

CENTRE FOR ECONOMETRIC ANALYSIS
CEA@Cass



<http://www.cass.city.ac.uk/cea/index.html>

Cass Business School
Faculty of Finance
106 Bunhill Row
London EC1Y 8TZ

***Conditional Correlation Models of Autoregressive Conditional
Heteroskedasticity with Nonstationary GARCH Equations***

Cristina Amado and Timo Teräsvirta

CEA@Cass Working Paper Series

WP-CEA-01-2013

Conditional Correlation Models of Autoregressive Conditional Heteroskedasticity with Nonstationary GARCH Equations

Cristina Amado*

*CREATES, School of Economics and Management, Aarhus University
Building 1322, DK-8000 Aarhus, Denmark*

and

*University of Minho and NIPE
Campus de Gualtar, 4710-057 Braga, Portugal*

Timo Teräsvirta†

*CREATES, School of Economics and Management, Aarhus University
Building 1322, DK-8000 Aarhus, Denmark*

June 2013

*Corresponding author. E-mail: camado@eeg.uminho.pt, Tel.: +351 253 601 383, Fax: +351 253 676 375.

†E-mail: tterasvirta@econ.au.dk, Tel.: +45 894 219 73.

Abstract

In this paper we investigate the effects of careful modelling the long-run dynamics of the volatilities of stock market returns on the conditional correlation structure. To this end we allow the individual unconditional variances in Conditional Correlation GARCH models to change smoothly over time by incorporating a nonstationary component in the variance equations as the Spline-GARCH model in Engle and Rangel (2008) and the TV-GARCH model in Amado and Teräsvirta (2013). The variance equations combine the long-run and the short-run dynamic behaviour of the volatilities. The structure of the conditional correlation matrix is assumed to be either time independent or to vary over time. We apply our model to pairs of seven daily stock returns belonging to the S&P 500 composite index and traded at the New York Stock Exchange. The results suggest that accounting for deterministic changes in the unconditional variances improves the fit of the multivariate Conditional Correlation GARCH models to the data. The effect of careful specification of the variance equations on the estimated correlations is variable: in some cases rather small, in others more discernible. We also show empirically that the Conditional Correlation GARCH models with time-varying unconditional variances using the TV-GARCH model outperform the other models under study in terms of out-of-sample forecasting performance. In addition, we find that portfolio volatility-timing strategies based on time-varying unconditional variances often outperforms the unmodelled long-run variances strategy in the out-of-sample. As a by-product, we generalize news impact surfaces to the situation in which both the GARCH equations and the conditional correlations contain a deterministic component that is a function of time.

JEL classification: C12; C32; C51; C52.

Key words: Multivariate GARCH model; Conditional correlations; Time-varying unconditional variance; Nonlinear time series; Forecasting; Portfolio allocation.

1 Introduction

Many financial issues, such as hedging and risk management, portfolio selection and asset allocation rely on information about the covariances or correlations between the underlying returns. This has motivated the modelling of volatility using multivariate financial time series rather than modelling individual returns separately. A number of multivariate generalized autoregressive conditional heteroskedasticity (GARCH) models have been proposed, and some of them have become standard tools for financial analysts. For recent surveys of multivariate GARCH models see Bauwens, Laurent and Rombouts (2006) and Silvennoinen and Teräsvirta (2009).

In the univariate setting, volatility models have been extensively investigated. Many modelling proposals of univariate financial returns have suggested that nonstationarities in return series may be the cause of the extreme persistence of shocks in estimated GARCH models. In particular, Mikosch and Stărică (2004) showed how the long-range dependence and the ‘integrated GARCH effect’ can be explained by unaccounted structural breaks in the unconditional variance. Previously, Diebold (1986) and Lamoureux and Lastrapes (1990) also argued that spurious long memory may be detected from a time series with structural breaks.

The problem of structural breaks in the conditional variance can be dealt with by assuming that the ARCH or GARCH model is piecewise stationary and detecting the breaks; see for example Berkes, Gombay, Horváth and Kokoszka (2004), or Lavielle and Teyssière (2006) for the multivariate case. It is also possible to assume, as Dahlhaus and Subba Rao (2006) recently did, that the parameters of the model change smoothly over time such that the conditional variance is locally but not globally stationary. These authors proposed a locally time-varying ARCH process for modelling the nonstationarity in variance. van Bellegem and von Sachs (2004), Engle and Rangel (2008) and, independently, Amado and Teräsvirta (2008) (for later versions, see Amado and Teräsvirta, 2012, 2013) assumed global nonstationarity and, among other things, developed an approach in which volatility is modelled by a multiplicative decomposition of the variance to a nonstationary and stationary component. The stationary component is modelled as a GARCH process, whereas the nonstationary one is a deterministic time-varying component. In van Bellegem and von Sachs (2004) this component is estimated nonparametrically using kernel estimation, in Engle and Rangel (2008) it is an exponential spline, whereas Mazur and Pipień (2012) use the

Fourier Flexible Form of Gallant (1981). Amado and Teräsvirta (2012, 2013) described the non-stationary component by a linear combination of logistic functions of time and their generalisations and developed a data-driven specification technique for determining the parametric structure of the time-varying component. The parameters of both the unconditional and the conditional component were estimated jointly by maximum likelihood. Asymptotic properties of these estimators were considered by Amado and Teräsvirta (2013).

Despite the growing literature on multivariate GARCH models, little attention has been devoted to modelling multivariate financial data by explicitly allowing for nonstationarity in variance. Recently, Hafner and Linton (2010) proposed what they called a semiparametric generalisation of the scalar multiplicative model of Engle and Rangel (2008). Their multivariate GARCH model is a first-order BEKK-GARCH model with a deterministic nonstationary or 'long run' component. In fact, their model is closer in spirit to that of van Bellegem and von Sachs (2004), because they estimate the nonstationary component nonparametrically. The authors suggested an estimation procedure for the parametric and nonparametric components and established semiparametric efficiency of their estimators.

In this paper we consider a parametric extension of the univariate multiplicative GARCH model of Amado and Teräsvirta (2012, 2013) to the multivariate case. We investigate the effects of careful modelling of the time-varying unconditional variance on the correlation structure of Conditional Correlation GARCH (CC-GARCH) models. To this end, we allow the individual unconditional variances in the multivariate GARCH models to change smoothly over time by incorporating a nonstationary component in the variance equations. The empirical analysis consists of first fitting bivariate conditional correlation GARCH models to pairs of daily return series and comparing the results from models with the time-varying unconditional variance component to models without such a component. Thereafter, we carry out an out-of-sample analysis to evaluate the forecasting performance for the conditional covariances matrices of all individual return series. We also assess the economic value of the time-varying unconditional variance based on CC-GARCH models. For this purpose, we implement volatility-timing strategies using both the unmodelled and the modelled time-varying unconditional variance components and evaluate the economic gains in portfolio allocation in the out-of-sample period associated with switching to the model with time-varying unconditional variance. By comparing covariance forecasts in the

portfolio selection framework we find that multivariate covariance forecasts based on time-varying unconditional variances are favoured over the ones obtained from CC-GARCH models with a constant unconditional variance.

As a by-product, we extend the concept of news impact surfaces of Kroner and Ng (1998) to the case where both the variances and conditional correlations are fluctuating deterministically over time. These surfaces illustrate how the impact of news on covariances between asset returns depends both on the state of the market and the time-varying dependence between the returns.

The paper is organised as follows. In Section 2 we describe the Conditional Correlation GARCH model in which the individual unconditional variances change smoothly over time. Estimation of parameters of these models is discussed in Section 3 and specification of the unconditional variance components in Section 4. Section 5 contains the empirical results of fitting bivariate CC-GARCH models to the 21 pairs of seven daily return series of stocks belonging to the S&P 500 composite index and results of an out-of-sample forecasting experiment. Section 6 comprises an out-of-sample evaluation of the economic value of modelling the time-varying unconditional variances. Generalisations of news impact surfaces are presented in Section 7. Conclusions can be found in Section 8.

2 The model

2.1 The general framework

Consider a $N \times 1$ vector of return time series $\{\mathbf{y}_t\}$, $t = 1, \dots, T$, described by the following vector process:

$$\mathbf{y}_t = \mathbf{E}(\mathbf{y}_t | \mathcal{F}_{t-1}) + \boldsymbol{\varepsilon}_t \quad (1)$$

where \mathcal{F}_{t-1} is the sigma-algebra generated by the available information up until $t-1$. For simplicity, we assume $\mathbf{E}(\mathbf{y}_t | \mathcal{F}_{t-1}) = \mathbf{0}$. The N -dimensional vector of innovations (or now, returns) $\{\boldsymbol{\varepsilon}_t\}$ is defined as

$$\boldsymbol{\varepsilon}_t = \mathbf{D}_t \boldsymbol{\zeta}_t \quad (2)$$

where \mathbf{D}_t is a diagonal matrix of time-varying standard deviations. The error vectors $\boldsymbol{\zeta}_t$ form a sequence of independent and identically distributed variables with mean zero and a positive

definite correlation matrix $\mathbf{P}_t = [\rho_{ijt}]$ such that $\rho_{iit} = 1$ and $|\rho_{ijt}| < 1$, $i \neq j$, $i, j = 1, \dots, N$. This implies $\mathbf{P}_t^{-1/2} \boldsymbol{\zeta}_t \sim iid(\mathbf{0}, \mathbf{I}_N)$. Under these assumptions, the error vector $\boldsymbol{\varepsilon}_t$ satisfies the following moment conditions:

$$\begin{aligned} \mathbf{E}(\boldsymbol{\varepsilon}_t | \mathcal{F}_{t-1}) &= \mathbf{0} \\ \mathbf{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' | \mathcal{F}_{t-1}) &= \boldsymbol{\Sigma}_t = \mathbf{D}_t \mathbf{P}_t \mathbf{D}_t' \end{aligned} \quad (3)$$

where the conditional covariance matrix $\boldsymbol{\Sigma}_t = [\sigma_{ijt}]$ of $\boldsymbol{\varepsilon}_t$ given the information set \mathcal{F}_{t-1} is a positive-definite $N \times N$ matrix. It is now assumed that \mathbf{D}_t consists of a conditionally heteroskedastic component and a deterministic time-dependent one such that

$$\mathbf{D}_t = \mathbf{S}_t \mathbf{G}_t \quad (4)$$

where $\mathbf{S}_t = \text{diag}(h_{1t}^{1/2}, \dots, h_{Nt}^{1/2})$ contains the conditional standard deviations $h_{it}^{1/2}$, $i = 1, \dots, N$, and $\mathbf{G}_t = \text{diag}(g_{1t}^{1/2}, \dots, g_{Nt}^{1/2})$. The elements g_{it} , $i = 1, \dots, N$, are positive-valued deterministic functions of rescaled time, whose structure will be defined in a moment. Equations (3) and (4) jointly define the time-varying covariance matrix

$$\boldsymbol{\Sigma}_t = \mathbf{S}_t \mathbf{G}_t \mathbf{P}_t \mathbf{G}_t \mathbf{S}_t. \quad (5)$$

It follows that

$$\sigma_{ijt} = \rho_{ijt} (h_{it} g_{it})^{1/2} (h_{jt} g_{jt})^{1/2}, \quad i \neq j \quad (6)$$

and that

$$\sigma_{iit} = h_{it} g_{it}, \quad i = 1, \dots, N. \quad (7)$$

From (7) it follows that $h_{it} = \sigma_{iit}/g_{it} = \mathbf{E}(\varepsilon_{it}^* \varepsilon_{it}^* | \mathcal{F}_{t-1})$, where $\varepsilon_{it}^* = \varepsilon_{it}/g_{it}^{1/2}$. When $\mathbf{G}_t \equiv \mathbf{I}_N$ and the conditional correlation matrix $\mathbf{P}_t \equiv \mathbf{P}$, one obtains the Constant Conditional Correlation (CCC-) GARCH model of Bollerslev (1990). More generally, when $\mathbf{G}_t \equiv \mathbf{I}_N$ and \mathbf{P}_t is a time-varying correlation matrix, the model belongs to the family of Conditional Correlation GARCH models.

Following Amado and Teräsvirta (2012, 2013), the diagonal elements of the matrix \mathbf{G}_t are

defined as follows:

$$g_{it} = \delta_{0l} + \sum_{l=1}^{r_i} \delta_{il} G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il}) \quad (8)$$

where $\gamma_{il} > 0$, $i = 1, \dots, N$, $l = 1, \dots, r_i$, and $r_i = 0, 1, 2, \dots, R$, such that R is a finite integer. The identification problem arising from the fact that both h_{it} and g_{it} contain an intercept is solved by setting $\delta_{0l} = 1$. Each g_{it} varies smoothly over time satisfying the conditions $\inf_{t=1, \dots, T} g_{it} > 0$, and $\delta_{il} \leq M_\delta < \infty$, $l = 1, \dots, r$, for $i = 1, \dots, N$. For identification reasons, in (8) $\delta_{i1} < \dots < \delta_{ir}$ and $\delta_{il} \neq 0$ for all i and l . The function $G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il})$ is a generalized logistic function, that is,

$$G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il}) = \left(1 + \exp \left\{ -\gamma_{il} \prod_{j=1}^{k_{il}} (t/T - c_{ilj}) \right\} \right)^{-1}, \quad \gamma_{il} > 0, \quad c_{i11} \leq \dots \leq c_{ilk}. \quad (9)$$

Function (9) is by construction continuous for $\gamma_{il} < \infty$, $i = 1, \dots, r$, and bounded between zero and one. The parameters, c_{ilj} and γ_{il} determine the location and the speed of the transition between regimes.

The parametric form of (8) with (9) is very flexible and capable of describing smooth changes in the amplitude of volatility clusters. Under $\delta_{i1} = \dots = \delta_{ir} = 0$ or $\gamma_{i1} = \dots = \gamma_{ir} = 0$, $i = 1, \dots, N$, in (8), the unconditional variance of ε_t becomes constant, otherwise it is time-varying. Assuming either $r_i > 1$ or $k_{il} > 1$ or both with $\delta_{il} \neq 0$ adds flexibility to the unconditional variance component g_{it} . In the simplest case, $r = 1$ and $k = 1$, g_{it} increases monotonically over time when $\delta_{i1} > 0$ and decreases monotonically when $\delta_{i1} < 0$. The slope parameter γ_{i1} in (9) controls the degree of smoothness of the transition: the larger γ_{i1} , the faster the transition between the extreme regimes. As $\gamma_{i1} \rightarrow \infty$, g_{it} approaches a step function with a switch at c_{i11} . For small values of γ_{i1} , the transition between regimes is very smooth.

In this work we shall account for potentially asymmetric responses of volatility to positive and negative shocks or returns by assuming the conditional variance components to follow the GJR-GARCH process of Glosten, Jagannathan and Runkle (1993). In the present context,

$$h_{it} = \omega_i + \sum_{j=1}^q \alpha_{ij} \varepsilon_{i,t-j}^{*2} + \sum_{j=1}^q \kappa_{ij} I(\varepsilon_{i,t-j}^* < 0) \varepsilon_{i,t-j}^{*2} + \sum_{j=1}^p \beta_{ij} h_{i,t-j}, \quad (10)$$

where the indicator function $I(A) = 1$ when A is valid, otherwise $I(A) = 0$. The assumption of a

discrete switch at $\varepsilon_{i,t-j}^* = 0$ can be generalised following Hagerud (1997), but this extension is left for later work.

2.2 The structure of the (un)conditional correlations

The purpose of this work is to investigate the effects of modelling changes in the unconditional variances on conditional correlation estimates. The idea is to compare the standard approach, in which the nonstationary component is left unmodelled, with the one relying on the decomposition (5) with $\mathbf{G}_t \neq \mathbf{I}_N$. As to modelling the time-variation in the correlation matrix \mathbf{P}_t , several choices exist. As already mentioned, the simplest multivariate correlation model is the CCC-GARCH model of Bollerslev (1990) in which $\mathbf{P}_t \equiv \mathbf{P}$. With h_{it} specified as in (10), this model will be called the CCC-TVGJR-GARCH model. When $g_{it} \equiv 1$, (10) defines the i th conditional variance of the CCC-GJR-GARCH model.

The CCC-GARCH model has considerable appeal due to its computational simplicity, but in many studies the assumption of constant correlations has been found to be too restrictive. There are several ways of relaxing this assumption using parametric representations for the correlations. Engle (2002) introduced the so-called Dynamic CC-GARCH (DCC-GARCH) model in which the conditional correlations are defined through GARCH(1,1) type equations. Tse and Tsui (2002) presented a rather similar model. In the DCC-GARCH model, the coefficient of correlation ρ_{ijt} is a typical element of the matrix \mathbf{P}_t with the dynamic structure

$$\mathbf{P}_t = \{\text{diag}\mathbf{Q}_t\}^{-1/2}\mathbf{Q}_t\{\text{diag}\mathbf{Q}_t\}^{-1/2} \quad (11)$$

where

$$\mathbf{Q}_t = (1 - \theta_1 - \theta_2)\overline{\mathbf{Q}} + \theta_1\boldsymbol{\zeta}_{t-1}\boldsymbol{\zeta}'_{t-1} + \theta_2\mathbf{Q}_{t-1} \quad (12)$$

such that $\theta_1 > 0$ and $\theta_2 \geq 0$ with $\theta_1 + \theta_2 < 1$, $\overline{\mathbf{Q}}$ is the unconditional correlation matrix of the standardised errors ζ_{it} , $i = 1, \dots, N$, and $\boldsymbol{\zeta}_t = (\zeta_{1t}, \dots, \zeta_{Nt})'$. In our case, each $\zeta_{it} = \varepsilon_{it}/(h_{it}g_{it})^{1/2}$, and this version of the model will be called the DCC-TVGJR-GARCH model. Accordingly, when $g_{it} \equiv 1$, the model becomes the DCC-GJR-GARCH model. In the Varying Correlation (VC-) GARCH model of Tse and Tsui (2002), \mathbf{P}_t has a definition that is slightly different from (12).

More specifically,

$$\mathbf{P}_t = (1 - \theta_1 - \theta_2)\mathbf{P} + \theta_1\boldsymbol{\Psi}_{t-1} + \theta_2\mathbf{P}_{t-1} \quad (13)$$

where \mathbf{P} is a constant positive definite parameter matrix with unit diagonal elements, θ_1 and θ_2 are non-negative parameters such that $\theta_1 + \theta_2 < 1$, and $\boldsymbol{\Psi}_{t-1}$ is the sample correlation matrix of $\{\boldsymbol{\zeta}_{t-1}, \dots, \boldsymbol{\zeta}_{t-M}\}$, $M \geq N$. The positive definiteness of \mathbf{P}_t is ensured if \mathbf{P}_0 and $\boldsymbol{\Psi}_{t-1}$ are positive definite matrices. In our application, when ζ_{it} is specified as $\zeta_{it} = \varepsilon_{it}/(h_{it}g_{it})^{1/2}$, the model will be called VC-TVGJR-GARCH model. When $g_{it} \equiv 1$, the model becomes the VC-GJR-GARCH model.

2.3 Multi-step ahead forecasting

Constructing one-step-ahead covariance forecasts for the CC-TVGJR-GARCH models discussed in the previous section is straightforward. Since the conditional standard deviations for the next period are known, we have

$$\mathbf{E}_t\boldsymbol{\Sigma}_{t+1} = \mathbf{S}_{t+1|t}\mathbf{G}_{t+1|t}\mathbf{P}_{t+1|t}\mathbf{G}_{t+1|t}\mathbf{S}_{t+1|t}.$$

where $\mathbf{S}_{t+1|t}$ is the diagonal matrix holding the one-step ahead conditional variance forecasts as described in Section (2.1). The low frequency volatility forecasts included in the matrix $\mathbf{G}_{t+1|t}$ are constructed under the assumption that $g_{i,t+1|t} = g_{i,t}$, for all $i = 1, \dots, N$. The d -steps-ahead conditional expectations of the covariance matrix do not have a closed form, and we construct the d -steps-ahead correlation forecasts as in Engle and Sheppard (2001). In the DCC-TVGJR-GARCH model, the correlation forecast $\mathbf{P}_{t+1|t}$ is the standardized version of $\mathbf{Q}_{t+1|t}$, whose one-step-ahead forecast is obtained by projecting (12) one step into the future. In this scheme, the standardized returns are $\mathbf{E}_t\tilde{\zeta}_{i,t+1} = \varepsilon_{i,t+1|t}/(h_{i,t+1|t}g_{i,t+1|t})^{1/2}$, and

$$\mathbf{Q}_{t+r|t} = (1 - \theta_1 - \theta_2)\overline{\mathbf{Q}} + (\theta_1 + \theta_2)\mathbf{Q}_{t+r-1|t}$$

for $r > 1$. We construct one-step-ahead forecasts for the VC-TVGJR-GARCH model in an identical fashion.

3 Estimation of parameters

The conditional log-likelihood function for observation t of the DCC-TVGJR-GARCH model assuming $\boldsymbol{\varepsilon}_t | \mathcal{F}_{t-1} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_t)$ is defined as

$$\begin{aligned}
\ell_t(\boldsymbol{\psi}, \boldsymbol{\varphi}, \boldsymbol{\phi}) &= -(N/2) \ln(2\pi) - (1/2) \ln |\boldsymbol{\Sigma}_t| - (1/2) \boldsymbol{\varepsilon}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\varepsilon}_t \\
&= -(N/2) \ln(2\pi) - (1/2) \ln |\mathbf{S}_t \mathbf{G}_t \mathbf{P}_t \mathbf{G}_t \mathbf{S}_t| - (1/2) \boldsymbol{\varepsilon}_t' \mathbf{S}_t^{-1} \mathbf{G}_t^{-1} \mathbf{P}_t^{-1} \mathbf{G}_t^{-1} \mathbf{S}_t^{-1} \boldsymbol{\varepsilon}_t \\
&= -(N/2) \ln(2\pi) - \ln |\mathbf{S}_t \mathbf{G}_t| - (1/2) \ln |\mathbf{P}_t| - (1/2) \boldsymbol{\zeta}_t' \mathbf{P}_t^{-1} \boldsymbol{\zeta}_t \\
&= -(N/2) \ln(2\pi) - \ln |\mathbf{G}_t| - (1/2) \tilde{\boldsymbol{\varepsilon}}_t' \mathbf{G}_t^{-2} \tilde{\boldsymbol{\varepsilon}}_t - \ln |\mathbf{S}_t| - (1/2) \boldsymbol{\varepsilon}_t^* \mathbf{S}_t^{-2} \boldsymbol{\varepsilon}_t^* \\
&\quad + \boldsymbol{\zeta}_t' \boldsymbol{\zeta}_t - (1/2) \ln |\mathbf{P}_t| - (1/2) \boldsymbol{\zeta}_t' \mathbf{P}_t^{-1} \boldsymbol{\zeta}_t
\end{aligned} \tag{14}$$

where

$$\begin{aligned}
\tilde{\boldsymbol{\varepsilon}}_t &= \mathbf{S}_t^{-1} \boldsymbol{\varepsilon}_t = (\varepsilon_{1t} / \{h_{1t}(\boldsymbol{\psi}_1, \boldsymbol{\varphi}_1)\}^{1/2}, \dots, \varepsilon_{Nt} / \{h_{Nt}(\boldsymbol{\psi}_N, \boldsymbol{\varphi}_N)\}^{1/2})' \\
\boldsymbol{\varepsilon}_t^* &= \mathbf{G}_t^{-1} \boldsymbol{\varepsilon}_t = (\varepsilon_{1t} / \{g_{1t}(\boldsymbol{\psi}_1)\}^{1/2}, \dots, \varepsilon_{Nt} / \{g_{Nt}(\boldsymbol{\psi}_N)\}^{1/2})' \\
\boldsymbol{\zeta}_t &= \mathbf{G}_t^{-1} \mathbf{S}_t^{-1} \boldsymbol{\varepsilon}_t = (\varepsilon_{1t} / \{g_{1t}(\boldsymbol{\psi}_1) h_{1t}(\boldsymbol{\psi}_1, \boldsymbol{\varphi}_1)\}^{1/2}, \dots, \varepsilon_{Nt} / \{g_{Nt}(\boldsymbol{\psi}_N) h_{Nt}(\boldsymbol{\psi}_N, \boldsymbol{\varphi}_N)\}^{1/2})'.
\end{aligned}$$

The N -block vector $\boldsymbol{\psi} = (\boldsymbol{\psi}'_1, \dots, \boldsymbol{\psi}'_N)'$ contains the parameters of the unconditional variance components, $\boldsymbol{\varphi} = (\boldsymbol{\varphi}'_1, \dots, \boldsymbol{\varphi}'_N)'$ the parameters of the conditional variances, and the elements of $\boldsymbol{\phi} = (\theta_1, \theta_2)'$ are the parameters of the correlation matrix. Furthermore, $\boldsymbol{\psi}_i = (\delta_{i0}, \boldsymbol{\delta}'_i, \boldsymbol{\gamma}'_i, \mathbf{c}'_i)'$, $\boldsymbol{\delta}_i = (\delta_{i1}, \dots, \delta_{ir})'$, $\boldsymbol{\gamma}_i = (\gamma_{i1}, \dots, \gamma_{ir})'$, $\mathbf{c}_i = (\mathbf{c}'_{i1}, \dots, \mathbf{c}'_{ir})'$, and $\boldsymbol{\varphi}_i = (\alpha_{i1}, \dots, \alpha_{iq}, \kappa_{i1}, \dots, \kappa_{iq}, \beta_{i1}, \dots, \beta_{ip})'$, $i = 1, \dots, N$. Details of estimation of the DCC model can be found in the Appendix A.

4 Specifying the unconditional variance component

In fitting a model belonging to the family of CC-TVGJR-GARCH models to the data, there are two specification problems. First, one has to determine p and q in (10) and r in (8). Furthermore, if $r \geq 1$, one also has to determine k for each transition function (9). Second, at least in principle one has to test the null hypothesis of constant conditional correlations against either the DCC- or VC-GARCH model. It appears, however, that in applications involving DCC-GARCH models the null hypothesis of constant conditional correlations is never tested, and we shall adhere to that

practice.

We shall thus concentrate on the first set of specification issues. We choose $p = q = 1$ and test for higher orders at the evaluation stage. As to selecting r and k , we follow Amado and Teräsvirta (2012) and briefly review their procedure. The conditional variances are estimated first, assuming $g_{it} \equiv 1$, $i = 1, \dots, N$. The number of deterministic functions g_{it} is determined thereafter equation by equation by sequential testing. For the i th equation, the first hypothesis to be tested is H_{01} : $\gamma_{i1} = 0$ against H_{11} : $\gamma_{i1} > 0$ in

$$g_{it} = 1 + \delta_{i1}G_{i1}(t/T; \gamma_{i1}, \mathbf{c}_{i1}).$$

The standard test statistic has a non-standard asymptotic distribution because δ_{i1} and \mathbf{c}_{i1} are unidentified nuisance parameters when H_{01} is true. This lack of identification may be circumvented by following Luukkonen, Saikkonen and Teräsvirta (1988). This means that $G_{i1}(t/T; \gamma_{i1}, \mathbf{c}_{i1})$ is replaced by its m th-order Taylor expansion around $\gamma_{i1} = 0$. Choosing $m = 3$, this yields

$$g_{it} = \alpha_0^* + \sum_{j=1}^3 \delta_{ij}^*(t/T)^j + R_3(t/T; \gamma_{i1}, \mathbf{c}_{i1}) \quad (15)$$

where $\delta_{ij}^* = \gamma_{i1}^j \tilde{\delta}_{ij}^*$, with $\tilde{\delta}_{ij}^* \neq 0$, and $R_3(t/T; \gamma_{i1}, \mathbf{c}_{i1})$ is the remainder. The new null hypothesis based on this approximation is H'_{01} : $\delta_{i1}^* = \delta_{i2}^* = \delta_{i3}^* = 0$ in (15). In order to test this null hypothesis, we use the Lagrange multiplier (LM) test. Furthermore, $R_3(t/T; \gamma_{i1}, \mathbf{c}_{i1}) \equiv 0$ under H_{01} , so the asymptotic distribution theory is not affected by the remainder. As discussed in Amado and Teräsvirta (2012), the LM-type test statistic has an asymptotic χ^2 -distribution with three degrees of freedom when H_{01} holds.

If the null hypothesis is rejected, the model builder also faces the problem of selecting the order $k \leq 3$ in the exponent of $G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il})$. It is solved by carrying out a short test sequence within (15); for details see Amado and Teräsvirta (2012). The next step is then to estimate the alternative with the chosen k , add another transition, and test the hypothesis $\gamma_{i2} = 0$ in

$$g_{it} = 1 + \delta_{i1}^*G_{i1}(t/T; \gamma_{i1}, \mathbf{c}_{i1}) + \delta_{i2}^*G_{i1}(t/T; \gamma_{i2}, \mathbf{c}_{i2})$$

using the same technique as before. Testing continues until the first non-rejection of the null

hypothesis. The LM-type test statistic still has an asymptotic χ^2 -distribution with three degrees of freedom under the null hypothesis.

The model-building cycle for TVGJR-GARCH models for the elements of $\mathbf{D}_t = \mathbf{S}_t \mathbf{G}_t$ of the CC-GARCH model defined by equations (3) and (4) consists on specification, estimation and evaluation stages. After specifying and estimating the model, the estimated individual TVGJR-GARCH equations will be evaluated by means of LM-type diagnostic tests considered in Amado and Teräsvirta (2012).

5 Empirical analysis I: Modelling and forecasting

5.1 Data

The effects of careful modelling the nonstationarity in return series on the conditional correlations are studied with price series of seven stocks of the S&P 500 composite index traded at the New York Stock Exchange. The time series are available at the website *Yahoo! Finance*. They consist of daily closing prices of American Express (AXP), Boeing Company (BA), Caterpillar (CAT), Intel Corporation (INTC), JPMorgan Chase & Co. (JPM), Whirlpool (WHR) and Exxon Mobil Corporation (XOM). The seven companies belong to different industries that are consumer finance (AXP), aerospace and defence (BA), machines (CAT), semiconductors (INTC), banking (JPM), consumption durables (WHR) and energy (XOM). The in-sample observation period begins September 29, 1998 and ends October 7, 2008, yielding a total of 2521 observations. All stock prices are converted into continuously compounded rates of return, whose values are plotted in Figure 1. The out-of-sample period extends from October 8, 2008 to December 31, 2009, which amounts to 311 trading days. As will be discussed later, the daily error covariance matrices for this period are constructed from 5-minute returns.

A common pattern is evident in the seven return series. There is a volatile period from the beginning until the middle of the observation period and a less volatile period starting around 2003 that continues almost to the end of the sample. At the very end, it appears that volatility increases again. Moreover, as expected, all series exhibit volatility clustering, but the amplitude of the clusters varies over time.

Descriptive statistics for the individual return series can be found in Table 1. Conventional

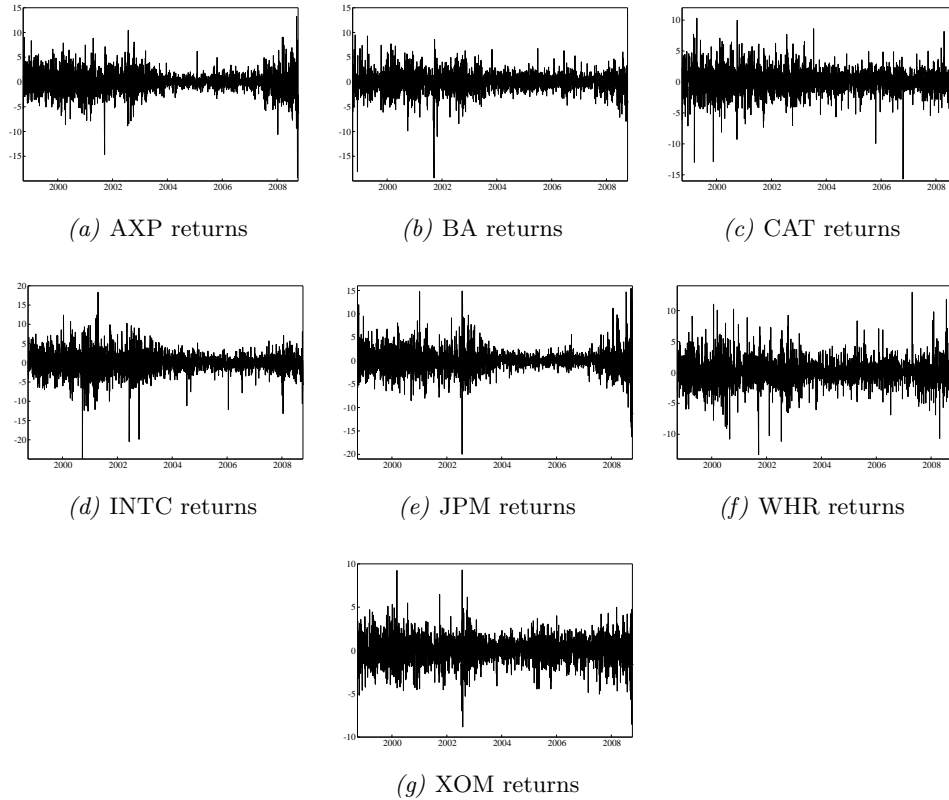


Figure 1: The seven stock returns of the S&P 500 composite index from September 29, 1998 until October 7, 2008 (2521 observations).

measures for skewness and kurtosis and also their robust counterparts are provided for all series. The conventional estimates indicate both non-zero skewness and excess kurtosis: both are typically found in financial asset returns. However, conventional measures of skewness and kurtosis are sensitive to outliers and should therefore be viewed with caution. Kim and White (2004) suggested to look at robust estimates of these quantities. The robust measures for skewness are all positive but very close to zero indicating that the return distributions show very little skewness. All robust kurtosis measures are positive, AXP and JPM being extreme examples of this, which suggests excess kurtosis (the robust kurtosis measure equals zero for normally distributed returns) but less than what the conventional measures indicate. The estimates are strictly univariate and any correlations between the series are ignored.

Table 1: Descriptive statistics of the asset returns

Asset	Min	Max	Mean	Std.dev.	Skew	Ex.Kurt	Rob.Sk.	Rob.Kr.
AXP	-19.35	13.23	0.011	2.280	-0.265	4.864	0.004	0.418
BA	-19.39	9.513	0.019	2.069	-0.611	7.215	-0.003	0.106
CAT	-15.68	10.32	0.037	2.073	-0.260	3.945	-0.022	0.108
INTC	-24.87	18.32	-0.009	2.896	-0.470	6.197	-0.001	0.160
JPM	-19.97	15.47	0.025	2.524	0.282	6.901	-0.010	0.396
WHR	-13.30	12.95	0.022	2.254	0.183	3.516	0.004	0.272
XOM	-8.83	9.29	0.039	1.579	-0.136	2.334	-0.058	0.060

Notes: The table contains summary statistics for the raw returns of the seven stocks of the S&P 500 composite index. The sample period is from September 29, 1998 until October 7, 2008 (2521 observations). Rob.Sk. denotes the robust measure for skewness based on quantiles proposed by Bowley and the Rob.Kr. is the robust centred coefficient for kurtosis proposed by Moors (see Kim and White (2004)).

5.2 Modelling the conditional variances and testing for the nonstationary component

We first construct an adequate GJR-GARCH(1,1) model for the conditional variance of each of the seven return series. The estimated models show a distinct IGARCH effect: for the AXP and JPM returns the estimate of $\alpha_{i1} + \kappa_{i1}/2 + \beta_{i1}$ even exceeds unity. In order to save space, the results are not shown here. Results of the constant unconditional variance against a time-varying structure appear in Table 2 under the heading 'single transition'. The null model is strongly rejected in all seven cases. From the same table it is seen when the single transition model is tested against two transitions ('double transition') that one transition is enough in all cases. The test sequence for selecting the type of transition shows that not all rejections imply a monotonically increasing function g_{it} .

The estimated TVGJR-GARCH models can be found in Tables 3 and 4. Table 4 shows how the persistence measure $\hat{\alpha}_{i1} + \hat{\kappa}_{i1}/2 + \hat{\beta}_{i1}$ is dramatically smaller in all cases than it is when $g_{it} \equiv 1$. In two occasions, remarkably low values, 0.782 for CAT and 0.888 for WHR, are obtained. For the remaining series the reduction in persistence is smaller but still distinct. From Table 3 it can be seen that \hat{g}_{it} changes monotonically only for BA, whereas for the other series this component first decreases and then increases again. In INTC and WHR, however, there is an increase very early on, after which the pattern is similar to that of the other four series. This is also clear from Figure 2 that contains the graphs of \hat{g}_{it} for the seven estimated models.

Figures 3 and 4 also illustrate the effects of explicitly modelling the nonstationarity in variance.

Table 2: Sequence of tests of the GJR-GARCH model against a TVGJR-GARCH model

Transitions	H ₀	H ₀₃	H ₀₂	H ₀₁
Single transition				
AXP	0.0184	0.1177	0.0071	0.5722
BA	0.0021	0.0616	0.0461	0.0072
CAT	0.0044	0.0260	0.0107	0.1971
INTC	5×10^{-5}	9×10^{-5}	0.1600	0.0197
JPM	9×10^{-4}	0.0073	0.0023	0.8500
WHR	6×10^{-5}	7×10^{-4}	0.0011	0.9401
XOM	0.0018	0.0836	7×10^{-4}	0.4271
Double transition				
AXP	0.0826	0.1953	0.0378	0.4032
BA	0.1208	0.1480	0.0547	0.8419
CAT	0.4011	0.1719	0.4961	0.4347
INTC	0.4307	0.8757	0.1050	0.7458
JPM	0.0947	0.0144	0.8678	0.5484
WHR	0.3059	0.8856	0.1450	0.2249
XOM	0.1111	0.1526	0.4198	0.0685

Notes: The entries are the p -values of the LM-type tests of constant unconditional variance applied to the seven stock returns of the S&P 500 composite index. The appropriate order k in (9) is chosen from the short sequence of hypothesis as follows: If the smallest p -value of the test corresponds to H₀₂, then choose $k = 2$. If either H₀₁ or H₀₃ are rejected more strongly than H₀₂, then select either $k = 1$ or $k = 3$. See Amado and Teräsvirta (2012) for further details.

Figure 3 shows the estimated conditional standard deviations from the GJR-GARCH models. The behaviour of these series looks rather nonstationary. The conditional standard deviations from the TVGJR-GARCH models can be found in Figure 4. These plots, in contrast to the ones in Figure 3, are rather flat and do not show signs of nonstationarity. The deterministic component g_{it} is able to absorb the changing ‘baseline volatility’, and only volatility clustering is left to be parameterized by h_{it} . This is clearly seen from the graphs in Figure 4 as they retain the peaks visible in Figure 3. This is what we would expect after the unconditional variance component has absorbed the long-run movements in the series. Results from the Spline-GARCH model of Engle and Rangel (2008) are also provided for comparison. The graphs of \hat{g}_{it} of the Spline-GARCH model for the seven return series are shown in Figure 5. The main differences between the graphs in Figures 2 and 5 appear at their ends, which may have implications for forecasting. Figure 6 shows that after the effect of the long-run volatility has been modelled, the conditional standard deviation series for the standardised returns look stationary as they do in Figure 4. Note that the scales are not comparable, as the ones in Figure 4 are based on the assumption $\delta_{0l} = 1$, which is not a unique

Table 3: Estimation results for the univariate TVGJR-GARCH models

Asset	$\hat{\delta}_1$	$\hat{\gamma}_1$	\hat{c}_{11}	\hat{c}_{12}	\hat{c}_{13}	r
<u>g_t component</u>						
AXP	4.3601 (0.1989)	300 (-)	0.4825 (0.0015)	0.9034 (0.0021)	—	1
BA	-0.651 (0.0135)	300 (-)	0.4686 (0.0011)	—	—	1
CAT	1.2366 (0.1102)	300 (-)	0.3021 (0.0011)	0.9726 (0.0028)	—	1
INTC	2.9973 (0.1553)	300 (-)	0.0262 (0.0004)	0.4775 (0.0031)	0.9127 (0.0039)	1
JPM	6.3688 (0.2737)	300 (-)	0.4821 (0.0012)	0.9042 (0.0020)	—	1
WHR	1.2272 (0.0917)	300 (-)	0.0892 (0.0008)	0.4195 (0.0072)	0.8497 (0.0024)	1
XOM	1.1063 (0.0809)	300 (-)	0.4106 (0.0029)	0.8672 (0.0037)	—	1

Notes: The table contains the parameter estimates of the g_{it} component from the TVGJR-GARCH(1,1) model for the seven stocks of the S&P 500 composite index, over the period September 29, 1998 - October 7, 2008. The estimated model has the form $g_{it} = 1 + \sum_{i=1}^r \delta_{il} G_{il}(t/T; \gamma_{il}, c_{il})$, where $G_{il}(t/T; \gamma_{il}, c_{il})$ is defined in (9) for all i . The numbers in parentheses are the standard errors.

Table 4: Estimation results for the univariate TVGJR-GARCH models

Asset	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\kappa}_1$	$\hat{\beta}_1$	$\hat{\alpha}_1 + \frac{\hat{\kappa}_1}{2} + \hat{\beta}_1$
<u>h_t component</u>					
AXP	0.0480 (0.0126)	0.0045 (0.0114)	0.1280 (0.0218)	0.9012 (0.0172)	0.9697
BA	0.2994 (0.1448)	0.0106 (0.0138)	0.0851 (0.0353)	0.9029 (0.0391)	0.9561
CAT	0.6641 (0.4611)	0.0477 (0.0290)	—	0.7340 (0.1696)	0.7817
INTC	0.1203 (0.0498)	0.0450 (0.0165)	—	0.9155 (0.0269)	0.9605
JPM	0.0474 (0.0152)	0.0213 (0.0110)	0.1135 (0.0262)	0.8890 (0.0229)	0.9670
WHR	0.3569 (0.2392)	0.0736 (0.0326)	—	0.8141 (0.1009)	0.8877
XOM	0.0644 (0.0222)	0.0272 (0.0113)	0.0578 (0.0215)	0.9008 (0.0235)	0.9568

Notes: The table contains the parameter estimates of the h_{it} component from the TVGJR-GARCH(1,1) model for the seven stocks of the S&P 500 composite index, over the period September 29, 1998 - October 7, 2008. The estimated model has the form $h_{it} = \omega_i + \alpha_{i1} \varepsilon_{it-1}^{*2} + \kappa_{i1} I_{it-1}(\varepsilon_{it-1}^*) \varepsilon_{it-1}^{*2} + \beta_{i1} h_{it-1}$, where $\varepsilon_{it}^* = \varepsilon_{it}/g_{it}^{1/2}$ and $I_{it}(\varepsilon_{it}^*) = 1$ if $\varepsilon_{it}^* < 0$ (and 0 otherwise) for all i . The numbers in parentheses are the Bollerslev-Wooldridge robust standard errors.

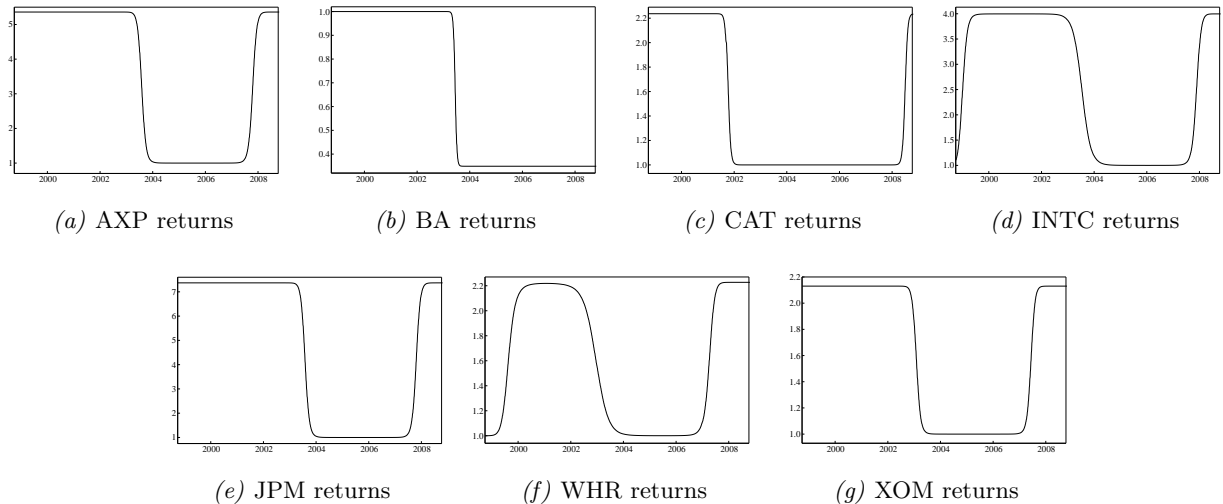


Figure 2: Estimated g_t functions for the TV-GJR-GARCH model for the seven stock returns of the S&P 500 composite index.

solution to the identification problem mentioned in Section 2.1.

5.3 Effects of modelling the long-run dynamics of volatility on correlations

We now study the effects of modelling nonstationary volatility equations on the correlations between pairs of stock returns. Since each individual return series belongs to a different industry, we first estimate bivariate Conditional Correlation GARCH models to investigate the effect on the conditional correlations at the industry level. A bivariate analysis of the returns may also give some idea of how different the correlations between firms representing different industries can be. For that purpose, we consider the Conditional Correlation GARCH models as defined in Section 2.2. For each model, two specifications will be estimated for modelling the univariate volatilities. One is the first-order GJR-GARCH model that corresponds to $\mathbf{G}_t \equiv \mathbf{I}_2$, whereas the other one is the TVGJR-GARCH model for which $\mathbf{G}_t \neq \mathbf{I}_2$ in (5).

We begin by comparing the rolling correlation estimates for the $(\varepsilon_{it}, \varepsilon_{jt})$ and $(\varepsilon_{it}/\hat{g}_{it}^{1/2}, \varepsilon_{jt}/\hat{g}_{jt}^{1/2})$ pairs, where \hat{g}_{jt} is given in (8). Figure 7 contains the pairwise correlations between the former and the latter computed over 100 trading days. This window size represents a compromise between randomness and smoothness in the correlation sequences. The differences are sometimes quite remarkable in the first half of the series where the correlations of rescaled returns are often smaller than those of the original returns. In a few cases this is true for the whole series. This might

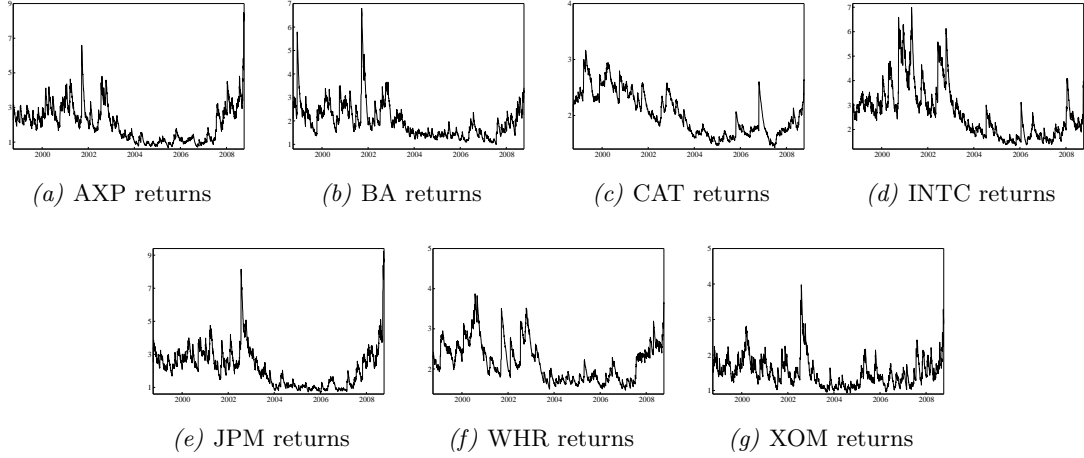


Figure 3: Estimated conditional standard deviations from the GJR-GARCH(1,1) model for the seven stock returns of the S&P 500 composite index.

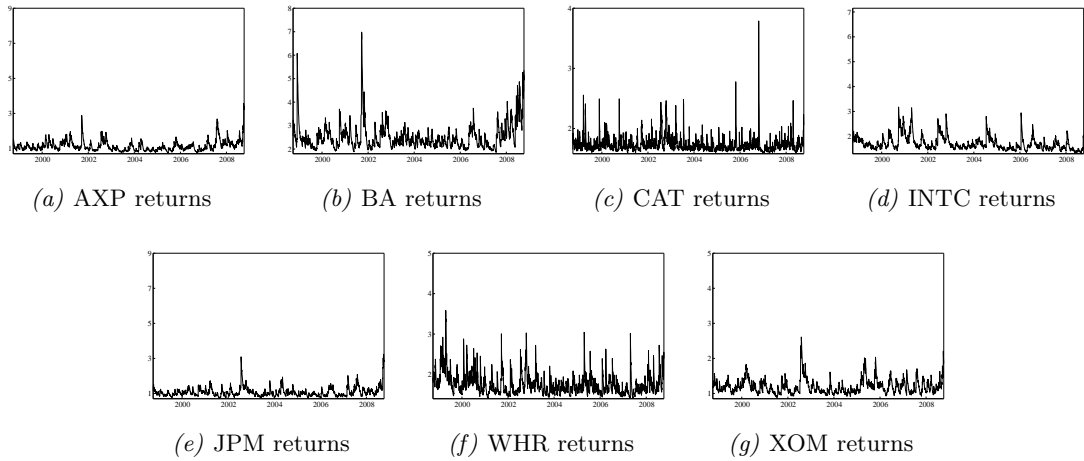


Figure 4: Estimated conditional standard deviations from the GJR-GARCH(1,1) model for the standardised variable $\varepsilon_t/\hat{g}_t^{1/2}$ as in the TVGJR-GARCH model for the seven stock returns of the S&P 500 composite index.

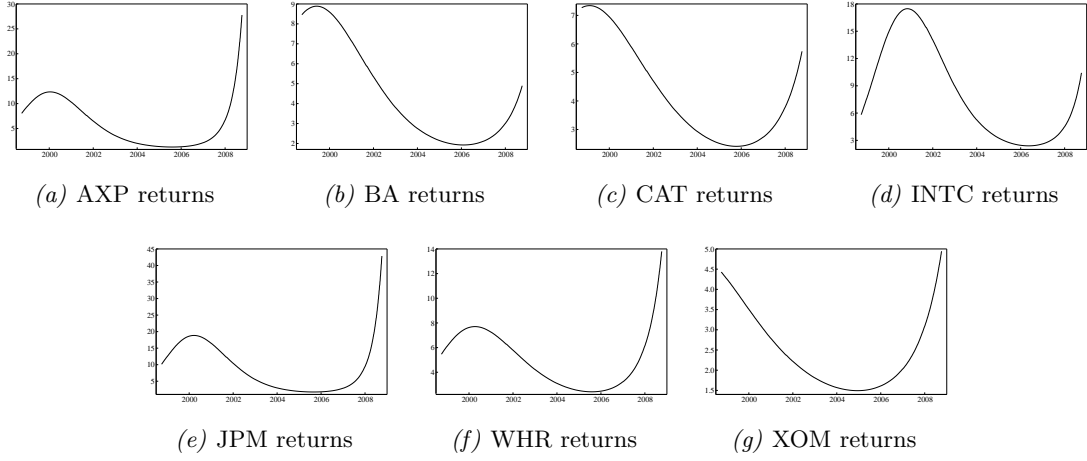


Figure 5: Estimated g_t functions for the Spline-GJR-GARCH model for the seven stock returns of the S&P 500 composite index.

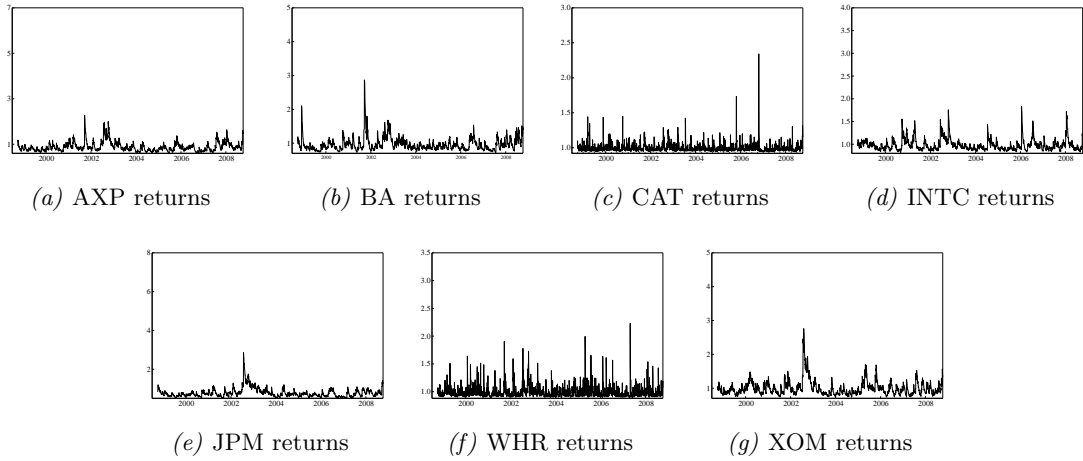


Figure 6: Estimated conditional standard deviations from the GJR-GARCH(1,1) model for the standardised variable $\varepsilon_t/\hat{g}_t^{1/2}$ as in the Spline-GJR-GARCH model for the seven stock returns of the S&P 500 composite index.

suggest that there are also differences in conditional correlations between models based on GJR-GARCH type variances and their TVGJR-GARCH and Spline-GJR-GARCH counterparts. A look at Figure 8 suggests, perhaps surprisingly, that when one compares DCC-GJR-GARCH models with DCC-TVGJR-GARCH and DCC-Spline-GJR-GARCH ones, this is not the case. The figure depicts the differences between the conditional correlations over time for the 21 bivariate models. They are generally rather small, and it is difficult to find any systematic pattern in them. The CAT-WHR pair is the only exception: the difference between the correlations lies within the interval $(-0.22, 0.30)$ in the DCC-TVGJR-GARCH case. The difference on the correlations for the DCC-Spline-GJR-GARCH model is generally larger than that of the DCC-TVGJR-GARCH model. To save space, the correlations estimated from the CCC- and VC-GJR-GARCH models are not shown. A general finding is that the correlations from the CCC-TVGJR-GARCH model remain very close to the ones obtained from the CCC-GJR-GARCH model. The same is true for the VC-TVGJR-GARCH model as the modelled nonstationarity in the variances only has a small effect on time-varying correlations. One may thus conclude that if the focus of the modeller is on conditional correlations, taking nonstationarity in the variance into account is not particularly important.

Nevertheless, the fit of the models considerably improves when the unconditional variance component is properly modelled. The log-likelihood values for each 7-dimensional CC-GJR-GARCH model are reported in Table 5. The maxima of the log-likelihood functions are substantially higher when g_{it} is estimated than when it is ignored. In particular, the best fitting model is the VC-Spline-GJR-GARCH model, followed by the DCC-Spline-GJR-GARCH and the VC-TVGJR-GARCH model.

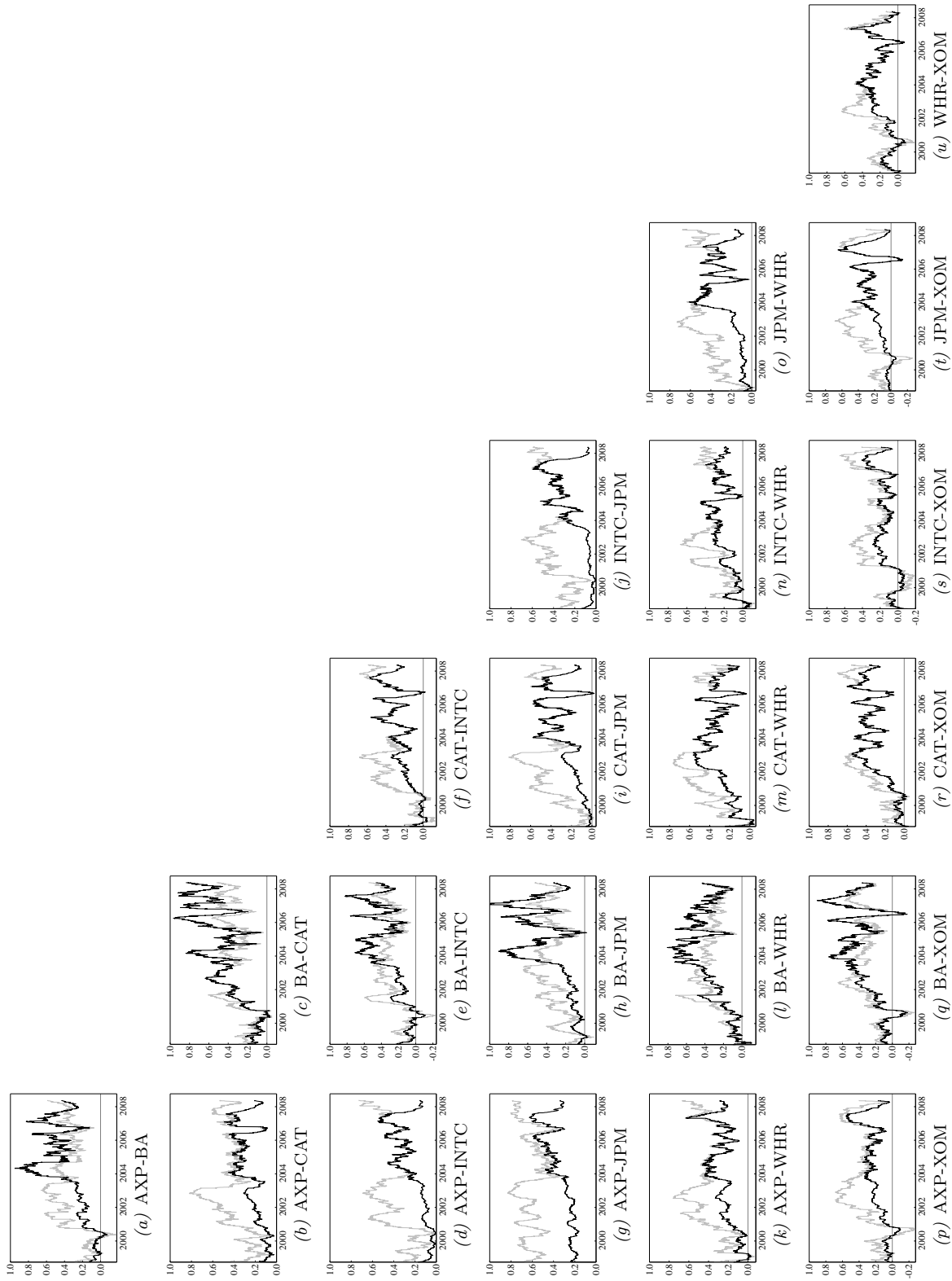


Figure 7: Differences between the estimated rolling correlation coefficients for pairs of the raw returns (grey curve) and pairs of the standardised returns (black curve).

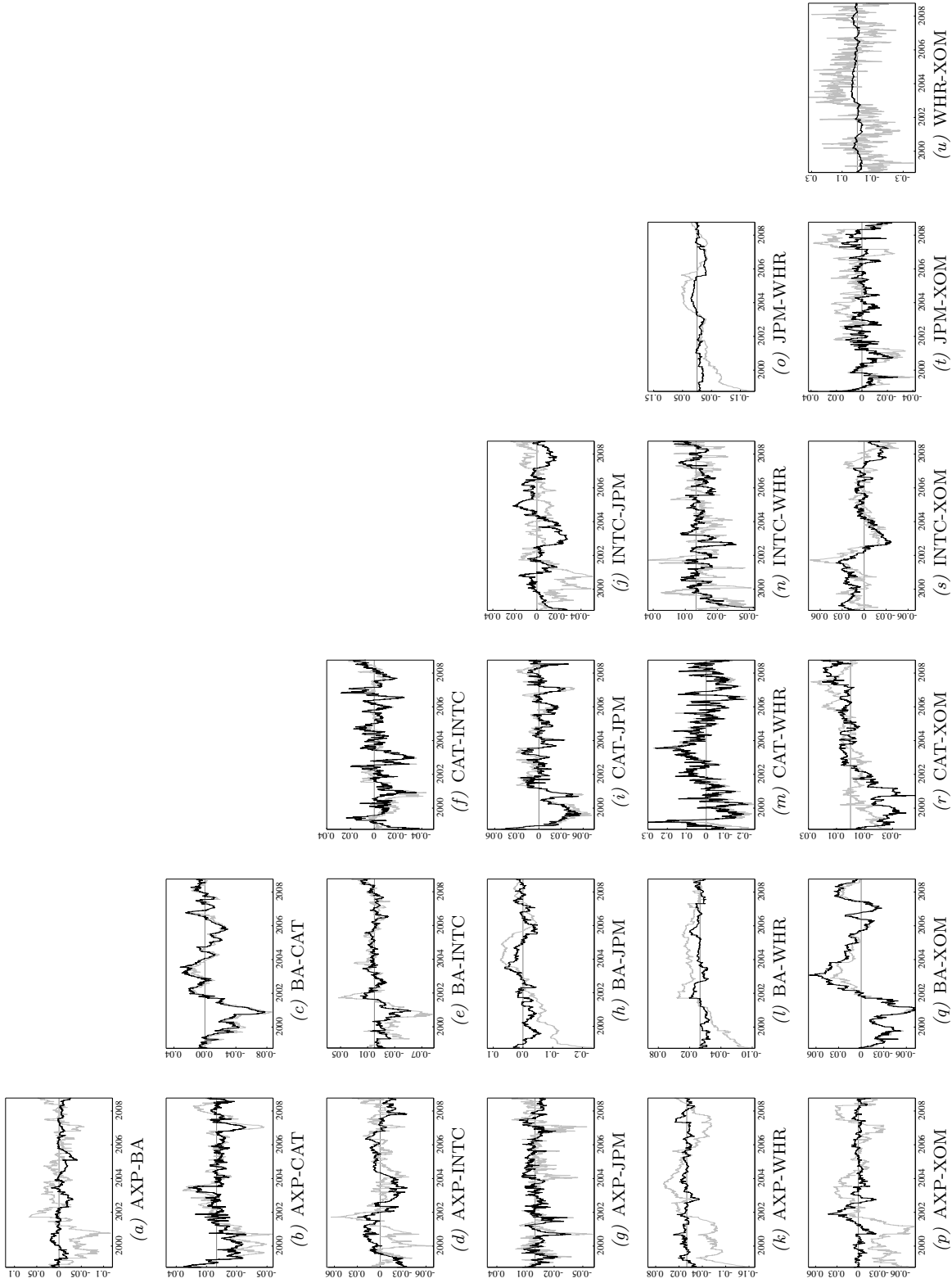


Figure 8: Differences between the estimated conditional correlations obtained from the bivariate DCC-GJR-GARCH and the DCC-TVJR-GARCH models (black curve), and difference between the estimated conditional correlations obtained from the bivariate DCC-GJR-GARCH and the DCC-Spline-GJR-GARCH models (grey curve) for the asset returns.

Table 5: Log-likelihood values from the 7-variate normal density for the CC-GJR-GARCH models

Models	GJR-GARCH	TVGJR-GARCH	Spline-GJR-GARCH
CCC-GARCH	-34772.3	-34608.4	-34497.2
DCC-GARCH	-34640.0	-34490.3	-34401.9
VC-GARCH	-34623.3	-34477.9	-34389.9

Notes: The GJR-GARCH column indicates that the unconditional variances are time-invariant functions. The TVGJR-GARCH column indicates that the unconditional variances vary over time according to function (8).

5.4 Evaluating forecasting performance

To evaluate the forecasting performance of the multivariate CC-GARCH models, we consider a rolling scheme for the estimation of parameters using a fixed window of 2521 daily observations. Specifically, the first set of one-step-ahead covariance forecasts is based on the estimation period from September 29, 1998 until October 7, 2008. To generate the next set of covariance forecasts, the window is rolled forward one day to obtain the second set of daily covariance forecasts. We repeat this process by adding the next observation and discarding the earliest return until we reach the end of the out-of-sample period. After performing this procedure, we compute 311 one-step-ahead covariance forecasts, each based on the estimation of 2521 returns. As already mentioned, the overall out-of-sample period ranges from October 8, 2008 to December 31, 2009.

The evaluation consists of comparing the predicted covariance matrix with the true matrix. Since the true covariance matrix is unobserved, following Andersen, Bollerslev, Diebold and Labys (2003) we use as a proxy the realized covariance estimator defined as the sum of the outer-product of intra-daily returns over the forecast horizon. All computations are based on one-day-ahead forecasts of the covariance matrix over 311 days. The intra-daily data consist of tick-by-tick trade prices for the seven series from the NYSE Trade and Quote database (TAQ) sampled from 9:30 until 16:00 at 5-minute intervals. Intraday returns, $r_{j,t}$, are computed as

$$r_{j,t} = p_{j\Delta,t} - p_{(j-1)\Delta,t}, \quad j = 1, \dots, M$$

where $\Delta = 1/M$ and $p_{j\Delta,t}$ is the log price at time $j\Delta$ in day t .

In Figure 9 we plot the differences between the estimated correlations obtained from the 7-dimensional VC-GJR-GARCH and the VC-TVGJR-GARCH models in the out-of-sample period.

The differences in correlations between the two DCC-GARCH models are not shown since they do not differ much from those found in-sample. The conditional correlations estimated from the VC-GJR-GARCH model are generally larger than the ones obtained from the VC-TVGJR-GARCH model, but the differences are rather small. With few exceptions, the difference reaches its maximum around the middle of the period, after which it suddenly decreases.

To compare the accuracy of the one-day-ahead covariance matrix forecasts we consider the following loss function based on the Frobenius distance of the forecast error; see Patton and Sheppard (2009):

$$L_{F,T+i} = (1/N^2) \left\{ \text{vec} \left(\boldsymbol{\Sigma}_{T+i} - \widehat{\boldsymbol{\Sigma}}_{T+i} \right)' \text{vec} \left(\boldsymbol{\Sigma}_{T+i} - \widehat{\boldsymbol{\Sigma}}_{T+i} \right) \right\}$$

where $\widehat{\boldsymbol{\Sigma}}_{T+i}$ is the one-step-ahead forecast of the covariance matrix for time $T+i$ and $\boldsymbol{\Sigma}_{T+i}$ is the true covariance matrix proxied by the realized covariance estimator. This function is the squared error generalised to matrix spaces. To measure the forecasting performance we shall consider the Root Mean Squared Error (RMSE) but also the Mean Absolute Deviation (MAD) based on the L_1 norm:

$$L_{1,T+i} = (1/N^2) \left\{ \left\| \text{vec} \left(\boldsymbol{\Sigma}_{T+i} - \widehat{\boldsymbol{\Sigma}}_{T+i} \right) \right\|_1 \right\}.$$

Criteria based on the absolute deviations are sometimes preferred because they are less affected by outliers than RMSE. To reduce the impact of outlying observations on forecasting evaluation even further, we also report values of the Median Squared Error (MedSE). Note, however, that MAD and MedSE are not consistent criteria in the sense that they do not necessarily preserve the correct ordering of the losses from the models under consideration when the true covariance matrix is unknown, see Laurent, Rombouts and Violante (2013) for discussion.

Table 6 presents values of the three criteria for the one-day horizon. According to MAD and MedSE, the models with time-varying unconditional variances perform better than the ones without this feature. Interestingly, the CCC-TVGJR-GARCH model outperforms the others, which suggests that modelling time-variation in correlations is not crucial in forecasting, at least not in the short run. If we instead consider RMSE as a measure of predictive ability, the differences between models are very small, and now the DCC-TVGJR-GARCH model is slightly superior to the rest. Obviously the CC-TVGJR-GARCH models generate some rather inaccurate forecasts

Table 6: Out-of-sample prediction accuracy for the conditional covariance matrices

Models	Statistics of predictive ability			MCS	
	RMSE ₁	MAD ₁	MedSE ₁	T_D	p -value
CCC-GJR-GARCH	9.834	4.307	28.04	96.71	0.000
CCC-TVGJR-GARCH	9.001	3.726	10.14	81.01	0.695
CCC-Spline-GJR-GARCH	9.330	4.284	18.87	87.06	0.000
DCC-GJR-GARCH	10.21	5.101	31.55	104.25	0.000
DCC-TVGJR-GARCH	8.952	4.125	12.88	80.13	1.000
DCC-Spline-GJR-GARCH	9.388	4.766	22.48	88.13	0.000
VC-GJR-GARCH	10.37	5.245	32.71	107.60	0.000
VC-TVGJR-GARCH	8.989	4.113	12.85	80.81	0.695
VC-Spline-GJR-GARCH	9.330	4.432	22.84	88.97	0.000

Notes: The out-of-sample forecast evaluation statistics are the Root Mean Squared Error (RMSE), the Mean Absolute Deviation (MAD) and the Median Squared Error (MedSE) criteria. The loss function is based on the Frobenius distance of the forecast error as in Patton and Sheppard (2009). The test statistic T_D denotes the deviation statistic of Hansen et al. (2011). The results are based on the out-of-sample data from October 8, 2008 to December 31, 2009 (311 observations).

that are downweighted by the use of MAD or MedSE.

In addition, to test joint forecasting performance across all models we make use of the Model Confidence Set (MCS) approach of Hansen, Lunde and Nason (2011). We set the confidence level of MCS equal to 0.1. The block length equals two, and the number of bootstrap samples is 10,000. The results can be found in Table 6. MCS contains the three CC-TVGJR-GARCH models. This means that out-of-sample performance of the CC-GJR-GARCH models with time-varying unconditional variance is superior to that of the models with constant unconditional variance. It also means that in this experiment the TV-GARCH approach performs somewhat better than the spline-GARCH one. A general conclusion is that careful modelling of the unconditional variance is very important in forecasting covariance matrices, at least in the short run.

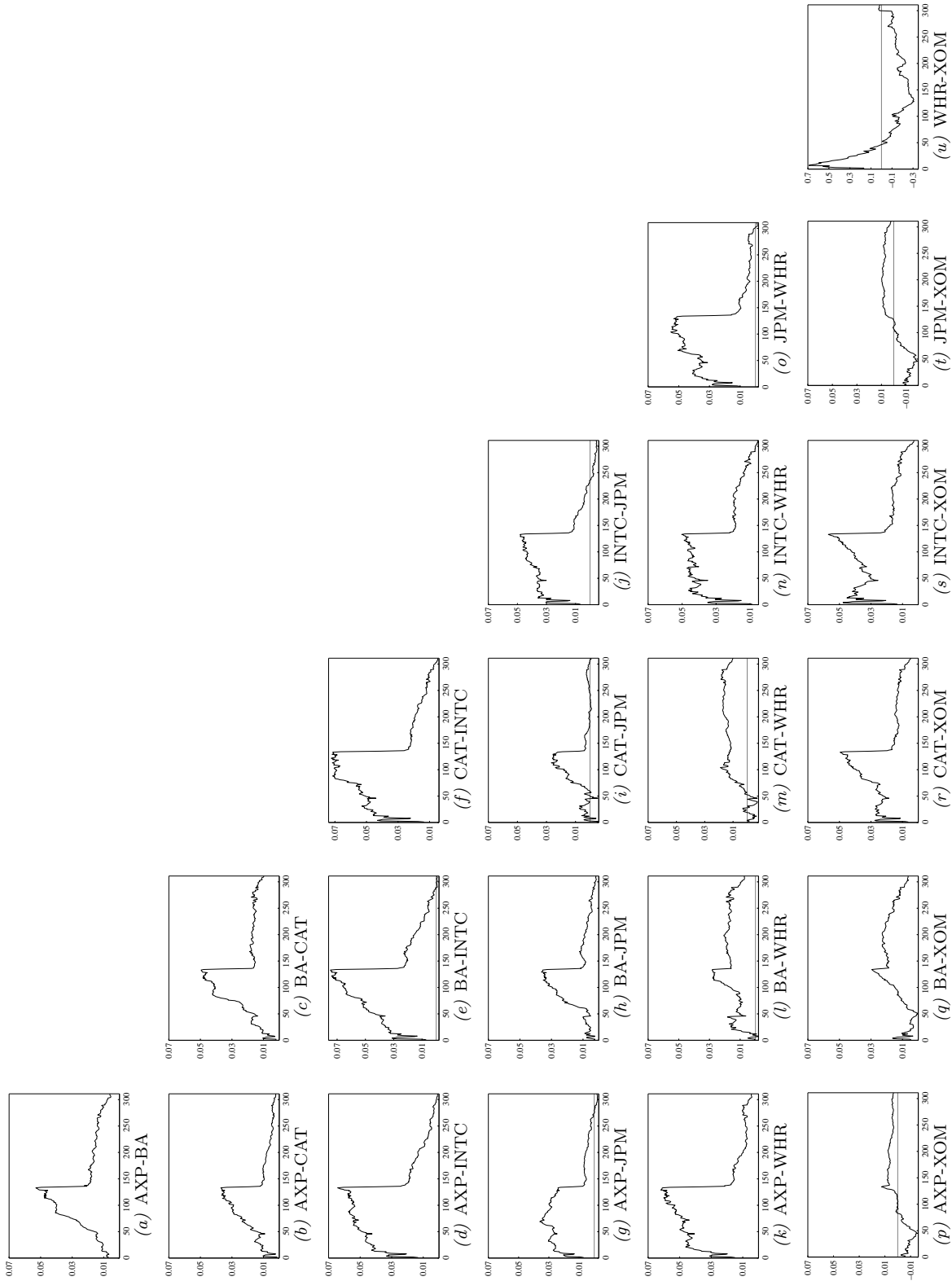


Figure 9: Differences between the estimated conditional correlations obtained from the 7-variate VC-GJR-GARCH and the 7-variate VC-TVGJR-GARCH models for the asset returns in the out-of-sample period from October 8, 2008 to December 31, 2009.

6 Empirical analysis II: Portfolio allocation

Another way of evaluating our models out-of-sample is to consider the economic value of volatility timing; see Fleming, Kirby and Ostdiek (2001) and Engle and Colacito (2006) among others. We do this by applying three portfolio allocation strategies both when the unconditional variance is time-varying and when it is not. These optimal asset-allocation strategies are the global minimum variance (GMV), the minimum variance with target expected return (Min-Variance) and the mean-variance (Mean-Variance) strategy. In implementing them, the weights are based on forecasts of the conditional covariance matrix associated with each CC-GARCH model.

According to the mean-variance strategy, for each time t the investor solves the quadratic programming problem

$$\min_{\mathbf{w}_t} \mathbf{w}_t' \boldsymbol{\Sigma}_t \mathbf{w}_t - \frac{1}{\gamma} \boldsymbol{\mu}_t' \mathbf{w}_t \quad \text{s.t.} \quad \mathbf{w}_t' \mathbf{1} = 1$$

where \mathbf{w}_t is an $N \times 1$ vector of portfolio weights, $\boldsymbol{\mu}_t$ is the $N \times 1$ vector of expected returns, $\mathbf{1}$ is an N -dimensional vector of ones, and γ is the risk aversion parameter. The GMV strategy corresponds to the mean-variance portfolio with an infinite risk aversion parameter. The Min-Variance strategy aims at finding the portfolio that has the smallest risk, measured by the portfolio variance, that achieves a target expected return

$$\min_{\mathbf{w}_t} \mathbf{w}_t' \boldsymbol{\Sigma}_t \mathbf{w}_t \quad \text{s.t.} \quad \mathbf{w}_t' \boldsymbol{\mu}_t = \mu_p \text{ and } \mathbf{w}_t' \mathbf{1} = 1$$

where μ_p is the target expected rate of return of the portfolio. We set $\mu_p = 8$. No short selling restrictions are imposed on any strategy.

For each strategy, we compute the annualized excess returns, the annualized standard deviation, and the annualized Sharpe ratio. We form six portfolios, each containing five stocks, each representing a different industry. The portfolios are rebalanced daily for each covariance estimator CC-GARCH model. As the risk-free asset needed for computing Sharpe ratios we use the three-month Treasury bill rate. Following Fleming et al. (2001), we also consider a utility-based measure denoted by Δ_γ to compare the performance of any CC-TVGJR-GARCH model to that of the corresponding CC-GJR-GARCH model. The value of Δ_γ is such that an investor with a quadratic utility function and the relative risk aversion parameter γ is indifferent between receiving r_t and

$r_{TVt} - \Delta_\gamma$, where r_t is the out-of-sample portfolio return under the the CC-GJR-GARCH model and r_{TVt} the corresponding return under the CC-TVGJR-GARCH model.

Tables 7 and 8 report the out-of-sample portfolio performance of the volatility-timing strategies for the CCC-GARCH and the DCC-GARCH models. As an indicator for the transaction costs we also show the portfolio turnover defined as the average of daily absolute changes in portfolio weights $|w_{it} - w_{it-}|$ over the period, where w_{it} is the optimal weight of asset i on day t and w_{it-} is the weight of the same asset at the end of the day $t - 1$, that is, before rebalancing the portfolio. The results show that when the time-varying unconditional variance is modelled, the annualized standard deviation of the portfolios tends to increase in two cases out of three, the GMV strategy being an exception. But then, with rather few exceptions, the strategies using covariance matrices with time-varying unconditional variances generate higher Sharpe ratios than those based on unmodelled long-run variances. This is mainly due to the increase of the annualized excess returns of the investment portfolios. The asset combination BA-CAT-INTC-JPM-XOM is the most notable example of this. A move from the DCC-GJR-GARCH model to the DCC-TVGJR-GARCH one pushes the expected excess returns of the Mean-Variance portfolio up from 2.24% to 21.45%. This in turn translates into an increase of the Sharpe ratio from 0.030 to 0.753. Large differences in the Sharpe ratio in favour of CC-TVGJR-GARCH models can be seen in most occasions. This is true both for the constant and for the dynamic conditional correlation models. Applying VC-GJR-GARCH models leads to similar conclusions, so the results are omitted to save space.

DeMiguel and Nogales (2009) indicate that the Mean-Variance portfolio usually has more unstable weights than the others. This implies a higher portfolio turnover and higher transaction costs. Our results accord with this observation. When the unconditional variances are time-varying, the weights of this portfolio strategy tend to fluctuate more over time than when they are not. On the other hand, the turnover rates of the GMV and Min-Variance strategies tend to be lower when the long-run variances are modelled than when they are not. One can thus argue that the CCC-TVGJR-GARCH and the DCC-TVGJR-GARCH models outperform their conventional counterparts when these two portfolio strategies are being applied.

Finally, Tables 7 and 8 also contain the estimates of Δ_γ as annualized fees in basis points using two different risk aversion parameters, $\gamma = 1$ and $\gamma = 10$. Over all six asset combinations and the

Table 7: Out-of-sample portfolio performance of the volatility-timing strategies: CCC-GJR-GARCH vs. CCC-TVGJR-GARCH

Portfolios	CCC-GJR-GARCH			CCC-TVGJR-GARCH			Δ_γ			
	μ	σ	SR	Turnover	μ	σ	SR	Turnover	$\gamma = 1$	$\gamma = 10$
AXP-BA-CAT-INTC-JPM										
GMV	11.82	17.64	0.670	0.017	7.136	16.22	0.440	0.007	11.43	13.98
Min-Variance	12.04	19.00	0.634	0.014	14.34	22.31	0.643	0.015	206.4	284.8
Mean-Variance	12.09	27.08	0.446	0.039	13.46	29.52	0.456	0.049	25.03	26.73
AXP-BA-CAT-INTC-WHR										
GMV	8.035	17.85	0.450	0.018	11.35	16.36	0.694	0.010	21.01	22.53
Min-Variance	5.502	19.20	0.287	0.011	14.20	22.09	0.643	0.014	27.00	282.0
Mean-Variance	7.578	27.12	0.279	0.030	10.83	29.44	0.368	0.050	19.96	21.47
AXP-BA-CAT-INTC-XOM										
GMV	-8.258	16.15	-0.511	0.015	-0.074	13.38	-0.006	0.018	3.03	0.814
Min-Variance	7.902	18.74	0.422	0.027	13.78	20.99	0.656	0.024	25.96	273.6
Mean-Variance	17.78	26.94	0.660	0.032	16.33	30.36	0.538	0.057	30.55	32.44
BA-CAT-INTC-JPM-WHR										
GMV	4.105	17.88	0.230	0.023	3.986	15.96	0.250	0.008	5.897	7.742
Min-Variance	2.722	20.43	0.133	0.019	10.99	25.41	0.433	0.018	20.76	218.0
Mean-Variance	1.970	21.91	0.090	0.024	14.90	26.54	0.561	0.035	28.70	29.68
BA-CAT-INTC-JPM-XOM										
GMV	-8.327	16.02	-0.520	0.021	1.691	12.94	0.131	0.014	5.137	3.537
Min-Variance	8.881	20.41	0.435	0.026	11.46	26.28	0.436	0.025	21.14	227.1
Mean-Variance	5.660	24.21	0.234	0.023	16.55	28.38	0.583	0.056	31.78	32.97
CAT-INTC-JPM-WHR-XOM										
GMV	-19.91	17.46	-1.140	0.033	-18.53	20.32	-0.912	0.038	0.065	0.002
Min-Variance	2.543	22.52	0.113	0.013	8.705	27.38	0.318	0.043	16.11	172.1
Mean-Variance	-1.520	25.48	-0.060	0.015	12.72	29.61	0.430	0.058	24.57	25.32

Notes: The table summarizes the out-of-sample performance for three sets of portfolio dynamic weights: global minimum variance (GMV), minimum variance with target expected return equal to 8% (Min-Variance) and mean-variance (Mean-Variance) portfolio strategy. For each set of weights, we report the annualized mean excess returns (μ), the annualized standard deviation (σ), the annualized Sharpe-ratio (SR) and the average daily turnover over the out-of-sample period from October 8, 2008 to December 31, 2009. We also report the average annualized basis point fees that an investor with quadratic utility and constant relative risk aversion of $\gamma = 1$ or $\gamma = 10$ would be willing to pay to switch from the constant to the time-varying unconditional variance strategy.

Table 8: Out-of-sample portfolio performance of the volatility-timing strategies: DCC-GJR-GARCH vs. DCC-TVGJR-GARCH

Portfolios	DCC-GJR-GARCH			DCC-TVGJR-GARCH			Δ_γ			
	μ	σ	SR	Turnover	μ	σ	SR	Turnover	$\gamma = 1$	$\gamma = 10$
AXP-BA-CAT-INTC-JPM										
GMV	10.09	22.81	0.442	0.013	5.325	16.59	0.321	0.007	7.442	10.41
Min-Variance	21.51	27.81	0.773	0.015	13.97	22.78	0.613	0.014	25.79	27.71
Mean-Variance	18.10	25.39	0.713	0.027	14.67	29.23	0.502	0.042	27.11	29.14
AXP-BA-CAT-INTC-WHR										
GMV	3.707	18.49	0.200	0.016	3.944	17.08	0.231	0.010	6.000	7.720
Min-Variance	5.435	19.90	0.273	0.011	7.602	22.71	0.335	0.013	13.47	14.98
Mean-Variance	2.881	27.19	0.106	0.030	4.166	29.20	0.143	0.043	6.690	8.142
AXP-BA-CAT-INTC-XOM										
GMV	-10.57	16.38	-0.645	0.018	-3.164	13.71	-0.231	0.012	1.415	0.070
Min-Variance	7.685	19.13	0.402	0.032	12.67	21.41	0.591	0.022	23.69	25.11
Mean-Variance	17.70	26.98	0.656	0.029	15.06	30.41	0.495	0.049	27.92	29.89
BA-CAT-INTC-JPM-WHR										
GMV	-0.601	18.18	-0.033	0.024	2.037	16.32	0.125	0.006	3.553	4.000
Min-Variance	4.325	20.98	0.206	0.015	9.648	25.90	0.373	0.017	17.88	19.14
Mean-Variance	2.485	22.11	0.112	0.021	18.30	26.98	0.678	0.054	35.49	36.49
BA-CAT-INTC-JPM-XOM										
GMV	-9.261	16.09	-0.576	0.022	-1.549	13.25	-0.117	0.009	2.079	0.289
Min-Variance	5.255	20.61	0.255	0.018	9.991	26.50	0.377	0.024	18.48	19.81
Mean-Variance	2.237	23.82	0.094	0.030	21.45	28.49	0.753	0.068	41.82	42.80
CAT-INTC-JPM-WHR-XOM										
GMV	-17.69	17.79	-0.995	0.026	-30.68	20.67	-1.485	0.039	-0.421	-0.051
Min-Variance	3.927	22.83	0.172	0.015	8.359	27.86	0.300	0.028	15.24	16.52
Mean-Variance	0.439	25.53	0.017	0.034	13.20	30.17	0.438	0.058	25.38	26.28

Notes: The table summarizes the out-of-sample performance for three sets of portfolio dynamic weights: global minimum variance (GMV), minimum variance with target expected return equal to 8% (Min-Variance) and mean-variance (Mean-Variance) portfolio strategy. For each set of weights, we report the annualized mean excess returns (μ), the annualized standard deviation (σ), the annualized Sharpe-ratio (SR) and the average daily turnover over the out-of-sample period from October 8, 2008 to December 31, 2009. We also report the average annualized basis point fees that an investor with quadratic utility and constant relative risk aversion of $\gamma = 1$ or $\gamma = 10$ would be willing to pay to switch from the constant to the time-varying unconditional variance strategy.

three strategies, a conservative investor, $\gamma = 10$, would be willing to pay an annual management fee between the minimum, 0.002 (−0.051) basis points, and maximum, 284.8 (42.80) points, when implementing the CCC-TVGJR-GARCH (DCC-TVGJR-GARCH) method. These outcomes strengthen the impression that the conditional correlation models with time-varying unconditional variances outperform their counterparts with constant unconditional variances.

7 Time-varying news impact surfaces

In this section we consider the impact of unexpected shocks to the asset returns on the estimated covariances. This is done by employing a generalization of the univariate *news impact curve* of Engle and Ng (1993) to the multivariate case introduced by Kroner and Ng (1998). The so-called *news impact surface* is the plot of the conditional covariance against a pair of lagged shocks, holding the past conditional covariances constant at their unconditional sample mean levels. The news impact surfaces of the multivariate correlation models with the volatility equations modelled as TVGJR-GARCH models are time-varying because they depend on the component $g_{i,t-1}$. They will be called *time-varying news impact surfaces*. The time-varying news impact surface for h_{ijt} is the three-dimensional graph of the function

$$h_{ijt} = f(\varepsilon_{i,t-1}/g_{i,t-1}^{1/2}, \varepsilon_{j,t-1}/g_{j,t-1}^{1/2}, \rho_{ij,t-1}; \mathbf{h}_{t-1})$$

where \mathbf{h}_{t-1} is a vector of conditional covariances at time $t - 1$ defined at their unconditional sample means. As an example, Figure 10 contains the time-varying news impact surface for the covariance generated by the CCC-TVGJR-GARCH model for the pair BA-XOM. The choice of this particular pair of assets is merely illustrative, but similar surfaces can be found for other pairs as well. It is seen how the surface can vary over time due to the nonstationary components $g_{i,t-1}$ and $g_{j,t-1}$. We are able to distinguish different reaction levels of covariance estimates to past shocks during tranquil and turbulent times. It shows that the response to the news of a given size on the estimated covariances is clearly stronger during periods of calm in the market ('lower regime') than it is during periods of high turbulence. According to the results, when calm prevails a minor piece of 'bad news' (unexpected negative shock) is rather big news compared to a big piece of 'good news' (unexpected positive shock) during turbulent periods. This is seen from the

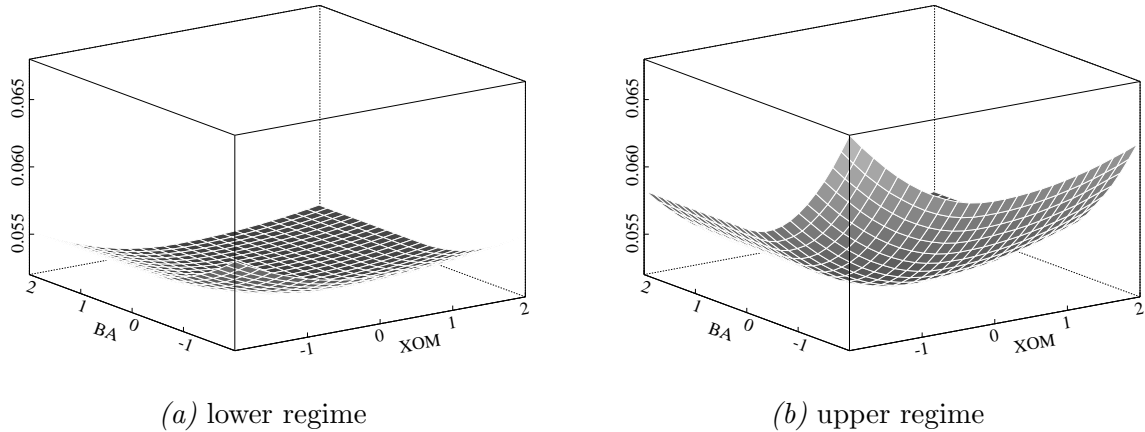


Figure 10: Estimated estimated time-varying news impact surfaces for the covariance between the BA and XOM returns under the CCC-TVGJR-GARCH model (a) in the lower regime and (b) in the upper regime of volatility.

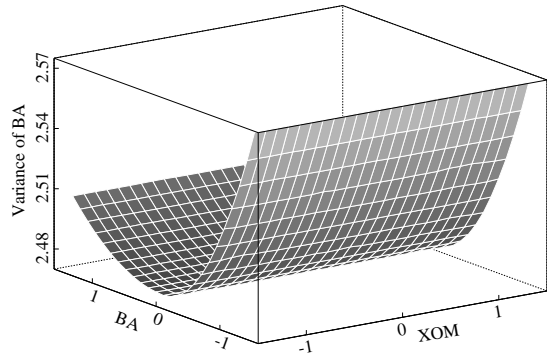
asymmetric bowl-shaped impact surface.

Figure 11 contains the time-varying news impact surfaces under low and high volatility from the CCC-TVGJR-GARCH model for the conditional variance of BA when there is no shock to XOM. Figure 12 contains a similar graph for XOM when there is no shock to BA. The asymmetric shape shows that a negative return shock has a greater impact than a positive return shock of the same size. Furthermore, as already seen from Figure 10, a piece of news of a given size has a stronger effect on the conditional variance when volatility is low than when it is high.

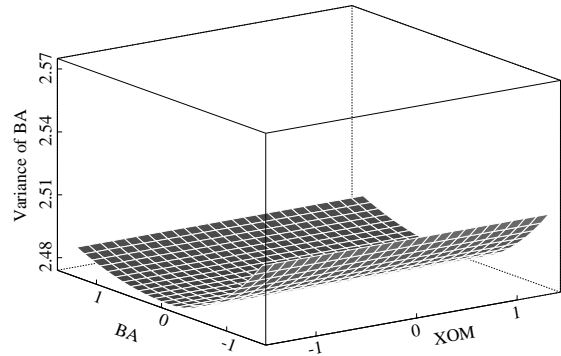
8 Conclusions

In this paper, we extend the univariate multiplicative TV-GARCH model of Amado and Teräsvirta (2012, 2013) to the multivariate CC-GARCH framework. The model allows the individual variances to vary smoothly over time according to the logistic transition function and its generalizations. We develop a modelling technique for specifying the parametric structure of the deterministic time-varying component that involves a sequence of Lagrange multiplier-type tests. In this respect, our model differs from the semiparametric model of Hafner and Linton (2010).

We consider a set of CC-GARCH models to investigate the effects of nonstationary variance

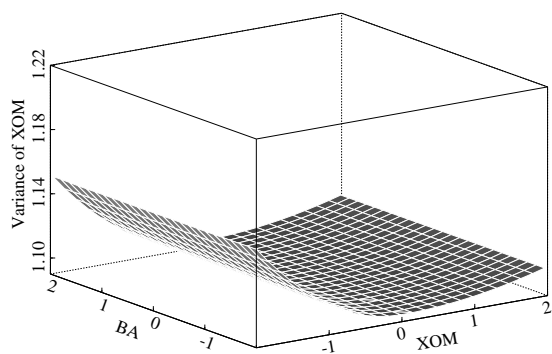


(a) lower regime

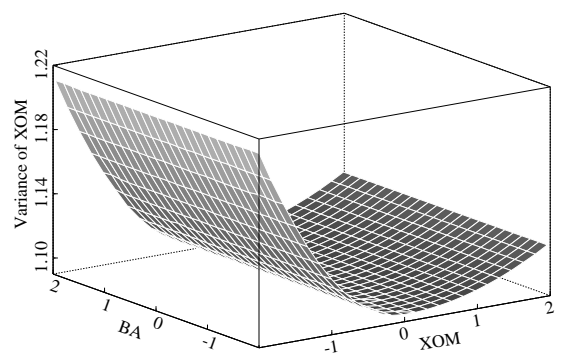


(b) upper regime

Figure 11: Estimated time-varying news impact surfaces for the conditional variance of the BA returns under the CCC-TVGJR-GARCH model (a) in the lower regime and (b) in the upper regime of volatility.



(a) lower regime



(b) upper regime

Figure 12: Estimated time-varying news impact surfaces for the conditional variance of the XOM returns under the CCC-TVGJR-GARCH model (a) in the lower regime and (b) in the upper regime of volatility.

equations on the conditional correlation matrix. The models are applied to pairs of seven daily stock returns belonging to the S&P 500 composite index and to the 7-variate case. We find that modelling the time-variation of the unconditional variances considerably improves the fit of the CC-GARCH models. The results show that multivariate correlation models combining both time-varying correlations and time-varying unconditional variances provide the best in-sample fit. They also indicate that modelling the nonstationary component in the variance has relatively little effect on correlation estimates when the conditional correlation model is the DCC-GARCH model.

The results on forecasting show that the CC-GARCH models with time-varying unconditional variances clearly outperform the others when the comparison is made using criteria robust to outliers. An interesting finding is that the CCC-TVGJR-GARCH model performs best, which suggests that modelling time-variation in correlations is not crucial in forecasting, at least not in the short run. Moreover, the out-of-sample portfolio analysis indicates that modelling the time-varying unconditional volatility is economically relevant. By applying three asset-allocation strategies we find that the conditional correlation models with time-varying unconditional variances tend to outperform their counterparts with constant unconditional variances.

Acknowledgments

This work has been supported by CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. The first author would also like to acknowledge financial support from the Louis Fraenckels Stipendiefond and the Danish Council for Independent Research (Grant No. 12-126018). Material from this paper has been presented at the 1st Humboldt - Copenhagen Conference in Financial Econometrics, Berlin, March 2009, Second Annual SoFiE Conference, Geneva, June 2009, European Time Series Econometrics Research Network Meeting, Lisbon, December 2010, 4th International Conference on Computational and Financial Econometrics (CFE 2010), London, December 2010, Aarhus-Xiamen Economics Workshop, Aarhus University, June 2012, Conference on Time Series Econometrics and Financial Econometrics, Southwestern University of Finance and Economics, Chengdu, October 2012, Exeter Workshop on Econometrics and its Applications in honour of Professor James Davidson, May 2013, the European Meeting of Statisticians, Budapest, July 2013, and in seminars at

Cass Business School, City University London, CREATES, Aarhus University, Emory University, Atlanta, Humboldt University, Berlin, Koç University, Istanbul, Queensland University of Technology, Brisbane, Soochow University, Suzhou, University of Heidelberg, University of Helsinki, and University of Sydney. We want to thank participants of their comments. We are also grateful to Francesco Violante, Gloria González-Rivera, Jose Gonzalo Rangel, José Faias, an Associate Editor and two anonymous referees for helpful remarks and suggestions. We thank Asger Lunde and Paolo Santucci de Magistris for providing us with the intra-daily data. Errors and shortcomings in this work are our own responsibility.

References

- Amado, C. and Teräsvirta, T.: 2008, Modelling unconditional and conditional heteroskedasticity with smoothly time-varying structure, *SSE/EFI Working Paper Series in Economics and Finance 691*, Stockholm School of Economics.
- Amado, C. and Teräsvirta, T.: 2012, Specification and testing of multiplicative time-varying GARCH models with applications, *Mimeo*, Aarhus University.
- Amado, C. and Teräsvirta, T.: 2013, Modelling volatility by variance decomposition, *Journal of Econometrics* **175**, 142–153.
- Andersen, T., Bollerslev, T., Diebold, F. and Labys, P.: 2003, Modelling and forecasting realized volatility, *Econometrica* **71**, 579–625.
- Bauwens, L., Laurent, S. and Rombouts, J.: 2006, Multivariate GARCH models: A survey, *Journal of Applied Econometrics* **21**, 79–109.
- Berkes, I., Gombay, E., Horváth, L. and Kokoszka, P.: 2004, Sequential change-point detection in GARCH(p,q) models, *Econometric Theory* **20**, 1140–1167.
- Bollerslev, T.: 1990, Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model, *Review of Economics and Statistics* **72**, 498–505.
- Dahlhaus, R. and Subba Rao, S.: 2006, Statistical inference for time-varying ARCH processes, *Annals of Statistics* **34**, 1075–1114.

- DeMiguel, V. and Nogales, F. J.: 2009, Portfolio selection with robust estimation, *Operations Research* **57**, 560–577.
- Diebold, F. X.: 1986, Comment on 'Modelling the persistence of conditional variance', by R.F. Engle and T. Bollerslev, *Econometric Reviews* **5**, 51–56.
- Doornik, J. A.: 2009, *An Object-Oriented Matrix Language Ox 6*, London: Timberlake Consultants Press and Oxford: www.doornik.com.
- 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 Colacito, R.: 2006, Testing and valuing dynamic correlations for asset allocation, *Journal of Business & Economic Statistics* **24**, 238–253.
- Engle, R. F. and Ng, V.: 1993, Measuring and testing the impact of news on volatility, *Journal of Finance* **48**, 1749–1778.
- Engle, R. F. and Rangel, J. G.: 2008, The spline-GARCH model for low-frequency volatility and its global macroeconomic causes, *Review of Financial Studies* **21**, 1187–1222.
- Engle, R. F. and Sheppard, K.: 2001, Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. NBER Working Paper 8554.
- Fleming, J., Kirby, C. and Ostdiek, B.: 2001, The economic value of volatility timing, *Journal of Finance* **56**, 329–352.
- Gallant, A. R.: 1981, On the bias in flexible functional forms and an essentially unbiased form. The Fourier flexible form, *Journal of Econometrics* **15**, 211–245.
- Glosten, L. R., Jagannathan, R. and Runkle, D. E.: 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* **48**, 1779–1801.
- Hafner, C. M. and Linton, O.: 2010, Efficient estimation of a multivariate multiplicative volatility model, *Journal of Econometrics* **159**, 55–73.

- Hagerud, G. E.: 1997, *A New Non-Linear GARCH Model*, PhD thesis, EFI, Stockholm School of Economics.
- Hansen, P. R., Lunde, A. and Nason, J. M.: 2011, Model confidence sets, *Econometrica* **79**, 453–497.
- Kim, T.-H. and White, H.: 2004, On more robust estimation of skewness and kurtosis, *Finance Research Letters* **1**, 56–73.
- Kroner, K. F. and Ng, V. K.: 1998, Modeling asymmetric comovements of asset returns, *Review of Financial Studies* **11**, 817–844.
- Lamoureux, C. G. and Lastrapes, W. D.: 1990, Persistence in variance, structural change, and the GARCH model, *Journal of Business & Economic Statistics* **8**, 225–234.
- Laurent, S., Rombouts, J. V. and Violante, F.: 2013, On loss functions and ranking forecasting performances of multivariate volatility models, *Journal of Econometrics* **173**, 1–10.
- Lavielle, M. and Teyssière, G.: 2006, Detection of multiple change-points in multivariate time series, *Lithuanian Mathematical Journal* **46**, 287–306.
- Ling, S. and McAleer, M.: 2003, Asymptotic theory for a new vector ARMA-GARCH model, *Econometric Theory* **19**, 280–310.
- Lütkepohl, H.: 1996, *Handbook of Matrices*, John Wiley & Sons, Chichester.
- Luukkonen, R., Saikkonen, P. and Teräsvirta, T.: 1988, Testing linearity against smooth transition autoregressive models, *Biometrika* **75**, 491–499.
- Mazur, B. and Pipień, M.: 2012, On the empirical importance of periodicity in the volatility of financial returns - time varying GARCH as a second order APC(2) process, *Central European Journal of Economic Modelling and Econometrics* **4**, 95–116.
- Mikosch, T. and Stărică, C.: 2004, Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects, *Review of Economics and Statistics* **86**, 378–390.

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

Silvennoinen, A. and Teräsvirta, T.: 2009, Multivariate GARCH models, *in* T. G. Andersen, R. A. Davis, J.-P. Kreiss and T. Mikosch (eds), *Handbook of Financial Time Series*, New York: Springer, pp. 201–229.

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

van Bellegem, S. and von Sachs, R.: 2004, Forecasting economic time series with unconditional time-varying variance, *International Journal of Forecasting* **20**, 611– 627.

A Maximum likelihood estimation of parameters

Equation (14) implies the following decomposition of the log-likelihood function for observation t :

$$\ell_t(\boldsymbol{\psi}, \boldsymbol{\varphi}, \boldsymbol{\phi}) = \ell_t^U(\boldsymbol{\psi}) + \ell_t^V(\boldsymbol{\psi}, \boldsymbol{\varphi}) + \ell_t^C(\boldsymbol{\psi}, \boldsymbol{\varphi}, \boldsymbol{\phi})$$

where $\ell_t^U(\boldsymbol{\psi}) = \sum_{i=1}^N \ell_{it}^U(\boldsymbol{\psi}_i)$ and $\ell_{it}^U(\boldsymbol{\psi}_i) = -(1/2)\{\ln g_{it}(\boldsymbol{\psi}_i) + \tilde{\varepsilon}_{it}^2/g_{it}(\boldsymbol{\psi}_i)\}$. Furthermore, $\ell_t^V(\boldsymbol{\psi}, \boldsymbol{\varphi}) = \sum_{i=1}^N \ell_{it}^V(\boldsymbol{\psi}_i, \boldsymbol{\varphi}_i)$, where $\ell_{it}^V(\boldsymbol{\psi}_i, \boldsymbol{\varphi}_i) = -(1/2)\{\ln h_{it}(\boldsymbol{\psi}_i, \boldsymbol{\varphi}_i) + \tilde{\varepsilon}_{it}^2/h_{it}(\boldsymbol{\psi}_i, \boldsymbol{\varphi}_i)\}$. Finally,

$$\ell_t^C(\boldsymbol{\psi}, \boldsymbol{\varphi}, \boldsymbol{\phi}) = -(1/2)\{\ln |\mathbf{P}_t(\boldsymbol{\psi}, \boldsymbol{\varphi}, \boldsymbol{\phi})| + \boldsymbol{\zeta}_t' \mathbf{P}_t^{-1}(\boldsymbol{\psi}, \boldsymbol{\varphi}, \boldsymbol{\phi}) \boldsymbol{\zeta}_t - 2\boldsymbol{\zeta}_t' \boldsymbol{\zeta}_t\}. \quad (16)$$

As is usual in the case of DCC-GARCH models, estimation is carried out in two steps. The GARCH equations are estimated first using maximization by parts, and $\boldsymbol{\phi}$ in (16) conditionally on GARCH estimates.

Maximization by parts works as follows: First reparameterise the deterministic component (8) as follows:

$$g_{it}^* = \delta_{i0}^* + \sum_{l=1}^r \delta_{il}^* G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il})$$

and set $\boldsymbol{\psi}_i^* = (\delta_{i0}^*, \boldsymbol{\delta}_i^{*'}, \boldsymbol{\gamma}_i', \mathbf{c}_i')'$, where $\delta_{i0}^* > 0$ and $\boldsymbol{\delta}_i^* = (\delta_{i1}^*, \dots, \delta_{ir}^*)'$ with $\delta_i^* = \delta_{i0}^* \delta_i$, so $g_{it}^* = \delta_{i0}^* g_{it}$.

1. Maximize

$$L_{iT}^U(\boldsymbol{\psi}^*) = \sum_{t=1}^T \ell_{it}^U(\boldsymbol{\psi}^*) = -(1/2) \sum_{t=1}^T \{\ln g_{it}^*(\boldsymbol{\psi}_i^*) + \tilde{\varepsilon}_{it}^2/g_{it}^*(\boldsymbol{\psi}_i^*)\}$$

for each i , $i = 1, \dots, N$, separately, assuming $\tilde{\varepsilon}_{it} = \varepsilon_{it}$, that is, setting $h_{it}(\boldsymbol{\psi}_i, \boldsymbol{\varphi}_i) \equiv 1$. The resulting estimators are $\hat{\boldsymbol{\psi}}_i^{*(1)} = (\hat{\delta}_{i0}^{*(1)}, \hat{\boldsymbol{\delta}}_i^{*(1)'}, \hat{\boldsymbol{\gamma}}_i^{(1)'}, \hat{\mathbf{c}}_i^{(1)'})'$, $i = 1, \dots, N$. Obtain $\hat{\boldsymbol{\delta}}_i^{(1)}$ as follows: $\hat{\boldsymbol{\delta}}_i^{(1)} = (\hat{\delta}_{i0}^{*(1)})^{-1} \hat{\boldsymbol{\delta}}_i^{*(1)}$ so that $\hat{\boldsymbol{\psi}}_i^{(1)} = (\hat{\boldsymbol{\delta}}_i^{(1)'}, \hat{\boldsymbol{\gamma}}_i^{(1)'}, \hat{\mathbf{c}}_i^{(1)'})'$. Note that $\hat{\delta}_{i0}^{*(1)} = \hat{\omega}_i^{(1)}$.

2. Setting $\boldsymbol{\psi}_i = \hat{\boldsymbol{\psi}}_i^{(1)}$, $i = 1, \dots, N$, in $\ell_t^V(\boldsymbol{\psi}, \boldsymbol{\varphi})$, maximize

$$L_{iT}^V(\hat{\boldsymbol{\psi}}_i^{(1)}, \boldsymbol{\varphi}_i) = \sum_{t=1}^T \ell_{it}^V(\hat{\boldsymbol{\psi}}_i^{(1)}, \boldsymbol{\varphi}_i) = -(1/2) \sum_{t=1}^T \{\ln h_{it}(\hat{\boldsymbol{\psi}}_i^{(1)}, \boldsymbol{\varphi}_i) + \varepsilon_{it}^{*2}/h_{it}(\hat{\boldsymbol{\psi}}_i^{(1)}, \boldsymbol{\varphi}_i)\}$$

with respect to $\boldsymbol{\varphi}_i$ assuming $\varepsilon_{it}^* = \varepsilon_{it}/g_{it}^{1/2}(\hat{\boldsymbol{\psi}}_i^{(1)})$, for each i , $i = 1, \dots, N$, separately. Call the i th resulting estimator $\hat{\boldsymbol{\varphi}}_i^{(1)}$.

The second iteration is as follows:

1. Maximize

$$L_{iT}^U(\boldsymbol{\psi}) = \sum_{t=1}^T \ell_{it}^U(\boldsymbol{\psi}_i) = -(1/2) \sum_{t=1}^T \{\ln g_{it}(\boldsymbol{\psi}_i) + \tilde{\varepsilon}_{it}^2/g_{it}(\boldsymbol{\psi}_i)\}$$

assuming $\tilde{\varepsilon}_{it} = \varepsilon_{it}/h_{it}^{1/2}(\hat{\boldsymbol{\psi}}_i^{(1)}, \hat{\boldsymbol{\varphi}}_i^{(1)})$, for each i , $i = 1, \dots, N$. This yields $\hat{\boldsymbol{\psi}}_i^{(2)}$. The important thing is that $\boldsymbol{\varphi}_i = \hat{\boldsymbol{\varphi}}_i^{(1)}$ (fixed) in the definition of $\tilde{\varepsilon}_{it}$.

2. Maximize

$$L_{iT}^V(\hat{\boldsymbol{\psi}}_i^{(2)}, \boldsymbol{\varphi}_i) = \sum_{t=1}^T \ell_{it}^V(\hat{\boldsymbol{\psi}}_i^{(2)}, \boldsymbol{\varphi}_i) = -(1/2) \sum_{t=1}^T \{\ln h_{it}(\hat{\boldsymbol{\psi}}_i^{(2)}, \boldsymbol{\varphi}_i) + \varepsilon_{it}^{*2}/h_{it}(\hat{\boldsymbol{\psi}}_i^{(2)}, \boldsymbol{\varphi}_i)\}$$

with respect to $\boldsymbol{\varphi}_i$ for each i , $i = 1, \dots, N$, separately, assuming $\varepsilon_{it}^* = \varepsilon_{it}/g_{it}(\hat{\boldsymbol{\psi}}_i^{(2)})$. This gives $\hat{\boldsymbol{\varphi}}_i^{(2)}$, $i = 1, \dots, N$.

Iterate until convergence. Call the resulting estimators $\hat{\boldsymbol{\psi}}_i$ and $\hat{\boldsymbol{\varphi}}_i$, $i = 1, \dots, N$, and set $\hat{\boldsymbol{\psi}} = (\hat{\boldsymbol{\psi}}_1', \dots, \hat{\boldsymbol{\psi}}_N')'$ and $\hat{\boldsymbol{\varphi}} = (\hat{\boldsymbol{\varphi}}_1', \dots, \hat{\boldsymbol{\varphi}}_N')'$.

After estimating the TV-GARCH equations, estimate ϕ given $\hat{\psi}_i$ and $\hat{\varphi}_i$ by maximizing

$$L_T^C(\phi) = \sum_{t=1}^T \ell_t^C(\phi) = -(1/2) \sum_{t=1}^T \{\ln |\mathbf{P}_t(\phi)| + \zeta_t' \mathbf{P}_t^{-1}(\phi) \zeta_t - 2\zeta_t' \zeta_t\}$$

where $\zeta_t = (\zeta_{1t}, \dots, \zeta_{Nt})'$ with $\zeta_{it} = \varepsilon_{it} / \{h_{it}(\hat{\psi}_i, \hat{\varphi}_i) g_{it}(\hat{\psi}_i)\}^{1/2}$, $i = 1, \dots, N$, and

$$\frac{\partial}{\partial \phi} L_T^C(\phi) = -(1/2) \sum_{t=1}^T \frac{\partial \text{vec}(\mathbf{P}_t)'}{\partial \phi} \text{vec}(\mathbf{P}_t^{-1} - \mathbf{P}_t^{-1} \zeta_t \zeta_t' \mathbf{P}_t^{-1}).$$

Under regularity conditions, Amado and Teräsvirta (2013) showed consistency and asymptotic normality of the ML estimators of the TVGJR-GARCH model. The results are valid for multivariate models under the assumption that the correlation matrix $\mathbf{P}_t = \mathbf{I}_N$.

Asymptotic properties of the maximum likelihood estimators only cover the CCC-GARCH model; see Ling and McAleer (2003). Due to a time-varying correlation matrix, deriving corresponding asymptotic results for the CC-TVGJR-GARCH models is a nontrivial problem and beyond the scope of the present paper.

All computations in this paper have been performed using the Ox programming language, versions 5.10 and 6.10, see Doornik (2009), the OxMetrics module G@RCH 6.1, MultCom 2.00 Ox package and a modified version of Matteo Pelagatti's source code¹.

¹The Ox estimation package is freely available at http://www.statistica.unimib.it/utenti/p_matteo/Ricerca/research.html