Conditional Correlation Models of Autoregressive Conditional Heteroskedasticity with Nonstationary GARCH Equations

Cristina Amado*
*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

August 2012

*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.

Acknowledgements: This research has been supported by the Danish National Research Foundation. The first author would like to acknowledge financial support from the Louis Fraenckels Stipendiefond. Part of this research was done while the first author was visiting CREATES, Aarhus University, whose kind hospitality is gratefully acknowledged. Material from this paper has been presented at the 1st Humboldt - Copenhagen Conference in Financial Econometrics, Berlin, Second Annual SoFiE Conference, Geneva, European Time Series Econometrics Research Network Meeting, Fall 2010, Lisbon, 4th International Conference on Computational and Financial Econometrics, London, and in seminars at CREATES, Aarhus University, Queensland University of Technology, Brisbane, and University of Sydney. We want to thank participants of their comments. We are also grateful to Gloria González-Rivera for helpful remarks. Errors and shortcomings in this work are our own responsibility.
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. The modelling technique to determine the parametric structure of this time-varying component is based on a sequence of specification Lagrange multiplier-type tests derived in Amado and Teräsvirta (2011). 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 considerably 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. 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; 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 (2009b).

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 Teyssiére (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 Gonzalo Rangel (2008) and, independently, Amado and Teräsvirta (2011) 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, whereas in Engle and Gonzalo Rangel (2008), it is an exponential spline. Amado and Teräsvirta (2011) described the nonstationary 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
compound. The parameters of both the unconditional and the conditional component were esti-
mated jointly.

Despite the growing literature on multivariate GARCH models, little attention has been de-
voted 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 Gonzalo Rangel (2008). Their multivariate GARCH
model is a first-order BEKK-GARCH model with a deterministic nonstationary or 'long run' com-
ponent. 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 estima-
tion 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 (2011) to the multivariate case. We investigate the effects of care-
ful 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 mod-
elled 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 $\{y_t\}, t = 1, ..., T$, described by the following vector
process:

$$y_t = \mathbb{E}(y_t|\mathcal{F}_{t-1}) + \varepsilon_t$$

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

$$\varepsilon_t = D_t \zeta_t$$

where $D_t$ is a diagonal matrix of time-varying standard deviations. The error vectors $\zeta_t$ form
a sequence of independent and identically distributed variables with mean zero and a positive
definite correlation matrix $P_t = [\rho_{ij}]$ such that $\rho_{ii} = 1$ and $|\rho_{ij}| < 1$, $i \neq j$, $i, j = 1, ..., N$. This
implies $P_t^{-1/2} \zeta_t \sim iid(0, I_N)$. Under these assumptions, the error vector $\varepsilon_t$ satisfies the following
moments conditions:

\[
E(\varepsilon_t | \mathcal{F}_{t-1}) = 0
\]

\[
E(\varepsilon_t \varepsilon'_t | \mathcal{F}_{t-1}) = \Sigma_t = D_t P_t D'_t
\]

(3)

where the conditional covariance matrix \( \Sigma_t = [\sigma_{ijt}] \) of \( \varepsilon_t \) given the information set \( \mathcal{F}_{t-1} \) is a positive-definite \( N \times N \) matrix. It is now assumed that \( D_t \) consists of a conditionally heteroskedastic component and a deterministic time-dependent one such that

\[
D_t = S_t G_t
\]

(4)

where \( S_t = \text{diag}(h_{1t}^{1/2}, \ldots, h_{Nt}^{1/2}) \) contains the conditional standard deviations \( h_{it}^{1/2}, i = 1, \ldots, N \), and \( G_t = \text{diag}(g_{1t}^{1/2}, \ldots, g_{Nt}^{1/2}) \). The elements \( g_{it}, i = 1, \ldots, N \), 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

\[
\Sigma_t = S_t G_t P_t G_t S_t.
\]

(5)

It follows that

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

(6)

and that

\[
\sigma_{iit} = h_{it} g_{it}, \ i = 1, \ldots, N.
\]

(7)

From (7) it follows that \( h_{it} = \sigma_{iit} / g_{it} = E(\varepsilon_{it}^* \varepsilon_{it}' | \mathcal{F}_{t-1}) \), where \( \varepsilon_{it}^* = \varepsilon_{it} / g_{it}^{1/2} \). When \( G_t \equiv I_N \) and the conditional correlation matrix \( P_t \equiv P \), one obtains the Constant Conditional Correlation (CCC-) GARCH model of Bollerslev (1990). More generally, when \( G_t \equiv I_N \) and \( P_t \) is a time-varying correlation matrix, the model belongs to the family of Conditional Correlation GARCH models.

Following Amado and Teräsvirta (2011), the diagonal elements of the matrix \( G_t \) are defined as follows:

\[
g_{it} = 1 + \sum_{i=1}^r \delta_{it} G_{it}(t/T; \gamma_{it}, c_{it})
\]

(8)
where $\gamma_{il} > 0$, $i = 1, \ldots, N$, $l = 1, \ldots, r$, and $r = 0, 1, 2, \ldots, R$, such that $R$ is a finite integer. Each $g_{it}$ varies smoothly over time satisfying the conditions $\inf_{t=1,\ldots,T} g_{it} > 0$, and $\delta_{il} \leq M_{il} < \infty$, $l = 1, \ldots, r$, for $i = 1, \ldots, N$. For identification reasons, in (8) $\delta_{i1} < \ldots < \delta_{ir}$ and $\delta_{il} \neq 0$ for all $i$ and $l$. The function $G_{il}(t/T; \gamma_{il}, c_{il})$ is a generalized logistic function, that is,

$$G_{il}(t/T; \gamma_{il}, c_{il}) = \left(1 + \exp \left\{-\gamma_{il} \prod_{j=1}^{k}(t/T - c_{il}) \right\}\right)^{-1}, \quad \gamma_{il} > 0, \quad c_{il1} \leq \ldots \leq c_{ilk}. \quad (9)$$

Function (9) is by construction continuous for $\gamma_{il} < \infty$, $i = 1, \ldots, r$, and bounded between zero and one. The parameters, $c_{ilj}$ and $\gamma_{il}$ 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 $\delta_{i1} = \ldots = \delta_{ir} = 0$ or $\gamma_{i1} = \ldots = \gamma_{ir} = 0$, $i = 1, \ldots, N$, in (8), the unconditional variance of $\varepsilon_{t}$ becomes constant, otherwise it is time-varying. Assuming either $r > 1$ or $k > 1$ 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 $\gamma_{i1}$ in (9) controls the degree of smoothness of the transition: the larger $\gamma_{i1}$, the faster the transition between the extreme regimes. As $\gamma_{i1} \to \infty$, $g_{it}$ approaches a step function with a switch at $c_{i11}$. For small values of $\gamma_{i1}$, 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 $G_t \neq I_N$. As to modelling the time-variation in the correlation matrix $P_t$, several choices exist. As already mentioned, the simplest multivariate correlation model is the CCC-GARCH model of Bollerslev (1990) in which $P_t \equiv P$. 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 $\rho_{ijt}$ is a typical element of the matrix $P_t$ with the dynamic structure

$$P_t = \{\text{diag}Q_t\}^{-1/2}Q_t\{\text{diag}Q_t\}^{-1/2}$$

(11)

where

$$Q_t = (1 - \theta_1 - \theta_2)\bar{Q} + \theta_1 \zeta_{t-1}^t\zeta_{t-1}' + \theta_2 Q_{t-1}$$

(12)

such that $\theta_1 > 0$ and $\theta_2 \geq 0$ with $\theta_1 + \theta_2 < 1$, $\bar{Q}$ is the unconditional correlation matrix of the standardised errors $\zeta_{it}$, $i = 1, ..., N$, and $\zeta_t = (\zeta_{1t}, ..., \zeta_{Nt})'$. 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), $P_t$ has a definition that is slightly different from (12). More specifically,

$$P_t = (1 - \theta_1 - \theta_2)P + \theta_1 S_{t-1} + \theta_2 P_{t-1}$$

(13)

where $P$ is a constant positive definite parameter matrix with unit diagonal elements, $\theta_1$ and $\theta_2$...
are non-negative parameters such that $\theta_1 + \theta_2 \leq 1$, and $S_{t-1}$ is a matrix whose elements are functions of the lagged standardised residuals. The positive definiteness of $P_t$ is ensured if $P_0$ and $S_{t-1}$ are positive definite matrices. In our application, when $\zeta_{it}$ 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.

Another way of introducing time-varying correlations is to assume that the correlation matrix $P_t$ varies smoothly over time between two extreme states of correlations $P_{(1)}$ and $P_{(2)}$; see Berben and Jansen (2005) and Silvennoinen and Teräsvirta (2009a, in press). More specifically,

$$P_t = \{1 - G(s_t; \gamma, c)\}P_{(1)} + G(s_t; \gamma, c)P_{(2)}$$ (14)

where $P_{(1)}$ and $P_{(2)}$ are positive definite $N \times N$ matrices with ones on the main diagonal and $P_{(1)} \neq P_{(2)}$. $G(s_t; \gamma, c)$ is a monotonic function bounded between zero and one, in which the stochastic or deterministic transition variable $s_t$ controls the correlations. It is defined as follows:

$$G(s_t; \gamma, c) = (1 + \exp\{-\gamma(s_t - c)\})^{-1}, \quad \gamma > 0$$ (15)

where, as in (9), the parameter $\gamma$ determines the smoothness and $c$ the location of the transition between the two extreme correlation regimes. In this work $s_t = t/T$, where $T$ is the number of observations. We call the resulting model the Time-Varying Correlation-TVGJR-GARCH (TVC-TVGJR-GARCH) model when the equations for $h_{it}$ are parameterised as in (10). When $g_{it} \equiv 1$, (14) reduces to the covariance matrix of the TVC-GJR-GARCH model. This model differs from the DCC-type model in the sense that the TVC-type model is a model of unconditional correlations while the former is a model of conditional correlations. Indeed, the covariances in the TVC-GJR-GARCH model are unconditional as they are independent of the past returns.

### 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

\[ E_t \Sigma_{t+1} = S_{t+1|t} G_{t+1|t} P_{t+1|t} G_{t+1|t} S_{t+1|t}. \]

where \( S_{t+1|t} \) 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 \( G_{t+1|t} \) are constructed under the assumption that \( g_{i,t+1|t} = g_{i,t} \), for all \( i = 1, \ldots, N \). 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-TVGSJR-GARCH model, the correlation forecast \( P_{t+1|t} \) is the standardized version of \( Q_{t+1|t} \), whose one-step-ahead forecast is obtained by projecting (12) one step into the future. In this scheme, the standardized returns are \( E_t \zeta_{i,t+1} = \varepsilon_{i,t+1|t}/(h_{i,t+1|t}g_{i,t+1|t})^{1/2} \), and

\[ Q_{t+r|t} = (1 - \theta_1 - \theta_2) \bar{Q} + (\theta_1 + \theta_2) Q_{t+r-1|t} \]

for \( r > 1 \). We construct one-step-ahead forecasts for the VC-TVGSJR-GARCH model in an identical fashion.

3 Estimation of parameters

3.1 Estimation of DCC-TVGSJR-GARCH models

In this section we assume that \( \omega_i = 1, i = 1, \ldots, N \), in (10) and that (8) has the form

\[ g_{i,t} = \delta_{i0} + \sum_{l=1}^{r} \delta_{il} G_{il}(t/T; \gamma_{il}, c_{il}) \]

where \( \delta_{i0} > 0 \). This facilitates the notation but does not change the gist of the argument. Under the assumption of normality, \( \varepsilon_t|\mathcal{F}_{t-1} \sim N(0, \Sigma_t) \), the conditional log-likelihood function for
observation $t$ is defined as

$$
\ell_t(\theta) = -(N/2) \ln(2\pi) - (1/2) \ln |\Sigma_t| - (1/2) \varepsilon_t^\prime \Sigma_t^{-1} \varepsilon_t
$$

$$
= -(N/2) \ln(2\pi) - (1/2) \ln |S_tG_tP_tG_tS_t| - (1/2) \varepsilon_t^\prime S_t^{-1}G_t^{-1}P_t^{-1}G_t^{-1}S_t^{-1} \varepsilon_t
$$

$$
= -(N/2) \ln(2\pi) - \ln |S_tG_t| - (1/2) \ln |P_t| - (1/2) \zeta_t^\prime P_t^{-1} \zeta_t
$$

$$
= -(N/2) \ln(2\pi) - \ln |G_t| - (1/2) \tilde{\varepsilon}_t^\prime G_t^{-2} \tilde{\varepsilon}_t - \ln |S_t| - (1/2) \varepsilon_t^\prime S_t^{-2} \varepsilon_t
$$

$$
+ \zeta_t^\prime \zeta_t - (1/2) \ln |P_t| - (1/2) \zeta_t^\prime P_t^{-1} \zeta_t
$$

(16)

where

$$
\tilde{\varepsilon}_t = S_t^{-1} \varepsilon_t = (\varepsilon_{1t}/\{h_{1t}(\psi_1, \varphi_1)\}^{1/2}, ..., \varepsilon_{Nt}/\{h_{Nt}(\psi_N, \varphi_N)\}^{1/2})'
$$

$$
\varepsilon_t^* = G_t^{-1} \varepsilon_t = (\varepsilon_{1t}/\{g_{1t}(\psi_1)\}^{1/2}, ..., \varepsilon_{Nt}/\{g_{Nt}(\psi_N)\}^{1/2})'
$$

$$
\zeta_t = G_t^{-1} S_t^{-1} \varepsilon_t = (\varepsilon_{1t}/\{g_{1t}(\psi_1)h_{1t}(\psi_1, \varphi_1)\}^{1/2}, ..., \varepsilon_{Nt}/\{g_{Nt}(\psi_N)h_{Nt}(\psi_N, \varphi_N)\}^{1/2})'.
$$

The parameter vector $\theta$ is partitioned into $\theta = (\psi', \varphi', \phi')'$, where $\psi = (\psi_1', ..., \psi_N')'$ is the subvector of parameters of the unconditional variances, $\varphi = (\varphi_1', ..., \varphi_N')'$ is the subvector of parameters of the conditional variances, and $\phi = (\theta_1, \theta_2)'$ contains the parameters of the correlation matrix. Furthermore, $\psi_i = (\delta_i, \gamma_i, c_i)'$, $\delta_i = (\delta_{i0}, ..., \delta_{ip})'$, $\gamma_i = (\gamma_{i1}, ..., \gamma_{ip})'$, $c_i = (c_{i1}', ..., c_{ip}')'$, and $\varphi_i = (\alpha_{i1}, ..., \alpha_{i\mu}, \kappa_{i1}, ..., \kappa_{i\mu}, \beta_{i1}, ..., \beta_{i\mu})'$, $i = 1, ..., N$.

Equation (16) implies the following decomposition of the log-likelihood function for observation $t$:

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

where first,

$$
\ell_t^U(\psi) = \sum_{i=1}^N \ell_{it}^U(\psi_i)
$$

(17)

and

$$
\ell_{it}^U(\psi_i) = -(1/2) \{\ln g_{it}(\psi_i) + \tilde{\varepsilon}_{it}^2/g_{it}(\psi_i)\}.
$$

Second,

$$
\ell_t^V(\psi, \varphi) = \sum_{i=1}^N \ell_{it}^V(\psi_i, \varphi_i)
$$

(18)
Finally,

\[ \ell_t^C(\psi, \varphi, \phi) = -(1/2)\{\ln |P_t(\psi, \varphi, \phi)| + \zeta_t'P_t^{-1}(\psi, \varphi, \phi)\zeta_t - 2\zeta_t'\zeta_t \}. \]  

The GARCH equations are estimated separately using maximization by parts. The first iteration consists of the following:

1. Reparameterise the deterministic component (8) as follows:

   \[ g_{it}^* = \delta_{i0}^* + \sum_{l=1}^{r} \delta_{il}^*G_{il}(t/T; \gamma_{il}, c_{il}). \]

   and set \( \psi_i^* = (\delta_{i0}^*, \delta_{i1}^*, \gamma_i, c_i)' \), where \( \delta_{i0}^* > 0 \) and \( \delta_i^* = (\delta_{i1}^*,...,\delta_{ir}^*)' \) with \( \delta_i^* = \delta_{i0}^* \delta_i \), so \( g_{i0}^* = \delta_{i0}^* g_{i0} \). Maximize

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

   for each \( i, i = 1, ..., N \), separately, assuming \( \tilde{\varepsilon}_{it} = \varepsilon_{it} \), that is, setting \( h_{it}(\psi_i, \varphi_i) = 1 \). The resulting estimators are \( \hat{\psi}_i^* = (\hat{\delta}_{i0}^{(1)}, \hat{\delta}_i^{(1)'}, \hat{\gamma}_i^{(1)'}, \hat{c}_i^{(1)'})' \), \( i = 1, ..., N \). Obtain \( \hat{\delta}_i^{(1)} \) as follows:

   \[ \hat{\delta}_i^{(1)} = (\hat{\delta}_{i0}^{(1)})^{-1}\hat{\delta}_i^{(1)} \]

   so that \( \hat{\psi}_i^{(1)} = (\hat{\delta}_i^{(1)'}, \hat{\gamma}_i^{(1)'}, \hat{c}_i^{(1)'})' \). Note that \( \hat{\delta}_{i0}^{(1)} = \hat{\omega}_i^{(1)} \).

2. Setting \( \psi_i = \hat{\psi}_i^{(1)} \), \( i = 1, ..., N \), in (18), maximize

   \[ L_{iT}^V(\hat{\psi}_i^{(1)}, \varphi_i) = \sum_{t=1}^{T} \ell_{it}^V(\hat{\psi}_i^{(1)}, \varphi_i) = -(1/2) \sum_{t=1}^{T} \{\ln h_{it}(\hat{\psi}_i^{(1)}, \varphi_i) + \varepsilon_{it}^2/h_{it}(\hat{\psi}_i^{(1)}, \varphi_i)\} \]

   with respect to \( \varphi_i \) assuming \( \varepsilon_{it} = \varepsilon_{it}/g_{it}^{1/2}(\hat{\psi}_i^{(1)}) \), for each \( i, i = 1, ..., N \), separately. Call the \( \varphi_i \) resulting estimators \( \hat{\varphi}_i^{(1)} \).

The second iteration is as follows:

1. Maximize

   \[ L_{iT}^U(\psi) = \sum_{t=1}^{T} \ell_{it}^U(\psi) = -(1/2) \sum_{t=1}^{T} \{\ln g_{it}(\psi_i) + \tilde{\varepsilon}_{it}^2/g_{it}(\psi_i)\} \]
assuming \( \tilde{e}_{it} = e_{it}/h_{it}^{1/2}(\hat{\psi}_i^{(1)}, \hat{\phi}_i^{(1)}) \), for each \( i, i = 1, ..., N \). Call the \( i \)th resulting estimator \( \hat{\psi}_i^{(2)} \). The important thing is that \( \phi_i = \hat{\phi}_i^{(1)} \) (fixed) in the definition of \( \tilde{e}_{it} \).

2. Maximize

\[
L_{VT}^{V}(\hat{\psi}_i^{(2)}, \phi_i) = \sum_{t=1}^{T} \ell_{it}^{V}(\hat{\psi}_i^{(2)}, \phi_i) = -(1/2) \sum_{t=1}^{T} \{ \ln h_{it}(\hat{\psi}_i^{(2)}, \phi_i) + \epsilon_{it}^{*2}/h_{it}(\hat{\psi}_i^{(2)}, \phi_i) \}
\]

with respect to \( \phi_i \) for each \( i, i = 1, ..., N \), separately, assuming \( \epsilon_{it}^{*} = \epsilon_{it}/g_{it}(\hat{\psi}_i^{(2)}) \). This yields \( \hat{\phi}_i^{(2)}, i = 1, ..., N \).

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

Maximization is carried out in the usual fashion by solving the equations

\[
\frac{\partial}{\partial \psi_i} L_{VT}^{V}(\psi_i) = (1/2) \sum_{t=1}^{T} \left( \frac{\epsilon_{it}^{2}}{g_{it}(\psi_i)} - 1 \right) \frac{1}{g_{it}(\psi_i)} \frac{\partial g_{it}(\psi_i)}{\partial \psi_i} = 0
\]

for \( \psi_i \)

\[
\frac{\partial}{\partial \phi_i} L_{VT}^{V}(\phi_i) = (1/2) \sum_{t=1}^{T} \left( \frac{\epsilon_{it}^{2}}{h_{it}(\hat{\psi}_i^{(n)}, \phi_i)} - 1 \right) \frac{1}{h_{it}(\hat{\psi}_i^{(n)}, \phi_i)} \frac{\partial h_{it}(\hat{\psi}_i^{(n)}, \phi_i)}{\partial \phi_i} = 0
\]

for \( \phi_i \) in the \( n \)th iteration. Writing \( G_{ilt} = G(t^*, \gamma_{il}, c_{il}) \), we have

\[
\frac{\partial g_{it}(\psi_i)}{\partial \psi_i} = (1, G_{ilt}^{(\gamma)}, G_{ilt}^{(c)}, ..., G_{irt}^{(\gamma)}, G_{irt}^{(c)})'
\]

where, for \( k = 1 \) in (9),

\[
G_{ilt}^{(\gamma)} = \frac{\partial G_{ilt}}{\partial \gamma_{il}} = \delta_{il} G_{ilt} (1 - G_{ilt})(t^* - c_{il}), \ l = 1, ..., r
\]

\[
G_{ilt}^{(c)} = \frac{\partial G_{ilt}}{\partial c_{il}} = -\gamma_{il} \delta_{il} G_{ilt} (1 - G_{ilt}), \ l = 1, ..., r
\]
and

\[
\frac{\partial h_{it}(\hat{\psi}_i^{(n)}, \varphi_i)}{\partial \varphi_i} = (1, \varepsilon_{i,t-1}^2, ..., \varepsilon_{i,t-q}^2 \varepsilon_{i,t-1}^2 I(\varepsilon_{i,t-1}^* < 0), ..., \varepsilon_{i,t-q}^2 I(\varepsilon_{i,t-q}^* < 0),
\]

\[
h_{i,t-1}(\hat{\psi}_i^{(n)}, \varphi_i), ..., h_{i,t-p}(\hat{\psi}_i^{(n)}, \varphi_i))' + \sum_{j=1}^{p} \beta_{ij} \frac{\partial h_{i,t-j}(\hat{\psi}_i^{(n)}, \varphi_i)}{\partial \varphi_i}
\]

when the conditional variance \( h_{it} \) is defined in (10).

After estimating the TVGARCH equations, estimate \( \phi \) given \( \hat{\psi}_i \) and \( \hat{\varphi}_i \) by maximizing

\[
L_C^T(\phi) = -\frac{1}{2} \sum_{t=1}^{T} \left\{ \ln |P_t(\phi)| + \zeta_t' P_t^{-1}(\phi) \zeta_t - 2 \zeta_t' \zeta_t \right\}
\]

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

\[
\frac{\partial}{\partial \phi} L_C^T(\phi) = -\frac{1}{2} \sum_{t=1}^{T} \frac{\partial \text{vec}(P_t)}{\partial \phi}' \text{vec}(P_t^{-1} - P_t^{-1} \zeta_t \zeta_t' P_t^{-1}).
\]

All computations in this paper have been performed using the Ox programming language, version 6.10, see Doornik (2009), the OxMetrics module G@RCH 6.1, and a modified version of Matteo Pelagatti’s source code\(^1\).

This approach is computationally feasible. Engle and Sheppard (2001) only estimate the GARCH equations once and show that for \( G_t = I_N \), the maximum likelihood estimators \( \hat{\varphi}_i, i = 1, ..., N, \) (in their framework \( g_{it}(\psi_i) \equiv 1 \)) are consistent. The two-step estimator is, however, asymptotically less efficient than the full maximum likelihood estimator. Further iteration in order to obtain efficient estimators is possible, see Fan, Pastorello and Renault (2007) for discussion, but it has not been undertaken here. Under regularity conditions, the maximum likelihood estimators of the TVGJR-GARCH equations are consistent and asymptotically normal; see Amado and Teräsvirta (2011).

\(^1\)The Ox estimation package is freely available at [http://www.statistica.unimib.it/utenti/p_matteo/Research/research.html](http://www.statistica.unimib.it/utenti/p_matteo/Research/research.html)
3.2 Estimation of TVC-TVGJR-GARCH models

The maximum likelihood estimation of the parameters of the model TVC-GJR-GARCH model can be carried out in three steps as in Silvennoinen and Teräsvirta (2009a, in press). The log-likelihood function can be decomposed as before. The components (17) and (18) remain the same, whereas (19) becomes

$$\ell_C^C(\psi, \varphi, \xi) = -(1/2)\{\ln |P_t(\xi)| + \zeta_t P_t^{-1}(\xi)\zeta_t - 2\zeta'_t\zeta_t\}$$

where the \(\{N(N - 1) + 2\} \times 1\) vector \(\xi = (\text{vecl}(P_{(1)}), \text{vecl}(P_{(2)}), \gamma, c)'\). (The \text{vecl} operator stacks the columns below the main diagonal into a vector.) In their scheme, the parameter vectors \(\psi\) and \(\varphi\) of the GARCH equations are estimated first, followed by the correlations in \(P_{(1)}\) and \(P_{(2)}\), given the transition function parameters \(\gamma\) and \(c\) in (15). Finally, \(\gamma\) and \(c\) are estimated given \(\psi\), \(\varphi\), \(P_{(1)}\) and \(P_{(2)}\). The next iteration begins by re-estimating \(\varphi\) given the previous estimates of \(P_{(1)}\), \(P_{(2)}\), \(\gamma\) and \(c\). The only modification required for the estimation of TVC-TVGJR-GARCH models compared to Silvennoinen and Teräsvirta (in press) is that for each main iteration there is an inside loop for iterative estimation (maximisation by parts) of \(\psi\) and \(\varphi\). In practice, compared to the two-step estimates, the extra iterations do not change the estimates much, but the estimators become fully efficient.

Asymptotic properties of the maximum likelihood estimators of the TVC-TVGJR-GARCH model are not yet known. The existing results only cover the CCC-GARCH model; see Ling and McAleer (2003). Due to a time-varying correlation matrix, deriving corresponding asymptotic results for the TVC-TVGJR-GARCH model is a nontrivial problem and beyond the scope of the present paper. Note that asymptotic normality has been proven for maximum likelihood estimates of the parameters of the TVGJR-GARCH model: our univariate GARCH components are of this form.

4 Specifying the unconditional variance component

In applying a model belonging to the family of CC-TVGJR-GARCH models, 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. We shall concentrate on the first set of issues. It appears that in applications involving DCC-GARCH models the null hypothesis of constant conditional correlations is never tested, and we shall adhere to that practice. In applications of the STCC-GARCH model, constancy of correlations is always tested before applying the larger model, see Silvennoinen and Teräsvirta (2009a, in press). The test can be extended to the current situation in which the GARCH equations are TVGJR-GARCH ones instead of plain GJR-GARCH ones. Nevertheless, in this work we assume that the correlations do vary over time as is done in the context of DCC-GARCH models and apply the TVC-GARCH model without carrying out a correlation constancy test.

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 (2011) and briefly review their procedure. The conditional variances are estimated first, assuming $g_{it} \equiv 1, \ i = 1, \ldots, N$. 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 $H_{11}: \gamma_{t1} = 0$ against $H_{11}: \gamma_{t1} > 0$ in

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

The standard test statistic has a non-standard asymptotic distribution because $\delta_{i1}$ and $c_{i1}$ 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 $G_{i1}(t/T; \gamma_{i1}, c_{i1})$ is replaced by its $m$th-order Taylor expansion around $\gamma_{i1} = 0$. Choosing $m = 3$, this yields

$$g_{it} = \alpha_{0}^{*} + \sum_{j=1}^{3} \delta_{ij}^{*}(t/T)^{j} + R_{3}(t/T; \gamma_{i1}, c_{i1})$$

(20)

where $\delta_{ij}^{*} = \gamma_{i1}^{j} \tilde{\delta}_{ij}$ with $\tilde{\delta}_{ij} \neq 0$, and $R_{3}(t/T; \gamma_{i1}, c_{i1})$ is the remainder. The new null hypothesis based on this approximation is $H'_{01}: \delta_{i1}^{*} = \delta_{i2}^{*} = \delta_{i3}^{*} = 0$ in (20). In order to test this null hypothesis, we use the Lagrange multiplier (LM) test. Furthermore, $R_{3}(t/T; \gamma_{i1}, c_{i1}) \equiv 0$ under $H_{01}$, so the asymptotic distribution theory is not affected by the remainder. As discussed in Amado and Teräsvirta (2011), the LM-type test statistic has an asymptotic $\chi^{2}$-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 $G_{il}(t/T; \gamma_{il}, c_{il})$. It is solved by carrying out a short test sequence within (20); for details see Amado and Teräsvirta (2011). The next step is then to estimate the alternative with the chosen $k$, add another transition, and test the hypothesis $\gamma_{i2} = 0$ in

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

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 $\chi^2$-distribution with three degrees of freedom under the null hypothesis.

The model-building cycle for TVGJR-GARCH models for the elements of $D_t = S_t G_t$ 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 proposed by Amado and Teräsvirta (2011).

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. To perform the out-of-sample analysis we use the next 311 observations for each return series from October 8,
2008, to December 31, 2009. All stock prices are converted into continuously compounded rates of return, whose values are plotted in Figure 1.

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

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 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
Table 1: Descriptive statistics of the asset returns

<table>
<thead>
<tr>
<th>Asset</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std.dev.</th>
<th>Skew</th>
<th>Ex.Kurt</th>
<th>Rob.Sk.</th>
<th>Rob.Kr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>-19.35</td>
<td>13.23</td>
<td>0.011</td>
<td>2.280</td>
<td>-0.265</td>
<td>4.864</td>
<td>0.004</td>
<td>0.418</td>
</tr>
<tr>
<td>BA</td>
<td>-19.39</td>
<td>9.513</td>
<td>0.019</td>
<td>2.069</td>
<td>-0.611</td>
<td>7.215</td>
<td>-0.003</td>
<td>0.106</td>
</tr>
<tr>
<td>CAT</td>
<td>-11.51</td>
<td>9.067</td>
<td>0.037</td>
<td>2.073</td>
<td>-0.260</td>
<td>3.945</td>
<td>-0.022</td>
<td>0.108</td>
</tr>
<tr>
<td>INTC</td>
<td>-24.87</td>
<td>18.32</td>
<td>-0.009</td>
<td>2.896</td>
<td>-0.470</td>
<td>6.197</td>
<td>-0.001</td>
<td>0.160</td>
</tr>
<tr>
<td>JPM</td>
<td>-19.97</td>
<td>15.47</td>
<td>0.025</td>
<td>2.524</td>
<td>0.282</td>
<td>6.901</td>
<td>-0.010</td>
<td>0.396</td>
</tr>
<tr>
<td>WHR</td>
<td>-13.30</td>
<td>12.95</td>
<td>0.022</td>
<td>2.254</td>
<td>0.183</td>
<td>3.516</td>
<td>0.004</td>
<td>0.272</td>
</tr>
<tr>
<td>XOM</td>
<td>-8.83</td>
<td>9.29</td>
<td>0.039</td>
<td>1.579</td>
<td>-0.136</td>
<td>2.334</td>
<td>-0.058</td>
<td>0.060</td>
</tr>
</tbody>
</table>

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 (see Kim and White (2004)) and the Rob.Kr. denotes the robust centred coefficient for kurtosis proposed by Moors (see Kim and White (2004)).

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.

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

<table>
<thead>
<tr>
<th>Transitions</th>
<th>$H_0$</th>
<th>$H_{03}$</th>
<th>$H_{02}$</th>
<th>$H_{01}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single transition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AXP</td>
<td>0.0184</td>
<td>0.1177</td>
<td>0.0071</td>
<td>0.5722</td>
</tr>
<tr>
<td>BA</td>
<td>0.0021</td>
<td>0.0616</td>
<td>0.0461</td>
<td>0.0072</td>
</tr>
<tr>
<td>CAT</td>
<td>0.0044</td>
<td>0.0260</td>
<td>0.0107</td>
<td>0.1971</td>
</tr>
<tr>
<td>INTC</td>
<td>$5 \times 10^{-5}$</td>
<td>$9 \times 10^{-5}$</td>
<td>0.1600</td>
<td>0.0197</td>
</tr>
<tr>
<td>JPM</td>
<td>$9 \times 10^{-4}$</td>
<td>0.0073</td>
<td>0.0023</td>
<td>0.8500</td>
</tr>
<tr>
<td>WHR</td>
<td>$6 \times 10^{-5}$</td>
<td>$7 \times 10^{-4}$</td>
<td>0.0011</td>
<td>0.9401</td>
</tr>
<tr>
<td>XOM</td>
<td>0.0018</td>
<td>0.0836</td>
<td>$7 \times 10^{-4}$</td>
<td>0.4271</td>
</tr>
<tr>
<td>Double transition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AXP</td>
<td>0.0826</td>
<td>0.1953</td>
<td>0.0378</td>
<td>0.4032</td>
</tr>
<tr>
<td>BA</td>
<td>0.1208</td>
<td>0.1480</td>
<td>0.0547</td>
<td>0.8419</td>
</tr>
<tr>
<td>CAT</td>
<td>0.4011</td>
<td>0.1719</td>
<td>0.4961</td>
<td>0.4347</td>
</tr>
<tr>
<td>INTC</td>
<td>0.4307</td>
<td>0.8757</td>
<td>0.1050</td>
<td>0.7458</td>
</tr>
<tr>
<td>JPM</td>
<td>0.0947</td>
<td>0.0144</td>
<td>0.8678</td>
<td>0.5484</td>
</tr>
<tr>
<td>WHR</td>
<td>0.3059</td>
<td>0.8856</td>
<td>0.1450</td>
<td>0.2249</td>
</tr>
<tr>
<td>XOM</td>
<td>0.1111</td>
<td>0.1526</td>
<td>0.4198</td>
<td>0.0685</td>
</tr>
</tbody>
</table>

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 $H_{02}$, then choose $k = 2$. If either $H_{01}$ or $H_{03}$ are rejected more strongly than $H_{02}$, then select either $k = 1$ or $k = 3$. See Amado and Teräsvirta (2011) for further details.

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

<table>
<thead>
<tr>
<th>Asset</th>
<th>$\hat{\gamma}_1$</th>
<th>$\hat{c}_{11}$</th>
<th>$\hat{c}_{12}$</th>
<th>$\hat{c}_{13}$</th>
<th>$\delta_t$ component</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>4.3601</td>
<td>0.4825</td>
<td>0.9034</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>BA</td>
<td>0.651</td>
<td>0.4686</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>CAT</td>
<td>1.2366</td>
<td>0.3021</td>
<td>0.9726</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>INTC</td>
<td>2.9973</td>
<td>0.0262</td>
<td>0.4775</td>
<td>0.9127</td>
<td>1</td>
</tr>
<tr>
<td>JPM</td>
<td>6.3688</td>
<td>0.4821</td>
<td>0.9042</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>WHR</td>
<td>1.2272</td>
<td>0.0892</td>
<td>0.4195</td>
<td>0.8497</td>
<td>1</td>
</tr>
<tr>
<td>XOM</td>
<td>1.1063</td>
<td>0.4106</td>
<td>0.8672</td>
<td>–</td>
<td>1</td>
</tr>
</tbody>
</table>

Notes: The table contains the parameter estimates of the $g_t$ 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 $g_t = 1 + \sum_{i=1}^{\infty} \delta_i G_i(t; \gamma_{it}, \kappa_{it})$, where $G_i(t; \gamma_{it}, \kappa_{it})$ 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

<table>
<thead>
<tr>
<th>Asset</th>
<th>$\hat{\omega}$</th>
<th>$\hat{\alpha}_1$</th>
<th>$\hat{\kappa}_1$</th>
<th>$\hat{\beta}_1$</th>
<th>$\hat{\alpha}_1 + \frac{\hat{\omega}}{2} + \hat{\beta}_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>0.4777</td>
<td>–</td>
<td>0.1309</td>
<td>0.9045</td>
<td>0.9699</td>
</tr>
<tr>
<td>BA</td>
<td>0.1050</td>
<td>–</td>
<td>0.0899</td>
<td>0.9103</td>
<td>0.9552</td>
</tr>
<tr>
<td>CAT</td>
<td>0.6641</td>
<td>0.0477</td>
<td>–</td>
<td>0.7340</td>
<td>0.7817</td>
</tr>
<tr>
<td>INTC</td>
<td>0.1203</td>
<td>0.0450</td>
<td>–</td>
<td>0.9155</td>
<td>0.6065</td>
</tr>
<tr>
<td>JPM</td>
<td>0.0474</td>
<td>0.0213</td>
<td>0.1135</td>
<td>0.8890</td>
<td>0.9670</td>
</tr>
<tr>
<td>WHR</td>
<td>0.3569</td>
<td>0.0736</td>
<td>–</td>
<td>0.8141</td>
<td>0.8877</td>
</tr>
<tr>
<td>XOM</td>
<td>0.0644</td>
<td>0.0272</td>
<td>0.0578</td>
<td>0.9008</td>
<td>0.9568</td>
</tr>
</tbody>
</table>

Notes: The table contains the parameter estimates of the $h_t$ 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 $h_t = \omega + \alpha_1 \hat{\varepsilon}_{it-1}^2 + \kappa_1 I_{it-1}(\hat{\varepsilon}_{it-1}^2)\hat{\varepsilon}_{it-1}^2 + \beta_1 h_{it-1}$, where $\hat{\varepsilon}_{it} = \hat{\varepsilon}_{it}/g_{it}^{1/2}$ and $I_d(\hat{\varepsilon}_{it}^2) = 1$ if $\hat{\varepsilon}_{it}^2 < 0$ (and 0 otherwise) for all $i$. The numbers in parentheses are the Bollerslev-Wooldridge robust standard errors.
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 $G_t \equiv I_2$, whereas the other one is the TVGJR-GARCH model for which $G_t \neq I_2$ in (5).

We begin by comparing the rolling correlation estimates for the $(\varepsilon_{it}, \varepsilon_{jt})$ and $(\varepsilon_{it}/\tilde{g}_{it}^{1/2}, \varepsilon_{jt}/\tilde{g}_{jt}^{1/2})$ pairs. Figure 5 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 suggest that there are also differences in conditional correlations between models based on GJR-GARCH type variances and their TVGJR-GARCH counterparts. A look at Figure 6 suggests, perhaps surprisingly, that when one compares DCC-GJR-GARCH models with DCC-TVGJR-GARCH ones, this is not the case. The figure
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 $\frac{\varepsilon_t}{\hat{g}_t^{1/2}}$ for the seven stock returns of the S&P 500 composite index.
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 on the correlations lies within the interval \((-0.22, 0.30)\). 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.

Figure 7 shows the estimated time-varying correlations for the bivariate TVC-GJR-GARCH and TVC-TVGJR-GARCH models. The parameter estimates are omitted to conserve space. For the majority of the estimated models, the estimate of the slope transition parameter \(\gamma\) attains its upper bound of 500. For these cases, the transition function is close to a step function. The differences in correlations have to do with the smoothness of the increase in correlations during the first quarter of the observations. These differences are not systematic: in some cases the increase is smoother in the former model, in others in the latter. In a few cases, the differences are very small. The main conclusion from these comparisons would be rather similar to that obtained from considering DCC-GARCH models. We also used other transition variables than time, such as a linear combination of past returns of the paired assets, but the fit of the model was inferior to that obtained when the transition variable was time.

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. A comparison between DCC- and VC-GJR-GARCH models suggests that the latter one fits the data better than the former in the 7-variate case.
Figure 5: Difference between the estimated rolling correlation coefficients for pairs of the raw returns (grey solid curve) and pairs of the standardised returns (red solid curve).
Figure 6: Difference between the estimated conditional correlations obtained from the bivariate DCC-GJR-GARCH and the bivariate DCC-TVGJR-GARCH models for the asset returns.
Figure 7: The estimated conditional correlations obtained from the bivariate TVC-GJR-GARCH (solid curve) and the bivariate TVC-TV-GJR-GARCH (dotted curve) models for the asset returns.
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 the parameters using a fixed window of 2521 daily observations. Specifically, the first set of one-step-ahead covariance forecasts are based on the estimation period from September 29, 1998 until October 7, 2008. To generate the next set of covariance forecasts, the window is then rolled forward one day to obtain the second set of daily covariance forecasts. We repeat this process by adding the next observation and dropping the earliest return until we reach the end of the out-of-sample period. After performing this procedure, we computed 311 one-step-ahead covariance forecasts based on the estimation of 2521 returns each. The overall out-of-sample period ranges from October 8, 2008 to December 31, 2009.

The evaluation consists of comparing the conditional covariance matrix with the true matrix. Since the true conditional covariance matrix is unobserved, following Pelletier (2006) we use a proxy based on the cross-product of the daily returns over the forecast horizon. All computations are based on one-day-ahead forecasts of the covariance matrix over 311 days. In Figure 8 we plot the differences between the estimated conditional 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. The graphs show a systematic pattern in the differences between these correlations. 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 use the following criteria (Pelletier, 2006):

\[
\text{RMSE}_1 = \left\{ \frac{1}{N^2} \sum_{i,j} E(\Sigma_{i,j,t+1|t} - y_{i,t+1}y_{j,t+1})^2 \right\}^{1/2}
\]

and

\[
\text{MAD}_1 = \frac{1}{N^2} \sum_{i,j} E|\Sigma_{i,j,t+1|t} - y_{i,t+1}y_{j,t+1}|
\]
that are multivariate versions of the root mean squared error (RMSE) and mean absolute deviation (MAD), respectively, and $\Sigma_{i,j,t+1|t}$ is the one-step-ahead forecast of the covariance between returns $y_{it}$ and $y_{jt}$. Criteria based on the absolute deviations are sometimes preferred because they are less affected by outliers than RMSE. To further reduce the impact of outlying observations on forecasting evaluation, we also consider the Median Squared Error (MedSE) criteria. 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-TVJR-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 models with constant unconditional variances are slightly superior to the rest. Obviously the CC-TVJR-GARCH models generate some rather inaccurate forecasts that are downweighted by the use of MAD or MedSE.

6 Empirical analysis II: Portfolio allocation

Another way of evaluating our models out-of-sample is to consider the economic value of volatility timing. 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

$$\min_{w_t} w_t' \Sigma_t w_t - \frac{1}{\gamma} \mu_t' w_t$$

s.t. $w_t' 1 = 1$

where $w_t$ is an $N \times 1$ vector of portfolio weights, $\mu_t$ is the $N \times 1$ vector of expected returns, $1$ is an $N$-dimensional vector of ones, and $\gamma$ 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 portfolio
Table 5: Log-likelihood values from the 7-variate normal density for the CC-GJR-GARCH models

<table>
<thead>
<tr>
<th>Models</th>
<th>GJR-GARCH</th>
<th>TVGJR-GARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCC-GARCH</td>
<td>-34761.1</td>
<td>-28861.3</td>
</tr>
<tr>
<td>DCC-GARCH</td>
<td>-34629.4</td>
<td>-28743.3</td>
</tr>
<tr>
<td>VC-GARCH</td>
<td>-34611.8</td>
<td>-28730.8</td>
</tr>
</tbody>
</table>

*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).

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

<table>
<thead>
<tr>
<th>Models</th>
<th>RMSE₁</th>
<th>MAD₁</th>
<th>MedSE₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCC-GJR-GARCH</td>
<td>27.687</td>
<td>11.587</td>
<td>97.217</td>
</tr>
<tr>
<td>CCC-TVGJR-GARCH</td>
<td>28.087</td>
<td>11.298</td>
<td>71.761</td>
</tr>
<tr>
<td>DCC-GJR-GARCH</td>
<td>27.430</td>
<td>12.008</td>
<td>106.15</td>
</tr>
<tr>
<td>DCC-TVGJR-GARCH</td>
<td>27.790</td>
<td>11.512</td>
<td>82.087</td>
</tr>
<tr>
<td>VC-GJR-GARCH</td>
<td>27.421</td>
<td>12.085</td>
<td>112.31</td>
</tr>
<tr>
<td>VC-TVGJR-GARCH</td>
<td>27.824</td>
<td>11.499</td>
<td>78.694</td>
</tr>
</tbody>
</table>

*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 criteria RMSE and MAD are described in equations (5.1) and (5.2) in Pelletier (2006). The results are based on the out-of-sample data from October 8, 2008 to December 31, 2009 (311 observations).
Figure 8: Difference 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.
variance, that achieves a target expected return

$$\min_{\mathbf{w}_t} \mathbf{w}_t' \Sigma_t \mathbf{w}_t \quad \text{s.t.} \quad \mathbf{w}_t' \mu_t = \mu_p \text{ and } \mathbf{w}_t' \mathbf{1} = 1$$

where $\mu_p$ 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 such that they represent five different industries. 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, Kirby and Ostdiek (2001), we also consider a utility-based measure denoted by $\Delta_\gamma$ to compare the performance of any CC-TVGJR-GARCH model to that of the corresponding CC-GJR-GARCH model. The value of $\Delta_\gamma$ is such that an investor with a quadratic utility function and the relative risk aversion parameter $\gamma$ is indifferent between receiving $r_t$ and $r_{TV_t} - \Delta_\gamma$, where $r_t$ is the out-of-sample portfolio return under the the CC-GJR-GARCH model and $r_{TV_t}$ the corresponding return under the CC-TVGJR-GARCH model.

Tables 7 and 8 contain the results of 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 $|w_{it} - w_{it-1}|$ over the period, where $w_{it}$ is the optimal weight of asset $i$ on day $t$ and $w_{it-1}$ 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.

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, it is seen that the GMV and the Min-Variance strategies tend to have better stability properties when the long-run variances are modelled, as the turnover often decreases. Given these results one can 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 $\Delta \gamma$ as annualized fees in basis points using two different risk aversion parameters, $\gamma = 1$ and $\gamma = 10$. Over all six asset combinations and the 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.
Table 7: Out-of-sample portfolio performance of the volatility-timing strategies: CCC-GJR-GARCH vs. CCC-TVGJR-GARCH

<table>
<thead>
<tr>
<th>Portfolios</th>
<th>CCC-GJR-GARCH</th>
<th>CCC-TVGJR-GARCH</th>
<th>Δγ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>μ</td>
<td>σ</td>
<td>Turnover</td>
</tr>
<tr>
<td>AXP–BA–CAT–INTC–JPM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMV</td>
<td>11.82</td>
<td>17.64</td>
<td>0.670</td>
</tr>
<tr>
<td>Min-Variance</td>
<td>12.04</td>
<td>19.00</td>
<td>0.634</td>
</tr>
<tr>
<td>Mean-Variance</td>
<td>12.09</td>
<td>27.08</td>
<td>0.446</td>
</tr>
<tr>
<td>AXP–BA–CAT–INTC–WHR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMV</td>
<td>8.035</td>
<td>17.85</td>
<td>0.450</td>
</tr>
<tr>
<td>Min-Variance</td>
<td>5.502</td>
<td>19.20</td>
<td>0.287</td>
</tr>
<tr>
<td>Mean-Variance</td>
<td>7.578</td>
<td>27.12</td>
<td>0.279</td>
</tr>
<tr>
<td>AXP–BA–CAT–INTC–XOM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMV</td>
<td>-8.258</td>
<td>16.15</td>
<td>-0.511</td>
</tr>
<tr>
<td>Min-Variance</td>
<td>7.902</td>
<td>18.74</td>
<td>0.422</td>
</tr>
<tr>
<td>Mean-Variance</td>
<td>17.78</td>
<td>26.94</td>
<td>0.660</td>
</tr>
<tr>
<td>BA–CAT–INTC–JPM–WHR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMV</td>
<td>4.105</td>
<td>17.88</td>
<td>0.230</td>
</tr>
<tr>
<td>Min-Variance</td>
<td>2.722</td>
<td>20.43</td>
<td>0.133</td>
</tr>
<tr>
<td>Mean-Variance</td>
<td>1.970</td>
<td>21.91</td>
<td>0.090</td>
</tr>
<tr>
<td>BA–CAT–INTC–JPM–XOM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMV</td>
<td>-8.327</td>
<td>16.02</td>
<td>-0.520</td>
</tr>
<tr>
<td>Min-Variance</td>
<td>8.881</td>
<td>20.41</td>
<td>0.435</td>
</tr>
<tr>
<td>Mean-Variance</td>
<td>5.660</td>
<td>24.21</td>
<td>0.234</td>
</tr>
<tr>
<td>CAT–INTC–JPM–WHR–XOM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMV</td>
<td>-19.91</td>
<td>17.46</td>
<td>-1.140</td>
</tr>
<tr>
<td>Min-Variance</td>
<td>2.543</td>
<td>22.52</td>
<td>0.113</td>
</tr>
<tr>
<td>Mean-Variance</td>
<td>-1.520</td>
<td>25.48</td>
<td>-0.060</td>
</tr>
</tbody>
</table>

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 γ = 1 or γ = 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

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

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 γ = 1 or γ = 10 would be willing to pay to switch from the constant to the time-varying unconditional variance strategy.
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 $g_{it-1}$. They will be called time-varying news impact surfaces. The time-varying news impact surface for $h_{ij}$ is the three-dimensional graph of the function

$$h_{ij} = f(\varepsilon_{i,t-1}/g_{i,t-1}, \varepsilon_{j,t-1}/g_{j,t-1}, \rho_{ij,t-1}; h_{t-1})$$

where $h_{t-1}$ is a vector of conditional covariances at time $t-1$ defined at their unconditional sample means. As an example, Figure 9 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
Figure 10: Estimated time-varying news impact surfaces for the conditional variance of the BA returns under the CCC-TVGJR-GARCH model in the (a) lower regime and (b) upper regime of volatility.

Figure 11: Estimated time-varying news impact surfaces for the conditional variance of the XOM returns under the CCC-TVGJR-GARCH model in the (a) lower regime and (b) upper regime of volatility.
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 component $g_{t-1}$. 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 asymmetric bowl-shaped impact surface.

Figure 10 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 11 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 9, 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.

Estimated news impact surfaces from the TVC-TVGJR-GARCH model of the BA-XOM pair are plotted in Figure 12. These news impact surfaces are able to distinguish between responses during low and high variance as well as low and high correlation levels. It is seen from Figure 12 that both the degree of turbulence in the market and the level of the correlations affect the impact of past shocks on the covariances. This indicates that both factors play an important role in assessing the effect of shocks on the covariances according to the TVC-TVGJR-GARCH model. High covariance estimates are related to strong correlations and a high degree of turbulence in the market.

8 Conclusions

In this paper, we extend the univariate multiplicative TV-GARCH model of Amado and Teräsvirta (2011) 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-
Lower regime for the correlations

![Diagram](a)

Upper regime for the correlations

![Diagram](b)

Figure 12: Estimated time-varying news impact surfaces for the covariance between the BA and XOM returns under the TVC-TVGJR-GARCH model in the (a) lower regime and in the (b) upper regime of volatility.
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 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 are, however, different for the STCC-GARCH model of Silvennoinen and Teräsvirta (2009a, in press). In a number of occasions, the correlations estimated from this model with time as the sole transition variable (TVC-GJR-GARCH) are quite different from what they are when the GJRGARCH equations are implicitly assumed stationary. The most conspicuous difference between the estimated TVC-TVGJR-GARCH and TVC-GJR-GARCH models can be found in the degree of smoothness of the change in the correlations, but the direction of the change is not systematic.

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.

Furthermore, the TVGJR-GARCH approach gives us the opportunity to generalize the news impact surfaces introduced by Kroner and Ng (1998) such that they can vary over time. In the TVC-TVGJR-GARCH model, the impact of news (shocks) on the covariances between returns is a function of both time-varying variances and time-varying correlations. As in the univariate case already considered in Amado and Teräsvirta (2011), it is seen that the impact of a piece of news of
a given size is larger when the market is calm than when it is during periods of high volatility. In
the multivariate case we can also conclude that high conditional correlation between two returns
adds to the impact as compared to the situation in which the correlation is low. We also reproduce
the old result that negative shocks or news have a stronger effect on volatility than positive news
of the same size.

An extension of this methodology to the case in which the conditional correlations are also con-
trolled by a stochastic variable is available through the Double STCC-GJR-GARCH model. Using
that model would make it possible, for example, to model asymmetric responses of conditional
correlations to functions of past returns. This, however, is a topic left for future research.

References

Amado, C. and Teräsvirta, T.: 2011, Modelling volatility by variance decomposition, CREATESES
Research Paper 2011-1, Aarhus University.


Berben, R.-P. and Jansen, W. J.: 2005, Comovements in international equity markets: A sectorial

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

Dahlhaus, R. and Subba Rao, S.: 2006, Statistical inference for time-varying ARCH processes,

Research 57, 560–577.

Diebold, F. X.: 1986, Comment on 'Modelling the persistence of conditional variance', by R.F.


Lamoureux, C. G. and Lastrapes, W. D.: 1990, Persistence in variance, structural change, and the

Lavielle, M. and Teyssi`ere, G.: 2006, Detection of multiple change-points in multivariate time


Mikosch, T. and St˘aric˘a, C.: 2004, Nonstationarities in financial time series, the long-range de-

473.

Silvennoinen, A. and Teräsvirta, T.: 2009a, Modelling multivariate autoregressive conditional
heteroskedasticity with the double smooth transition conditional correlation GARCH model,

Silvennoinen, A. and Teräsvirta, T.: 2009b, Multivariate GARCH models, *in* T. G. Andersen,

Silvennoinen, A. and Teräsvirta, T.: in press, Modelling conditional correlations of asset returns:
A smooth transition approach, *Econometric Reviews* .

Tse, Y. and Tsui, A. K.: 2002, A multivariate GARCH model with time-varying correlations,