unrestricted across regimes. In this case, under (27), the LM test statistic is asymptotically distributed as χ^2 with $kM(M-1)/2$ degrees of freedom.

b. The conditional correlation dynamics are the same across regimes, i.e. $\delta_{i\ell,1} = \delta_{i\ell,2} = \cdots = \delta_{i\ell,k}$, $i = 1, \dots, M-1$, $\ell = i+1, \dots, M$. In this case, under (27), the LM test statistic is asymptotically distributed as χ^2 with $M(M-1)/2$ degrees of freedom.

For the purpose of the current section, it is convenient to decompose the parameter vector of the model as $\boldsymbol{\theta} = (\text{vec}(\boldsymbol{P})', \boldsymbol{\vartheta}')'$,[18] where $\boldsymbol{\vartheta}$ consists of the parameters of the conditional regime densities, i.e. $\boldsymbol{\vartheta} = (\boldsymbol{\vartheta}_1', \dots, \boldsymbol{\vartheta}_k', \nu)'$, where ν is the (common) shape parameter of the t distribution and $\boldsymbol{\vartheta}_j = (\boldsymbol{\vartheta}_{1j}', \dots, \boldsymbol{\vartheta}_{Mj}', \boldsymbol{\rho}_j', \boldsymbol{\delta}_j')'$, $j = 1, \dots, k$, where $\boldsymbol{\vartheta}_{ij} = (\omega_{ij}, a_{ij}, b_{ij})'$, and $\boldsymbol{\rho}_j$ and $\boldsymbol{\delta}_j$ are the $M(M-1)/2$ vectors which stack, respectively, parameters $\bar{\rho}_{i\ell,j}$ and $\delta_{i\ell,j}$ in Equation (26), i.e.

$$\begin{aligned}\boldsymbol{\rho}_j &= (\bar{\rho}_{12,j}, \bar{\rho}_{13,j}, \dots, \bar{\rho}_{1M,j}, \bar{\rho}_{23,j}, \dots, \bar{\rho}_{2M,j}, \dots, \bar{\rho}_{M-1,M,j})', \\ \boldsymbol{\delta}_j &= (\delta_{12,j}, \delta_{13,j}, \dots, \delta_{1M,j}, \delta_{23,j}, \dots, \delta_{2M,j}, \dots, \delta_{M-1,M,j})'.\end{aligned}$$

The log-likelihood of the model for a sample of size T is given by

$$\begin{aligned}\log L(\boldsymbol{\theta}) &= \sum_{t=1}^{T} \log f(\epsilon_t|\Omega_{t-1}; \boldsymbol{\theta}) \\ &= \sum_{t=1}^{T} \log\left\{\sum_{j=1}^{k} p(\Delta_t = j|\Omega_{t-1}; \boldsymbol{\theta})f(\epsilon_t|\Omega_{t-1}, \Delta_t = j; \boldsymbol{\vartheta}_j, \nu)\right\},\end{aligned} \tag{28}$$

where $\Omega_t = \{\boldsymbol{\epsilon}_t, \boldsymbol{\epsilon}_{t-1}, \dots, \boldsymbol{\epsilon}_0\}$, $p(\Delta_t = j|\Omega_{t-1}; \boldsymbol{\theta})$ are the one-step predicted regime inferences (cf. Hamilton, 1994, Ch. 22), and the conditional regime densities are

$$\begin{aligned}f(\epsilon_t|\Omega_{t-1}, \Delta_t = j; \boldsymbol{\vartheta}_j, \nu) &= \frac{\Gamma\left(\frac{\nu+M}{2}\right)}{\Gamma(\nu/2)(\pi(\nu-2))^{M/2}|\boldsymbol{R}_{jt}|^{1/2}\prod_{i=1}^{M}\sigma_{ijt}}\left\{1 + \frac{d_{jt}^2}{\nu-2}\right\}^{-(\nu+M)/2} \\ &=: f_{jt}(\boldsymbol{\vartheta}_j, \nu),\end{aligned} \tag{29}$$

where the squared Mahalanobis distance

$$d_{jt}^2 = \boldsymbol{\epsilon}_t'\boldsymbol{D}_{jt}^{-1}\boldsymbol{R}_{jt}^{-1}\boldsymbol{D}_{jt}^{-1}\boldsymbol{\epsilon}_t = \boldsymbol{\epsilon}_{jt}^{\star'}\boldsymbol{R}_{jt}^{-1}\boldsymbol{\epsilon}_{jt}^{\star} = \boldsymbol{\epsilon}_{jt}^{\star'}\tilde{\boldsymbol{\epsilon}}_{jt},$$

with

$$\boldsymbol{\epsilon}_{jt}^{\star} = \boldsymbol{D}_{jt}^{-1}\boldsymbol{\epsilon}_t, \quad \tilde{\boldsymbol{\epsilon}}_{jt} = \boldsymbol{R}_{jt}^{-1}\boldsymbol{\epsilon}_{jt}^{\star}, \quad j = 1, \dots, k. \tag{30}$$

The regime-specific log-density for observation t is

$$
\begin{aligned}
\log f_{jt}(\boldsymbol{\vartheta}_j, \nu) &= -\frac{M}{2}(\log\pi + \log(\nu-2)) + \log\Gamma\left(\frac{\nu+M}{2}\right) - \log\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\log|\boldsymbol{R}_{jt}| \\
&\quad - \sum_{i=1}^{M}\log\sigma_{ijt} + \frac{\nu+M}{2}\log\left(1 + \frac{d_{jt}^2}{\nu-2}\right) \\
&= -\frac{M}{2}\log\pi + \log\Gamma\left(\frac{\nu+M}{2}\right) - \log\Gamma\left(\frac{\nu}{2}\right) + \frac{\nu}{2}\log(\nu-2) - \frac{1}{2}\log|\boldsymbol{R}_{jt}| \\
&\quad - \sum_{i=1}^{M}\log\sigma_{ijt} - \frac{\nu+M}{2}\log(\nu-2+d_{jt}^2), \quad j = 1,\ldots,k.
\end{aligned}
\tag{31}
$$

The partial derivatives of (31) are obtained as

$$
\begin{aligned}
\frac{\partial\log f_{jt}(\boldsymbol{\vartheta}_j,\nu)}{\partial\nu} &= \frac{1}{2}\left[\psi\left(\frac{\nu+M}{2}\right) - \psi\left(\frac{\nu}{2}\right)\right] - \frac{1}{2}\log\left(1 + \frac{d_{jt}^2}{\nu-2}\right) \\
&\quad + \frac{1}{2}\left[\frac{\nu}{\nu-2} - \frac{\nu+M}{\nu-2+d_{jt}^2}\right], \\
\frac{\partial\log f_{jt}(\boldsymbol{\vartheta}_j,\nu)}{\partial\boldsymbol{\vartheta}_{ij}} &= \frac{\partial\sigma_{ijt}}{\partial\boldsymbol{\nu}_{ij}}\frac{1}{\sigma_{ijt}}\left(\frac{\epsilon_{ijt}^\star\tilde{\epsilon}_{ijt}(\nu+M)}{\nu-2+d_{jt}^2} - 1\right), \quad i = 1,\ldots,M, \\
\frac{\partial\log f_{jt}(\boldsymbol{\vartheta}_j,\nu)}{\partial\boldsymbol{\rho}_j} &= \frac{1}{2}\boldsymbol{U}\left(\frac{\nu+M}{\nu-2+d_{jt}^2}(\tilde{\boldsymbol{\epsilon}}_{jt}\otimes\tilde{\boldsymbol{\epsilon}}_{jt}) - \mathrm{vec}\boldsymbol{R}_{jt}^{-1}\right), \\
\frac{\partial\log f_{jt}(\boldsymbol{\vartheta}_j,\nu)}{\partial\boldsymbol{\delta}_j} &= \frac{1}{2}\boldsymbol{U}\left\{\left(\frac{\nu+M}{\nu-2+d_{jt}^2}(\tilde{\boldsymbol{\epsilon}}_{jt}\otimes\tilde{\boldsymbol{\epsilon}}_{jt}) - \mathrm{vec}\boldsymbol{R}_{jt}^{-1}\right)\odot(\boldsymbol{\epsilon}_{t-1}\otimes\boldsymbol{\epsilon}_{t-1})\right\},
\end{aligned}
\tag{32}
$$

where $\psi(x) = \mathrm{d}\log\Gamma(x)/\mathrm{d}x$ is the logarithmic derivative of the gamma function (the digamma function); $\epsilon_{ijt}^\star$ and $\tilde{\epsilon}_{ijt}$ are, respectively, the ith elements of vectors $\boldsymbol{\epsilon}_{jt}^\star$ and $\tilde{\boldsymbol{\epsilon}}_{jt}$ defined in (30) $i = 1,\ldots,M$; and the $M(M-1)/2 \times M^2$ matrix $\boldsymbol{U}$ is defined as in Silvennoinen and Teräsvirta (2009a), i.e. with its $[(i-1)M - i(i+1)/2 + \ell]$th row given by

$$
\mathrm{vec}(\tilde{\boldsymbol{e}}_i\tilde{\boldsymbol{e}}_\ell' + \tilde{\boldsymbol{e}}_\ell\tilde{\boldsymbol{e}}_i')', \quad i = 1,\ldots,M-1, \quad \ell = i+1,\ldots,M,
$$

where $\tilde{\boldsymbol{e}}_i$ is the ith column of the M–dimensional identity matrix. The derivative in (32) is

$$
\frac{\partial\sigma_{ijt}}{\partial\boldsymbol{\vartheta}_{ij}} = \boldsymbol{\eta}_{ij,t-1} + b_{ij}\frac{\partial\sigma_{ij,t-1}}{\partial\boldsymbol{\vartheta}_{ij}}, \quad t = 2,\ldots,T,
\tag{33}
$$

where $\boldsymbol{\eta}_{ijt} = (1, |\epsilon_{it}|, \sigma_{ijt})'$, and the starting value in recursion (33) is

$$
\frac{\partial\sigma_{ij,t=1}}{\partial\boldsymbol{\vartheta}_{ij}} = (1, |\epsilon_{i0}|, \sigma_{ij0})',
\tag{34}
$$

where we initialize all regime-specific conditional standard deviations with the sample standard deviation, i.e. in (34),

$$
\sigma_{ij0} = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\epsilon_{it}^2}, \quad i = 1,\ldots,M, \quad j = 1,\ldots,k.
$$

The score of the tth observation is given by the derivative of the conditional log-density of $\boldsymbol{\epsilon}_t$ as given in (28) and (29),

$$
\frac{\partial\log f(\epsilon_t|\Omega_{t-1};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \frac{\partial\log\left\{\sum_{j=1}^{k}p(\Delta_t = j|\Omega_{t-1};\boldsymbol{\theta})f_{jt}(\boldsymbol{\vartheta}_j,\nu)\right\}}{\partial\boldsymbol{\theta}}.
\tag{35}
$$

Hamilton (1996) has shown that the derivatives in (35) involving elements of $\boldsymbol{\vartheta}$ can be evaluated as

$$
\begin{aligned}
\frac{\partial\log f(\epsilon_t|\Omega_{t-1};\boldsymbol{\theta})}{\partial\boldsymbol{\vartheta}} &= \sum_{j=1}^{k}\frac{\partial\log f_{jt}(\boldsymbol{\vartheta}_j,\nu)}{\partial\boldsymbol{\vartheta}}p(\Delta_t = j|\Omega_t;\boldsymbol{\theta}) \\
&\quad + \sum_{\tau=1}^{t-1}\sum_{j=1}^{k}\frac{\partial\log f_{j\tau}(\boldsymbol{\vartheta}_j,\nu)}{\partial\boldsymbol{\vartheta}}\left[p(\Delta_\tau = j|\Omega_t;\boldsymbol{\theta}) - p(\Delta_\tau = j|\Omega_{t-1};\boldsymbol{\theta})\right], \\
t &= 1,\ldots,T,
\end{aligned}
\tag{36}
$$

where the second line in (36) is set to zero for $t = 1$. For $\tau = t$ and $\tau < t$, the regime probabilities $p(\Delta_\tau = j|\Omega_t; \boldsymbol{\theta})$ in (36) are known as *filtered* and *smoothed* regime inferences, respectively; see, e.g. Hamilton (1994, Ch. 22) for their recursive calculation. We initialize these recursions by assuming that Δ_1 is drawn from the stationary distribution of the chain, i.e.

$$
\begin{aligned}
p(\Delta_1 = 1|\boldsymbol{\theta}) &= \pi_{1,\infty} = \frac{1-p_{22}}{2-p_{11}-p_{22}}, \\
p(\Delta_1 = 2|\boldsymbol{\theta}) &= \pi_{2,\infty} = \frac{1-p_{11}}{2-p_{11}-p_{22}}.
\end{aligned}
\tag{37}
$$

Note that, in (36), many terms are zero for parameters that appear in only one regime; e.g. if all parameters except ν are regime-specific, then

$$
\begin{aligned}
\frac{\partial \log f(\epsilon_t|\Omega_{t-1};\boldsymbol{\theta})}{\partial \boldsymbol{\vartheta}_j} &= \frac{\partial \log f_{jt}(\boldsymbol{\vartheta}_j, \nu)}{\partial \boldsymbol{\vartheta}_j} p(\Delta_t = j|\Omega_t; \boldsymbol{\theta}) \\
&+ \sum_{\tau=1}^{t-1} \frac{\partial \log f_{j\tau}(\boldsymbol{\vartheta}_j, \nu)}{\partial \boldsymbol{\vartheta}_j} \left[p(\Delta_\tau = j|\Omega_t; \boldsymbol{\theta}) - p(\Delta_\tau = j|\Omega_{t-1}; \boldsymbol{\theta}) \right] \\
t &= 1, \ldots, T, \quad j = 1, \ldots, k.
\end{aligned}
$$

For the score with respect to the parameters of $\boldsymbol{P}$, see Hamilton (1996).[19]

Now let $\boldsymbol{S}$ be the $T \times N$ matrix (where N is the dimension of $\boldsymbol{\theta}$) the tth row of which is given by (the transpose of) the score (35), $t = 1, \ldots, T$, and let $\hat{\boldsymbol{S}}$ be $\boldsymbol{S}$ evaluated at $\hat{\boldsymbol{\theta}}^c$ the constrained MLE of $\boldsymbol{\theta}$ under (27). Then the LM test statistic for H_0 given by (27) can be calculated as in the outer gradient product form as (cf. Hamilton, 1996)

$$
\text{LM} = \mathbf{1}_T' \hat{\boldsymbol{S}} (\hat{\boldsymbol{S}}' \hat{\boldsymbol{S}})^{-1} \hat{\boldsymbol{S}}' \mathbf{1}_T \xrightarrow{d} \chi^2(c_0),
\tag{38}
$$

where $\mathbf{1}_T$ is a T–dimensional column of ones, and c_0 is the number of parameter constraints under H_0, see the discussion following Equation (27). Note that the elements of $\mathbf{1}_T' \hat{\boldsymbol{S}}$ corresponding to unrestricted parameters are zero, so that a slight simplification of the LM statistic (38) can be obtained (cf. Tse, 2000).

D The mixed normal BEKK-GARCH (MixN BEKK) model

In this appendix, we briefly describe the mixed normal BEKK-GARCH (MixN BEKK) model applied in Section 4.3, detail the calculation of multi-step-ahead conditional covariance matrices, and discuss a few of its characteristics in comparison with those reported in Table 5.

The asymmetric normal mixture GARCH model in Haas, Mittnik, and Paolella (2009) is an asymmetric extension of the multivariate normal mixture GARCH process of Bauwens, Hafner, and Rombouts (2007).[20] With two mixture components, as in Section 4.3, the conditional distribution of the M–dimensional vector of shocks, $\boldsymbol{\epsilon}_t$, is a two-component normal mixture distribution with constant mixing weights $p_1 \in (0, 1)$ and $p_2 = 1 - p_1$, i.e. the conditional density is

$$
f_{t-1}(\boldsymbol{\epsilon}_t) = p_1 \phi(\boldsymbol{\epsilon}_t; \boldsymbol{\mu}_1, \boldsymbol{H}_{1t}) + p_2 \phi(\boldsymbol{\epsilon}_t; \boldsymbol{\mu}_2, \boldsymbol{H}_{2t}),
\tag{39}
$$

where $\phi(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{H})$ is the multivariate normal density with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{H}$, $\boldsymbol{\mu}_2 = -p_1\boldsymbol{\mu}_2/p_2$ (so that $E_{t-1}(\boldsymbol{\epsilon}_t) = E(\boldsymbol{\epsilon}_t) = \mathbf{0}$), and $\boldsymbol{H}_{1t}$ and $\boldsymbol{H}_{2t}$ are the conditional *component covariance matrices*. As in Haas, Mittnik, and Paolella (2009), we specify the latter as asymmetric BEKK processes, i.e. (cf. Engle & Kroner, 1995),

$$
\boldsymbol{H}_{jt} = \boldsymbol{C}_j \boldsymbol{C}_j' + \boldsymbol{A}_j (\boldsymbol{\epsilon}_{t-1} - \boldsymbol{\theta}_j)(\boldsymbol{\epsilon}_{t-1} - \boldsymbol{\theta}_j)' \boldsymbol{A}_j' + \boldsymbol{B}_j \boldsymbol{H}_{j,t-1} \boldsymbol{B}_j', \quad j = 1, 2,
\tag{40}
$$

where $\boldsymbol{C}_j$ is lower triangular with positive diagonal, $j = 1, 2$. In line with the other GARCH specifications considered in this paper, the MixN BEKK model used in Section 4.3 imposes the restrictions that

i. in (39), $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \mathbf{0}$,

ii. in (40), $\boldsymbol{A}_j$ and $\boldsymbol{B}_j$, $j = 1, 2$, are diagonal, and $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$.

To compute covariance matrix forecasts, write (40) in vech form,

$$
\begin{aligned}
\boldsymbol{h}_{jt} &= \boldsymbol{\omega}_j + \tilde{\boldsymbol{A}}_j \mathrm{vech}(\boldsymbol{\epsilon}_{t-1} - \boldsymbol{\theta}_j)(\boldsymbol{\epsilon}_{t-1} - \boldsymbol{\theta}_j)' + \tilde{\boldsymbol{B}}_j \boldsymbol{h}_{j,t-1} \\
&= \tilde{\boldsymbol{\omega}}_j + \tilde{\boldsymbol{A}}_j \boldsymbol{\eta}_{t-1} + \tilde{\boldsymbol{B}}_j \boldsymbol{h}_{j,t-1} - \boldsymbol{G}_j \boldsymbol{\epsilon}_{t-1}, \quad j = 1, 2,
\end{aligned}
\tag{41}
$$

where $\boldsymbol{\eta}_t = \mathrm{vech}(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t)'$, $\boldsymbol{h}_{jt} = \mathrm{vech}(\boldsymbol{H}_{jt})$, $\tilde{\boldsymbol{A}}_j = \boldsymbol{D}_M^+ (\boldsymbol{A}_j \otimes \boldsymbol{A}_j) \boldsymbol{D}_M$, $\tilde{\boldsymbol{B}}_j = \boldsymbol{D}_M^+ (\boldsymbol{B}_j \otimes \boldsymbol{B}_j) \boldsymbol{D}_M$, $\tilde{\boldsymbol{\omega}}_j = \boldsymbol{\omega}_j + \tilde{\boldsymbol{A}}_j \mathrm{vech}(\boldsymbol{\theta}_j \boldsymbol{\theta}_j')$, $\boldsymbol{G}_j = 2 \tilde{\boldsymbol{A}}_j \boldsymbol{D}_M^+ (\boldsymbol{I}_M \otimes \boldsymbol{\theta}_j)$, $j = 1, 2$, $\boldsymbol{D}_M$ denotes the $M^2 \times M(M+1)/2$ *duplication matrix* such that, for symmetric $M \times M$ matrix A, $\boldsymbol{D}_M \mathrm{vech}(\boldsymbol{A}) = \mathrm{vec}(\boldsymbol{A})$, and $\boldsymbol{D}_M^+ = (\boldsymbol{D}_M' \boldsymbol{D}_M)^{-1} \boldsymbol{D}_M'$. Matrix $\boldsymbol{D}_M^+$ has the properties $\boldsymbol{D}_M^+ \mathrm{vech}(\boldsymbol{A}) = \mathrm{vec}(\boldsymbol{A})$ for symmetric A, and $2 \boldsymbol{D}_M^+ \mathrm{vec}(\boldsymbol{A}) = \mathrm{vech}(\boldsymbol{A} + \boldsymbol{A}')$ (this explains the expression for $\boldsymbol{G}_j$; cf. Magnus, 1988, p. 80).

In more compact form, (41) can be written as

$$
\boldsymbol{h}_t = \tilde{\boldsymbol{\omega}} + \tilde{\boldsymbol{A}} \boldsymbol{\eta}_{t-1} + \tilde{\boldsymbol{B}} \boldsymbol{h}_{t-1} - \boldsymbol{G} \boldsymbol{\epsilon}_{t-1},
$$

where

$$
\boldsymbol{h}_t = \begin{pmatrix} \boldsymbol{h}_{1t} \\ \boldsymbol{h}_{2t} \end{pmatrix}, \quad \tilde{\boldsymbol{A}} = \begin{pmatrix} \tilde{\boldsymbol{A}}_1 \\ \tilde{\boldsymbol{A}}_2 \end{pmatrix}, \quad \tilde{\boldsymbol{B}} = \begin{pmatrix} \tilde{\boldsymbol{B}}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \tilde{\boldsymbol{B}}_2 \end{pmatrix}, \quad \boldsymbol{G} = \begin{pmatrix} \boldsymbol{G}_1 \\ \boldsymbol{G}_2 \end{pmatrix}.
$$

We have[21]

$$
\mathrm{E}_t(\boldsymbol{h}_{t+d}) = \tilde{\boldsymbol{\omega}} + \tilde{\boldsymbol{C}} \mathrm{E}_t(\boldsymbol{h}_{t+d-1}),
$$

where

$$
\tilde{\boldsymbol{C}} = \boldsymbol{p}' \otimes \tilde{\boldsymbol{A}} + \tilde{\boldsymbol{B}},
$$

with the vector of mixing weights $\boldsymbol{p} = (p_1, p_2)'$, and by repeated substitution,

$$
\mathrm{E}_t(\boldsymbol{h}_{t+d}) = \sum_{i=0}^{d-2} \tilde{\boldsymbol{C}}^i \tilde{\boldsymbol{\omega}} + \tilde{\boldsymbol{C}}^{d-1} \boldsymbol{h}_{t+1} = \bar{\boldsymbol{h}} + \tilde{\boldsymbol{C}}^{d-1}(\boldsymbol{h}_{t+1} - \bar{\boldsymbol{h}}),
$$

where $\bar{\boldsymbol{h}} = \mathrm{E}(\boldsymbol{h}_t) = (\boldsymbol{I}_{2N} - \tilde{\boldsymbol{C}})^{-1} \tilde{\boldsymbol{\omega}}$, and $N = M(M+1)/2$. Consequently, the (vech of the) d–step-ahead conditional covariance matrix is

$$
\mathrm{E}_t(\boldsymbol{\eta}_{t+d}) = (\boldsymbol{p}' \otimes \boldsymbol{I}_N) \mathrm{E}_t(\boldsymbol{h}_{t+d}) = \mathrm{E}(\boldsymbol{\eta}_t) + (\boldsymbol{p}' \otimes \boldsymbol{I}_N) \tilde{\boldsymbol{C}}^{d-1}(\boldsymbol{h}_{t+1} - \bar{\boldsymbol{h}}),
$$

and for the cumulative shock, $\boldsymbol{\epsilon}_{t:t+D} = \sum_{d=1}^{D} \boldsymbol{\epsilon}_{t+d}$,

$$
\mathrm{vech}\left(\mathrm{Cov}_t\left(\boldsymbol{\epsilon}_{t:t+D}\right)\right) = D\mathrm{E}(\boldsymbol{\eta}_t) + (\boldsymbol{p}' \otimes \boldsymbol{I}_N)(\boldsymbol{I}_{2N} - \tilde{\boldsymbol{C}}^D)(\boldsymbol{I}_{2N} - \tilde{\boldsymbol{C}})^{-1}(\boldsymbol{h}_{t+1} - \bar{\boldsymbol{h}}).
$$

Bauwens, Hafner, and Rombouts (2007) and Haas, Mittnik, and Paolella (2009) consider the model only with Gaussian mixture components. However, it is clear that, as in (17), a unit-variance Student's t distribution can be assumed for the mixture components in (39) as well. However, as already reported in Haas, Mittnik, and Paolella (2004a) for daily stock returns and confirmed for our data, use of fat-tailed component densities typically does not lead to significant improvements in this kind of model. This is in sharp contrast to the Markov-switching models investigated herein and thus may appear a bit surprising at first glance. It can be explained by the fact that, in contrast to the Markov-switching process (with time-varying conditional regime probabilities), the independent switching process of the MixN BEKK model as such does not add to the dynamics of the conditional moment structure. Rather, the conditional mixed normal distribution in (39) mainly contributes to conditional leptokurticity and thus essentially serves the same purpose as a conditional Student's t distribution. As conditional mixed normality is often sufficient to capture the excess kurtosis in the data, the degrees of freedom parameter in t–mixture GARCH models tends to be somewhat ill-identified and unstable and erratic over time. We have observed the same phenomenon for our data and thus only use the *Gaussian* mixture GARCH in Section 4.3.

To illustrate, we briefly compare the parameter estimates over the entire sample as reported in Table 5 with those from the two-component MixN BEKK model, as shown in Table 9. Table 9 reports the estimated mixing weights p_1 and $p_2 = 1 - p_1$ as well as the unconditional component-specific variances and correlations implied by the estimated parameters. Comparing the results in Table 9 with those for the MS CCC models in Table 8, the following differences can be observed. First, in the MixN BEKK model, the correlation is essentially regime-independent. This can be explained by the fact that, in the MixN BEKK model, conditional correlations are driven by the BEKK dynamics for the component covariance matrices, whereas the (independent) switching process mainly accounts for the conditional leptokurtosis. The kurtosis of a two-component normal mixture distribution is particularly large when there are sizeable differences between the component variances and the mixing weight of the high-variance regime is small (cf. Timmermann, 2000). The differences in volatility between the regimes are more pronounced in the MixN BEKK process,[22] and the mixing weight of its high-volatility component is $p_2 = 0.126$, which is about one third of the unconditional high-volatility regime probability $\pi_{2,\infty} = 0.390$ in the Gaussian MS CCC model.

Table 9: Characteristics of the estimated MixN BEKK model.

$E(\epsilon_{1t}^2|\Delta = 1)$	$E(\epsilon_{2t}^2|\Delta_t = 1)$	$\text{Corr}(\epsilon_{1t}, \epsilon_{2t} \mid \Delta_t = 1)$	p_1
3.133	3.713	0.730	0.874
			(0.031)
$E(\epsilon_{1t}^2|\Delta_t = 2)$	$E(\epsilon_{2t}^2|\Delta_t = 2)$	$\text{Corr}(\epsilon_{1t}, \epsilon_{2t}|\Delta_t = 2)$	p_2
10.16	14.89	0.718	0.126
			(0.031)

Due to the independent switching, $E(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t'|\Delta_t = j) = E(\boldsymbol{H}_{jt})$, $j = 1, 2$, where the latter expectation is obtained by rearranging the elements of $E(\boldsymbol{h}_t) = (\boldsymbol{I}_{2N} - \tilde{\boldsymbol{C}})^{-1}\tilde{\boldsymbol{\omega}}$, where $\boldsymbol{h}_t = (\text{vech}(\boldsymbol{H}_{1t})', \text{vech}(\boldsymbol{H}_{2t})')'$. The component-specific correlations $\text{Corr}(\epsilon_{1t}, \epsilon_{2t}|\Delta_t = j)$, $j = 1, 2$, are calculated from these component-specific unconditional covariance matrices.

Notes

1 Francq, Roussignol, and Zakoïan (2001) and Francq and Zakoïan (2005) consider stationarity and moment properties of an alternative univariate MS GARCH process; see also Abramson and Cohen (2007).

2 Over a shorter time span, these indices were analyzed in Haas (2010), who shows that Pelletier's (2006) model improves global minimum variance portfolios relative to the standard CCC model.

3 Strong evidence for time-varying conditional correlations between these markets has been reported by Lee (2014).

4 The model with one regime is just the AGARCH version of the single-component CCC.

5 But note that expected regime durations may be subject to considerable estimation error; see the legend of Table 5.

6 The effect of "outliers" on parameter estimates in Gaussian Markov-switching models has already been reported in Ryden, Teräsvirta, and Åsbrink (1998).

7 In copula terminology, this reflects the fact that the t copula, in contrast to the Gaussian copula, displays *tail dependence* (cf. Jondeau & Rockinger, 2006; and Manner & Reznikova, 2012).

8 The conditional standard deviations have been initialized by appropriate unconditional expectations, i.e. if $\pi_{\ell t} = 1$, then we set $\sigma_{ij,t+1} = E(\sigma_{ij,t} \mid \Delta_{t-1} = \ell)$, $i, j, \ell = 1, 2$.

9 Silvennoinen and Teräsvirta (2009a) follow a similar approach by deriving LM tests for a CCC against a smooth transition conditional correlation model.

10 For purpose of comparison with the other models discussed herein, the DCC is likewise coupled with asymmetric absolute value GARCH for the volatilities.

11 As argued by Pelletier (2006), "[o]ne interesting implication of smoother patterns for the correlations is for the computation of VaR and portfolio allocation. If the time-varying correlations are smoother, then the gain from portfolio diversification will also be smoother which might imply a smoother pattern for the VaR and portfolio weights."

12 Those for the Gaussian DCC are quite similar and not shown.

13 The GMVP allows to focus on the covariance matrix without "having to specify the vector of conditional expected returns, which is more a task for the portfolio manager than a statistical problem" (Ledoit, Santa-Clara & Wolf, 2003).

14 Putintseva (2012) suggests to estimate a mixture GARCH model with dynamic conditional correlations by means of the expectation-maximization (EM) algorithm. This somewhat simplifies the problem, but high-dimensional numerical optimization is still required in each M-step.

15 Calculation of these probabilities is described in detail in Hamilton (1994) and Kim and Nelson (1999).

16 The extension to allow for asymmetric response of volatility is straightforward.

17 As noted by Silvennoinen and Teräsvirta (2009b), (26) does not represent a specific alternative to the CCC as positive definite correlation matrices are not guaranteed for every t: "For this reason we interpret the test as a general misspecification test". Smith (2008) reports simulation results in favor of Hamilton's (1996) LM specification tests for Markov-switching models.

18 Of course the last row of $\boldsymbol{P}$ is redundant.

19 In particular, for $k = 2$ regimes and initial probabilities (37), the relevant derivatives with respect to p_{11} and p_{22} are given by Equations (3.15) and (3.16) in Hamilton (1996).

20 These models are multivariate extensions of the models in Haas, Mittnik, and Paolella (2004a) and Alexander and Lazar (2006, 2009).

21 The following formulas apply to the case $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$; see Appendix C of Haas, Mittnik, and Paolella (2009) for the modifications required when different component means are allowed for.

22 The ratio of expected high– to expected low-volatility regime variance in the MixN BEKK model is 3.24 and 4.01 for the MSCI and FTSE EPRA/NAREIT returns, respectively; the corresponding numbers for the Gaussian MS CCC are 1.65 and 2.22.

Code and Datasets: The author(s) published code and data associated with this article is on Code Ocean, a computational reproducibility platform. We recommend Code Ocean to SNDE contributors who wish share, discover, and run code in published research articles. (See: https://doi.org/10.24433/CO.2bf91781-8f40-4e16-89bb-35c7ce77f43e).

References

Abramson, A., and I. Cohen. 2007. "On the Stationarity of Markov-Switching GARCH Processes." *Econometric Theory* 23: 485–500.

Alexander, C. 2008. *Practical Financial Econometrics*. Chichester: John Wiley & Sons.

Alexander, C., and E. Lazar. 2006. "Normal Mixture GARCH(1,1). Applications to Exchange Rate Modelling." *Journal of Applied Econometrics* 21: 307–336.

Alexander, C., and E. Lazar. 2009. "Modelling Regime-Specific Stock Price Volatility." *Oxford Bulletin of Economics and Statistics* 71: 761–797.

Ang, A., and A. Timmermann. 2012. "Regime Changes and Financial Markets." *Annual Review of Financial Economics* 4: 313–337.

Ardia, D. 2009. "Bayesian Estimation of a Markov-Switching Threshold Asymmetric GARCH Model with Student-*t* Innovations." *Econometrics Journal* 12: 105–126.

Augustyniak, M. 2014. "Maximum Likelihood Estimation of the Markov-Switching GARCH Model." *Computational Statistics and Data Analysis* 76: 61–75.

Bauwens, L., C. M. Hafner, and J. V. K. Rombouts. 2007. "Multivariate Mixed Normal Conditional Heteroskedasticity." *Computational Statistics and Data Analysis* 51: 3551–3566.

Billio, M., and M. Caporin. 2005. "Multivariate Markov Switching Dynamic Conditional Correlation GARCH Representations for Contagion Analysis." *Statistical Methods & Applications* 14: 145–161.

Bollerslev, T. 1990. "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model." *Review of Economics and Statistics* 73: 498–505.

Bougerol, P., and N. Picard. 1992. "Strict Stationarity of Generalized Autoregressive Processes." *Annals of Probability* 20: 1714–1730.

Brandt, A.. 1986. "The Stochastic Equation $Y_{n+1} = A_n Y_n + B_n$ with Stationary Coefficients." Advances in Applied Probability 18: 211–220.

Broda, S. A., M. Haas, J. Krause, M. S. Paolella, and S. C. Steude. 2013. "Stable Mixture GARCH Models." *Journal of Econometrics* 172: 292–306.

Bulla, J. 2011. "Hidden Markov Models with *t* Components. Increased Persistence and Other Aspects." *Quantitative Finance* 11: 459–475.

Campbell, R., K. Koedijk, and P. Kofman. 2002. "Increased Correlation in Bear Markets: A Downside Risk Perspective." *Financial Analysts Journal* 58: 87–94.

Case, B., M. Guidolin, and Y. Yildrim. 2014. "Markov Switching Dynamics in Reit Returns: Univariate and Multivariate Evidence on Forecasting Performance." *Real Estate Economics* 42: 279–342.

Conrad, C. and M. Karanasos. 2010. "Negative Volatility Spillovers in the Unrestricted ECCC-GARCH Model." *Econometric Theory* 26: 838–862.

Conrad, C., and M. Karanasos. 2015. "Modeling the Link between us Inflation and Output: The Importance of the Uncertainty Channel." *Scottish Journal of Political Economy* 62: 431–453.

Dean, W. G., and R. W. Faff. 2008. "Evidence of Feedback Trading with Markov Switching Regimes." *Review of Quantitative Finance and Accounting* 30: 133–151.

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

Dueker, M. J. 1997. "Markov Switching in GARCH Processes and Mean-Reverting Stock-Market Volatility." *Journal of Business and Economic Statistics* 15: 26–34.

Engle, R. F. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of united kingdom Inflation." *Econometrica* 50: 987–1008.

Engle, R. F. 2002. "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models." *Journal of Business and Economic Statistics* 20: 339–350.

Engle, R. F., and K. F. Kroner. 1995. "Multivariate Simultaneous Generalized ARCH." *Econometric Theory* 11: 122–150.

Francq, C., and J.-M. Zakoïan. 2005. "The l^2–Structures of Standard and Switching-Regime GARCH Models." *Stochastic Processes and their Applications* 115: 1557–1582.

Francq, C., and J.-M. Zakoïan. 2012. "Qml Estimation of a Class of Multivariate Asymmetric GARCH Models." *Econometric Theory* 28: 179–206.

Francq, C., M. Roussignol, and J.-M. Zakoïan. 2001. "Conditional Heteroskedasticity Driven by Hidden Markov Chains." *Journal of Time Series Analysis* 22: 197–220.

Giot, P., and S. Laurent. 2003. "Value-at-Risk for Long and Short Trading Positions." *Journal of Applied Econometrics* 18: 641–664.

Glosten, L. R., R. Jagannathan, and D. E. Runkle. 1993. "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *Journal of Finance* 48: 1779–1801.

Gray, S. F. 1996. "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process." *Journal of Financial Economics* 42: 27–62.

Gribisch, B. 2016. "Multivariate Wishart Stochastic Volatility and Changes in Regime." *AStA - Advances in Statistical Analysis* 100: 443–473.

Guidolin, M. 2011. "Markov Switching Models in Empirical Finance." In *Missing Data Methods: Time-Series Methods and Applications (Advances in Econometrics, Volume 27)*, edited by D. M. Drukker, 1–86. Bingley: Emerald.

Haas, M. 2010. "Covariance Forecasts and Long-Run Correlations in a Markov-Switching Model for Dynamic Correlations." *Finance Research Letters* 7: 86–97.

Haas, M., and J.-C. Liu. 2014. "Theory for a Multivariate Markov-Switching GARCH Model with an Application to Stock Markets." QBER Discussion Paper 7/2014, University of Kiel.

Haas, M., S. Mittnik, and M. S. Paolella. 2004a. "Mixed Normal Conditional Heteroskedasticity." *Journal of Financial Econometrics* 2: 211–250.

Haas, M., S. Mittnik, and M. S. Paolella. 2004b. "A New Approach to Markov-Switching GARCH Models." *Journal of Financial Econometrics* 2: 493–530.

Haas, M., S. Mittnik, and M. S. Paolella. 2009. "Asymmetric Multivariate Normal Mixture GARCH." *Computational Statistics and Data Analysis* 53: 2129–2154.

Haas, M., and M. S. Paolella. 2012. "Mixture and Regime-Switching GARCH Models." In *Handbook of Volatility Models and their Applications*, edited by L. Bauwens, C. M. Hafner, and S. Laurent. John Wiley & Sons.

Hamilton, J. D. 1994. *Time Series Analysis*. Princeton, New Jersey: Princeton University Press.

Hamilton, J. D. 1996. "Specification Testing in Markov-Switching Time-Series Models." *Journal of Econometrics* 70: 127–157.

He, C., and T. Teräsvirta. 2004. "An Extended Constant Correlation GARCH Model and its Fourth-Moment Structure." *Econometric Theory* 20: 904–926.

Heaney, R., and S. Sriananthakumar. 2012. "Time-Varying Correlation between Stock Market Returns and Real Estate Returns." *Journal of Empirical Finance* 19: 583–594.

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

Jondeau, E., and M. Rockinger. 2006. "The Copula-GARCH Model of Conditional Dependencies: An International Stock Market Application." *Journal of International Money and Finance* 25: 827–853.

Kasch, M., and M. Caporin. 2013. "Volatility Threshold Dynamic Conditional Correlations: An International Analysis." *Journal of Financial Econometrics* 11: 706–742.

Keribin, C. 2000. "Consistent Estimation of the Order of Mixture Models." *Sankhyā: The Indian Journal of Statistics, Series A* 62: 49–66.

Kim, C.-J., and C. R. Nelson. 1999. *State-Space Models with Regime Switching*. Cambridge, Massachusetts: MIT Press.

Klaassen, F. 2002. "Improving GARCH Volatility Forecasts with Regime-Switching GARCH." *Empirical Economics* 27: 363–394.

Ledoit, O., P. Santa-Clara, and M. Wolf. 2003. "Flexible Multivariate GARCH Modeling with an Application to International Stock Markets." *Review of Economics and Statistics* 85: 735–747.

Lee, Y.-H. 2014. "An International Analysis of REITs and Stock Portfolio Management based on Dynamic Conditional Correlation Models." *Financial Markets and Portfolio Management* 28: 165–180.

Lejeune, B. 2009. "A Diagnostic m-test for Distributional Specification of Parametric Conditional Heteroscedasticity Models for Financial Data." *Journal of Empirical Finance* 16: 507–523.

Liu, J.-C. 2006. "Stationarity of a Markov-Switching GARCH Model." *Journal of Financial Econometrics* 4: 573–593.

Liu, J.-C. 2007. "Stationarity for a Markov-Switching Box-Cox Transformed Threshold GARCH Process." *Statistics and Probability Letters* 77: 1428–1438.

Magnus, J. R. 1988. *Linear Structures*. London: Griffin.

Manner, H., and O. Reznikova. 2012. "A Survey on Time-Varying Copulas: Specification, Simulations, and Application." *Econometric Reviews* 31: 654–687.

McAleer, M., S. Hoti, and F. Chan. 2009. "Structure and Asymptotic Theory for Multivariate Asymmetric Conditional Volatility." *Econometric Reviews* 28: 422–440.

Mittnik, S. 2014. "Var-Implied Tail-Correlation Matrices." *Economics Letters* 122: 69–73.

Morawski, J., H. Rehkugler, and R. Füss. 2008. "The Nature of Listed Real Estate Companies: Property or Equity Market?" *Financial Markets and Portfolio Management* 22: 101–126.

Nabeya, S. 1951. "Absolute Moments in 2-Dimensional Normal Distribution." *Annals of the Institute of Statistical Mathematics* 3: 2–6.

Nakatani, T., and T. Teräsvirta. 2008. "Positivity Constraints on the Conditional Variances in the Family of Conditional Correlation GARCH Models." *Finance Research Letters* 5: 88–95.

Nakatani, T., and T. Teräsvirta. 2009. "Testing for Volatility Interactions in the Constant Conditional Correlation GARCH Model." *Econometrics Journal* 12: 147–163.

Otranto, E. 2010. "Asset Allocation using Flexible Dynamic Correlation Models with Regime Switching." *Quantitative Finance* 10: 325–338.

Pelletier, D. 2006. "Regime Switching for Dynamic Correlations." *Journal of Econometrics* 131: 445–473.

Putintseva, M. 2012. "Mixture Normal Conditional Correlation Models." Research Paper No. 12-41, Swiss Finance Institute.

Ramchand, L., and R. Susmel. 1998. "Volatility and Cross Correlation Across Major Stock Markets." *Journal of Empirical Finance* 5: 397–416.

Reher, G., and B. Wilfling. 2016. "A Nesting Framework for Markov-Switching GARCH Modelling with an Application to the German Stock Market." *Quantitative Finance* 16: 411–426.

Ryden, T., T. Teräsvirta, and S. Åsbrink. 1998. "Stylized Facts of Daily Return Series and the Hidden Markov Model." *Journal of Applied Econometrics* 13: 217–244.

Schwarz, G. 1978. "Estimating the Dimension of a Model." *Annals of Statistics* 6: 461–464.

Shi, Y., and L. Feng. 2016. "A Discussion on the Innovation Distribution of the Markov Regime-Switching GARCH Model." *Economic Modelling* 53: 278–288.

Silvennoinen, A., and T. Teräsvirta. 2009a. "Modeling Multivariate Autoregressive Conditional Heteroskedasticity with the Double Smooth Transition Conditional Correlation GARCH Model." *Journal of Financial Econometrics* 7: 373–411.

Silvennoinen, A., and T. Teräsvirta. 2009b. "Multivariate GARCH Models." In *Handbook of Financial Time Series*, edited by T. Mikosch, J.-P. Kreiß, R. A. Davis, and T. G. Andersen, 201–229. Berlin: Springer.

Smith, D. R. 2008. "Evaluating Specification Tests for Markov-Switching Time-Series Models." *Journal of Time Series Analysis* 29: 629–652.

Taylor, S. J. 1986. *Modelling Financial Time Series*. Chichester: John Wiley & Sons.

Timmermann, A. 2000. "Moments of Markov Switching Models." *Journal of Econometrics* 96: 75–111.

Tse, Y. K. 2000. "A Test for Constant Correlations in a Multivariate GARCH Model." *Journal of Econometrics* 98: 107–127.