Building multiplicative time-varying smooth transition conditional correlation GARCH models

EXTENDED ABSTRACT

Anthony D. Hall*, Annastiina Silvennoinen** and Timo Teräsvirta†‡

*University of Technology, Sydney

**NCER, Queensland University of Technology, Brisbane

†CREATES, Aarhus University

‡C.A.S.E., Humboldt-Universität zu Berlin

December 6, 2017

Abstract

This paper explores the complete modelling process for an MTV–STCC–GARCH model. Building such nonlinear models is a data-driven process, however, recommendations on the systematic approach using statistical inference can be made. The focus is first on the deterministic components of the model. Specification tests for the univariate parts of the model draw from the authors' earlier work. A new contribution is the specification test for the multivariate part, that is, the correlation component. Simulation studies provide support for the robustness of the proposed model building approach. The full estimation procedure is also laid out, followed by misspecification tests for the purpose of model evaluation. The application of the proposed methodology to Australian four largest banks illustrates the modelling cycle.

JEL Classification Codes: C32; C52; C58

Keywords. unconditional correlation; time-varying correlation; modelling volatility; multivariate autoregressive conditional heteroskedasticity;

1 Introduction

Multivariate models of generalized autoregressive conditional heteroskedasticity (GARCH) provide a useful way of generating (short-run) forecasts for portfolio management. The first multivariate GARCH model was formulated and applied to financial time series by Bollerslev, Engle and Wooldridge (1988). The area has expanded rapidly and has been reviewed in surveys such as Bauwens, Laurent and Rombouts (2006) and Silvennoinen and Teräsvirta (2009b). The family of multivariate GARCH models we focus on in this work is the class of conditional correlation GARCH (CC–GARCH) models introduced by Bollerslev (1990). In the original model the conditional correlations were constant, hence the name Constant Conditional Correlation (CCC–GARCH) model. This assumption that made the resulting model rather parsimonious was later found too restrictive and time-varying correlations were introduced by Engle (2002) and Tse and Tsui (2002). In this paper we are interested in extensions to the Smooth Transition CC (STCC–) GARCH model, see Silvennoinen and Teräsvirta (2005, 2009a, 2015). One such extension was made by Amado and Teräsvirta (2014) who completed the GARCH equations of the model by a deterministically time-varying component. The correlations in their model were parameterised as in Bollerslev (1990), Engle (2002) and Tse and Tsui (2002). The idea of a deterministically time-varying component to describe smoothly changing variation in univariate GARCH models has been explored by several authors; see Feng (2004), van Bellegem and von Sachs (2004), Engle and Rangel (2008), Brownlees and Gallo (2010), Mazur and Pipień (2012) or Amado and Teräsvirta (2008, 2013, 2017) for examples.

Silvennoinen and Teräsvirta (2017a) assume that the time-varying correlations are also deterministic. This shifts the emphasis from short-run to smooth long run movements in correlations, see Engle (2002) and Tse and Tsui (2002) for modelling the former. At the same time, similarly to Amado and Teräsvirta (2014), the GARCH equations in the Multiplicative Time-Varying Smooth Transition Correlation GARCH model, MTV model for short, contain a multiplicative deterministic component. Silvennoinen and Teräsvirta (2017a) show that under regularity conditions, the maximum likelihood estimates of the parameters of the MTV model are consistent and asymptotically normal.

Building MTV models is a data-driven process. Silvennoinen and Teräsvirta (2017a) focus on estimation and do not consider the modelling process, which is the topic of this paper. We argue that the MTV model has to be built up in a systematic fashion using statistical inference. The first step is to specify the model, that is, select a member from the MTV family. After this has been done, the selected model is estimated. Finally, the estimated model has to be evaluated, which can be done by misspecification tests.

The plan of the paper is as follows. The MTV model is introduced in Section 2. The three stages of building MTV models are briefly mentioned in Section 3. Specification, estimation and evaluation stages of model building are considered in Sections 4 and 5. Finite sample properties of specification tests are investigated in Section 6 by simulation.

An empirical illustration is chosen from the Australian banking sector. The "Big Four", National Australia Bank (NAB), Commonwealth Bank of Australia (CBA), Australia and New Zealand Banking Group (ANZ), and Westpac (WBC), are the largest banks by market share, holding 80% of the home loan markets in Australia. The daily frequency data set spans from early 90's until present, thus providing a long time period over which a few major incidents and changes have taken place, e.g. the dot-com boom in the late 90's, the GFC nearly ten years later, a number of regulatory changes (Basel guidelines), technology driven market disruptions (non bank lenders and payment providers). The impact of such events may be identified as the timing of the changes in the correlation between the four banks. Furthermore, the analysis may reveal some interesting aspects of the effectiveness of the Four Pillars policy established by the Keating government in 1990.

2 The model

As already indicated, the topic of this paper is building multivariate multiplicative time-varying Smooth Transition Conditional Correlation GARCH (MTV–STCC–GARCH) models with deterministically time-varying correlations. Main features of the model, defined in Silvennoinen and Teräsvirta (2017a), are reproduced here. The observable stochastic $N \times 1$ vector $\boldsymbol{\varepsilon}_t$ is decomposed in a customary fashion as

$$\boldsymbol{\varepsilon}_t = \boldsymbol{H}_t^{1/2}\boldsymbol{z}_t = \boldsymbol{S}_t\boldsymbol{D}_t\boldsymbol{z}_t \tag{1}$$

where $\boldsymbol{z}_t$ is a vector of independent random variables with $\mathsf{E}\boldsymbol{z}_t = \boldsymbol{0}$ and a positive definite deterministically varying covariance matrix $\text{cov}(\boldsymbol{z}_t) = \boldsymbol{P}_t$. The structure of $\boldsymbol{P}_t$ will be defined later. It follows that $\boldsymbol{P}_t^{-1/2}\boldsymbol{z}_t = \boldsymbol{\zeta}_t \sim \text{iid}(\boldsymbol{0}, \boldsymbol{I}_N)$. The diagonal $N \times N$ matrix $\boldsymbol{H}_t^{1/2} = \boldsymbol{S}_t\boldsymbol{D}_t$, where the deterministic matrix $\boldsymbol{S}_t = \text{diag}(g_{1t}^{1/2}, ..., g_{Nt}^{1/2})$ has positive diagonal elements for all t, and $\boldsymbol{D}_t = \text{diag}(h_{1t}^{1/2}, ..., h_{Nt}^{1/2})$ contains the conditional standard deviations of the elements of $\boldsymbol{S}_t^{-1}\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}/g_{1t}^{1/2}, ..., \varepsilon_{Nt}/g_{Nt}^{1/2})'$. As in Silvennoinen and Teräsvirta (2017a) and earlier univariate papers, beginning with Amado and

Teräsvirta (2008), and in the multivariate time-varying GARCH article by Amado and Teräsvirta (2014), the diagonal elements of $\boldsymbol{S}_t^2$ are defined as follows:

$$g_{it} = g_i(t/T) = \delta_{i0} + \sum_{j=1}^{r_i} \delta_{ij} G_{ij}(\gamma_{ij}, \boldsymbol{c}_{ij}; t/T) \tag{2}$$

$i = 1, ..., N$, where $\delta_{i0} > 0$ is a known constant, and the (generalised) logistic function

$$G_{ij}(\gamma_{ij}, \boldsymbol{c}_{ij}; t/T) = (1 + \exp\{-\gamma_{ij} \prod_{k=1}^{K_{ij}} (t/T - c_{ijk})\})^{-1},\ \gamma_{ij} > 0 \tag{3}$$

and $\boldsymbol{c}_{ij} = (c_{ij1}, ..., c_{ijK_{ij}})'$ such that $c_{ij1} \leq ... \leq c_{ijK_{ij}}$. Both $\gamma_{ij} > 0$ and $c_{ij1} \leq ... \leq c_{ijK_{ij}}$ are identification restrictions. Assuming δ_{i0} in (2) is known (or setting its value equal to e.g. one) is another one.

The conditional variances have a GARCH or GJR–GARCH(1,1) structure, see Glosten, Jagannathan and Runkle (1993) for the latter:

$$h_{it} = \alpha_{i0} + \alpha_{i1}\varepsilon_{i,t-1}^2 + \kappa_{i1} I(\varepsilon_{t-1} < 0)\varepsilon_{i,t-1}^2 + \beta_{i1} h_{i,t-1} \tag{4}$$

where $I(A)$ is an indicator function: $I(A) = 1$ when A occurs, zero otherwise. A higher-order structure is possible, although there do not seem to exist applications of the GJR-GARCH model of order greater than one.

As discussed in earlier papers, the idea of g_{it} is to normalise or rescale the observations. Left-multiplying (1) by $\boldsymbol{S}_t^{-1}$ yields

$$\boldsymbol{\phi}_t = \boldsymbol{S}_t^{-1}\boldsymbol{\varepsilon}_t = \boldsymbol{D}_t \boldsymbol{z}_t \tag{5}$$

where $\boldsymbol{\phi}_t$ is assumed to have a standard weakly stationary GARCH representation with the conditional covariance matrix $\mathsf{E}\{\boldsymbol{\phi}_t\boldsymbol{\phi}_t'|\mathcal{F}_{t-1}\} = \boldsymbol{D}_t\boldsymbol{P}_t\boldsymbol{D}_t$. In this work, $\boldsymbol{P}_t$ is defined as in the smooth transition conditional correlation model of Silvennoinen and Teräsvirta (2005, 2015):

$$\boldsymbol{P}_t = G_t(t/T, \gamma, c)\boldsymbol{P}_{(1)} + \{1 - G_t(t/T, \gamma, c)\}\boldsymbol{P}_{(2)} \tag{6}$$

where $\boldsymbol{P}_{(1)}$ and $\boldsymbol{P}_{(2)}$, $\boldsymbol{P}_{(1)} \neq \boldsymbol{P}_{(2)}$, are positive definite correlation matrices, and

$$G_t(t/T, \gamma, c) = (1 + \exp\{-\gamma \prod_{k=1}^{K} (t/T - c_k)\})^{-1},\ \gamma > 0 \tag{7}$$

and $c_1 \leq ... \leq c_K$. As a convex combination of $\boldsymbol{P}_{(1)}$ and $\boldsymbol{P}_{(2)}$, $\boldsymbol{P}_t$ is positive definite. The resulting model is a Multiplicative Time-Varying Smooth Transition Conditional Correlation (MTV–STCC) GARCH model. A bivariate

STCC–GARCH model with time as the transition variable was first considered by Berben and Jansen (2005). A similar MTV–Conditional Correlation GARCH model but with different definitions of $\boldsymbol{P}_t$ was discussed in Amado and Teräsvirta (2014).

3 The three stages of model building

The MTV–STCC–GARCH model, or MTV model for short, is rather general and nests many models that are not identified. This being the case, fitting a general model when a more specific nested submodel actually generates the data leads to inconsistent parameter estimates. For this reason, building adequate MTV models requires care. A systematic approach is necessary. Selecting a candidate from this family of models is a data-driven process, and statistical inference has to be used to obtain an acceptable end product that passes the available misspecification tests and can then be used for forecasting and portfolio management.

In this work we follow the classical approach to model building already advocated by Box and Jenkins (1970) and applied to nonlinear models of the conditional mean, see for example Teräsvirta, Tjøstheim and Granger (2010, Ch. 16). It has also been applied to building single-equation multiplicative TV–GARCH models; see Amado and Teräsvirta (2017) and Silvennoinen and Teräsvirta (2016). The idea is to first specify the model (select a member from the family of MTV models) and once this has been done estimate its parameters. At the evaluation stage the estimated model is then subjected to a battery of misspecification tests. These three stages, specification, estimation and evaluation will be considered in next sections. The emphasis will be on specification and evaluation as maximum likelihood estimation of the parameters of the MTV model has already been considered in Silvennoinen and Teräsvirta (2017a).

4 Specification of the MTV model

4.1 Specification of GARCH equations

Specification of the MTV model is begun by specifying the GARCH equations. This was first discussed in Amado and Teräsvirta (2017). The idea is to begin with a GARCH(1,1) model by Bollerslev (1986) or the GJR-GARCH model by Glosten et al. (1993) and test the hypothesis that the multiplicative deterministic component is constant. The Multiplicative TV–GARCH model has the following form:

$$\varepsilon_t = z_t h_t^{1/2} g_t^{1/2} \tag{8}$$

where $z_t \sim \text{iid}(0,1)$. The conditional variance

$$h_t = \alpha_0 + \alpha_1 \phi_{t-1}^2 + \kappa_1 I(\phi_{t-1} < 0)\phi_{t-1}^2 + \beta_1 h_{i,t-1}$$

where $\phi_t = \varepsilon_t / g^{1/2}(t/T;)$. The deterministic positive-valued function $g_t = g(t/T) = g(t/T; \boldsymbol{\theta}_1)$ is defined as in (2):

$$g(t/T) = \delta_0 + \sum_{j=1}^{r} \delta_j G_j(t/T, \gamma, \boldsymbol{c}) \tag{9}$$

where, following (3),

$$G_j(t/T, \gamma, c) = (1 + \exp\{-\gamma \prod_{k=1}^{K_j} (t/T - c_k)\})^{-1}, \ \gamma > 0. \tag{10}$$

Positivity of (9) imposes restrictions on δ_i, $i = 1, ..., r$. Typically in applications, $K_j = 1, 2$. There are two specification issues, determining r and choosing K_j, $j = 1, ..., r$. It is possible that $g(t/T) = \delta_0 > 0$, that is, $g(t/T)$ is a positive constant. In this case the GARCH model collapses into a standard GARCH or GJR–GARCH equation.

Amado and Teräsvirta (2017) solved the problem of choosing r by first estimating the GARCH model and testing the hypothesis of a constant $g(t/T)$ against the alternative $r = 1$ in (9) thereafter using a Lagrange multiplier type of test. The test can be viewed as a misspecification test of the estimated GARCH model. If the null hypothesis is rejected, a TV–GARCH model with a single transition is estimated, and the hypothesis $r = 1$ is tested against $r = 2$. Sequential testing continues until the first non-rejection of the null hypothesis. The number of transitions is determined in this order because of an identification problem: the model with $r + 1$ transitions is not identified if the true number of transitions is r. The shape of the logistic function, controlled by the parameter K_j, can be determined using the sequence of tests familiar from the specification of smooth transition autoregressive (STAR) models, see Teräsvirta (1994) or Teräsvirta, Tjøstheim and Granger (2010, Chapter 16). Details can be found in Amado and Teräsvirta (2017).

More recently, Silvennoinen and Teräsvirta (2016) considered testing the constancy of $g(t/T)$ before estimating the GARCH model, that is, assuming $h_t = 1$ in (8). This implies that the size of the test is distorted because conditional heteroskedasticity is ignored, so the size of the test has to be adjusted by simulation. It turned out that power of the size-adjusted test improved considerably compared to the case where the test is a misspecification test. Reasons for this are discussed in Silvennoinen and Teräsvirta (2016).

4.2 Test statistic

In this work we take this idea a step further and not only test constancy but even specify the number of transitions before estimating the GARCH component of the model. Amado and Teräsvirta (2013) showed that maximum likelihood estimators of the corresponding time-varying variance (TVV) model, assuming that there is no conditional heteroskedasticity, are consistent and asymptotically normal. This forms the base for constructing Lagrange multiplier type tests for testing r against $r+1$ transitions. For notational simplicity consider testing one transition against two. The TVV model is (8) with $h_t = 1$, and

$$g_t = \delta_0 + \delta_1 G_1(t/T, \gamma_1, \boldsymbol{c}_1) + \delta_2 G_2(t/T, \gamma_2, \boldsymbol{c}_2),\ \gamma_i > 0,\ i = 1, 2.$$

The null hypothesis is $\gamma_2 = 0$, in which case $G_2(t/T, \gamma_2, \boldsymbol{c}_2) \equiv 1/2$. To circumvent the identification problem (the model with one transition is only identified when the alternative is true, $\gamma_2 > 0$) we follow Luukkonen, Saikkonen and Teräsvirta (1988) and approximate the second transition by a third-order Taylor expansion around the null hypothesis. After reparameterisation this yields

$$g_t = \delta_0 + \delta_1 G_1(t/T, \gamma_1, \boldsymbol{c}_1) + \psi_1 t/T + \psi_2 (t/T)^2 + \psi_3 (t/T)^3,\ \gamma_1 > 0. \tag{11}$$

We may call (8) with (11) the auxiliary TVV model. The parameters $\psi_i = \gamma_2 \widetilde{\psi}_i$, where $\widetilde{\psi}_i \neq 0$, $i = 1, 2, 3$. The new null hypothesis in (11) equals H_0': $\psi_1 = \psi_2 = \psi_3 = 0$. The remainder term of the expansion can be ignored because when we construct a Lagrange multiplier test, the model is only estimated under H_0 (or H_0') and under this hypothesis the order of the Taylor expansion equals zero. The remainder does appear under the alternative, and so ignoring it when H_0 is valid does not affect the asymptotic size of the test but makes a positive contribution to its power when H_0 does not hold.

Assume (again for notational simplicity) that $K_1 = 1$ in (11), so $\boldsymbol{c}_1 = c_1$ (a scalar). The log-likelihood for observation t of the auxiliary TVV model equals

$$\ell_t = k - (1/2) \ln g_t - (1/2) \frac{\varepsilon_t^2}{g_t}$$

and the corresponding element of the score is

$$\frac{\partial \ell_t}{\partial \boldsymbol{\theta}_1} = \frac{1}{2} (\frac{\varepsilon_t^2}{g_t} - 1) \frac{1}{g_t} \frac{\partial g_t}{\partial \boldsymbol{\theta}_1} \tag{12}$$

where $\boldsymbol{\theta}_1 = (\delta_0, \delta_1, \gamma_1, c_1, \psi_1, \psi_2, \psi_3)'$. Denoting $G_1(t/T) = G_1(t/T, \gamma_1, c_1)$, the partial derivative in (12) is

$$\frac{\partial g_t}{\partial \boldsymbol{\theta}_1} = \{1, G_1(t/T), G_{1\gamma}(t/T), G_{1c}(t/T), t/T, (t/T)^2, (t/T)^3\}'$$

where $G_{1\gamma}(t/T) = G_1(t/T)(1-G_1(t/T))(t/T-c_1)$ and $G_{1c}(t/T) = -\gamma_1 G_1(t/T)(1-G_1(t/T))$. Define the true parameter vector under H_0 as $\boldsymbol{\theta}_1^0 = (\delta_0^0, \delta_1^0, \gamma_1^0, c_1^0, 0, 0, 0)'$. If $\boldsymbol{z}_t$ is normally distributed, the corresponding element of the information matrix under H_0 has the form

$$\begin{aligned}\boldsymbol{B}_t &= \frac{1}{4}\mathsf{E}(\frac{\varepsilon_t^2}{g_t}-1)^2 \left[\begin{array}{cc} \boldsymbol{B}_{11t} & \boldsymbol{B}_{12t} \\ \boldsymbol{B}_{21t} & \boldsymbol{B}_{22t} \end{array}\right] \\ &= \frac{1}{2}\left[\begin{array}{cc} \boldsymbol{B}_{11t} & \boldsymbol{B}_{12t} \\ \boldsymbol{B}_{21t} & \boldsymbol{B}_{22t} \end{array}\right].\end{aligned}$$

where, letting $g_t^0 = \delta_0^0 + \delta_1^0 G_1(t/T, \gamma_1^0, \boldsymbol{c}_1^0)$ and denoting $G_1^0(t/T) = G_1(t/T, \gamma_1^0, \boldsymbol{c}_1^0)$,

$$\boldsymbol{B}_{11t} = \frac{1}{2(g_t^0)^2}\left[\begin{array}{cccc} 1 & G_1^0(t/T) & G_{1\gamma}^0(t/T) & G_{1c}^0(t/T) \\ & (G_1^0(t/T))^2 & G_1^0(t/T)G_{1\gamma}^0(t/T) & G_1^0(t/T)G_{1c}^0(t/T) \\ & & (G_{1\gamma}^0(t/T))^2(t/T) & G_{1\gamma}^0(t/T)G_{1c}^0(t/T) \\ & & & (G_{1c}^0(t/T))^2(t/T) \end{array}\right]$$

and

$$\boldsymbol{B}_{12t} = \frac{1}{2(g_t^0)^2}\left[\begin{array}{ccc} t/T & (t/T)^2 & (t/T)^3 \\ G_1^0(t/T)(t/T) & G_1^0(t/T)(t/T)^2 & G_1^0(t/T)(t/T)^3 \\ G_{1\gamma}^0(t/T)(t/T) & G_{1\gamma}^0(t/T)(t/T)^2 & G_{1\gamma}^0(t/T)(t/T)^3 \\ G_{1c}^0(t/T)(t/T) & G_{1c}^0(t/T)(t/T)^2 & G_{1c}^0(t/T)(t/T)^3 \end{array}\right].$$

Denoting $\boldsymbol{t}_t = (t/T, (t/T)^2, (t/T)^3)'$,

$$\boldsymbol{B}_{22t} = \frac{1}{2(g_t^0)^2}\boldsymbol{t}_t\boldsymbol{t}_t'.$$

We state the following lemma:

Lemma 1 *Under the null hypothesis and assuming $z_t \sim iid\mathcal{N}(0,1)$, the information matrix*

$$\boldsymbol{B} = \frac{1}{2}\left[\begin{array}{cc} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{array}\right] = \lim_{T\to\infty}\frac{1}{2T}\sum_{t=1}^{T}\left[\begin{array}{cc} \boldsymbol{B}_{11t} & \boldsymbol{B}_{12t} \\ \boldsymbol{B}_{21t} & \boldsymbol{B}_{22t} \end{array}\right]$$

has the following form:

$$\boldsymbol{B}_{11} = \frac{1}{2}\begin{bmatrix} \int_0^1 (g_r^0)^{-2}\,dr & \int_0^1 (g_r^0)^{-2} G_1^0(r)\,dr & \int_0^1 (g_r^0)^{-2} G_{1\gamma}^0(r)\,dr & \int_0^1 (g_r^0)^{-2} G_{1c}^0(r)\,dr \\ & \int_0^1 (g_r^0)^{-2} G_1^0(r)^2\,dr & \int_0^1 (g_r^0)^{-2} G_1^0(r) G_{1\gamma}^0(r)\,dr & \int_0^1 (g_r^0)^{-2} G_1^0(r) G_{1c}^0(r)\,dr \\ & & \int_0^1 (g_r^0)^{-2} G_{1\gamma}^0(r)^2\,dr & \int_0^1 (g_r^0)^{-2} G_{1\gamma}^0(r) G_{1c}^0(r)\,dr \\ & & & \int_0^1 (g_r^0)^{-2} G_{1c}^0(r)^2\,dr \end{bmatrix}$$

$$\boldsymbol{B}_{12} = \frac{1}{2}\begin{bmatrix} \int_0^1 r(g_r^0)^{-2}\,dr & \int_0^1 r^2(g_r^0)^{-2}\,dr & \int_0^1 r^3(g_r^0)^{-2}\,dr \\ \int_0^1 r(g_r^0)^{-2} G_1^0(r)\,dr & \int_0^1 r^2(g_r^0)^{-2} G_1^0(r)\,dr & \int_0^1 r^3(g_r^0)^{-2} G_1^0(r)\,dr \\ \int_0^1 r(g_r^0)^{-2} G_{1\gamma}^0(r)\,dr & \int_0^1 r^2(g_r^0)^{-2} G_{1\gamma}^0(r)\,dr & \int_0^1 r^3(g_r^0)^{-2} G_{1\gamma}^0(r)\,dr \\ \int_0^1 r(g_r^0)^{-2} G_{1c}^0(r)\,dr & \int_0^1 r^2(g_r^0)^{-2} G_{1c}^0(r)\,dr & \int_0^1 r^3(g_r^0)^{-2} G_{1c}^0(r)\,dr \end{bmatrix}$$

and

$$\boldsymbol{B}_{22} = \lim_{T\to\infty} \frac{1}{2T}\sum_{t=1}^{T} (g_t^0)^{-2} \boldsymbol{t}_t \boldsymbol{t}_t' = \frac{1}{2}\begin{bmatrix} \int_0^1 r^2(g_r^0)^{-2}\,dr & \int_0^1 r^3(g_r^0)^{-2}\,dr & \int_0^1 r^4(g_r^0)^{-2}\,dr \\ & \int_0^1 r^4(g_r^0)^{-2}\,dr & \int_0^1 r^5(g_r^0)^{-2}\,dr \\ & & \int_0^1 r^6(g_r^0)^{-2}\,dr \end{bmatrix}.$$

Proof. See Appendix A.

Since the maximum likelihood estimators of the parameters of the auxiliary TVV model under H_0 are consistent, we may construct the LM test for the hypothesis H_0': $\boldsymbol{\psi} = (\psi_1, \psi_2, \psi_3)' = \mathbf{0}$. Denoting the relevant block of the score by

$$\boldsymbol{s}_2(\widehat{\boldsymbol{\theta}}_1) = \frac{1}{2T}\sum_{t=1}^{T}(\frac{\varepsilon_t^2}{\widehat{g}_t} - 1)\frac{1}{\widehat{g}_t}\frac{\partial g_t}{\partial \boldsymbol{\psi}}$$

where $\partial g_t / \partial \boldsymbol{\psi} = \boldsymbol{t}_t$, and

$$\widehat{g}_t = \widehat{\delta}_0 + \widehat{\delta}_1(1 + \exp\{-\widehat{\gamma}_1(t/T - \widehat{c}_1)\})^{-1}$$

the test statistic has the following form:

$$LM_T = (T/2)\boldsymbol{s}_2'(\widehat{\boldsymbol{\theta}}_1)(\boldsymbol{B}_{22} - \boldsymbol{B}_{21}\boldsymbol{B}_{11}^{-1}\boldsymbol{B}_{12})^{-1}\boldsymbol{s}_2(\widehat{\boldsymbol{\theta}}_1) \tag{13}$$

where $\widehat{\boldsymbol{\theta}}_1 = (\widehat{\delta}_0, \widehat{\delta}_1, \widehat{\gamma}_1, \widehat{c}_1, 0, 0, 0)'$, see for example Godfrey (1988, p. 14). In order to make (13) operational, the blocks of $\boldsymbol{B}$ are replaced by their consistent counterparts.

As already mentioned, conditional heteroskedasticity is ignored in setting up the test. For this reason, the test statistic (13) is likely to be size distorted when applied to financial time series of sufficiently high frequency. In applications its size thus has to be adjusted by resampling. Finite sample properties of the test statistic, both its adjusted size and size-adjusted power, will be examined in Section 6.

4.3 Specification of time-varying correlations

After the GARCH equations have been specified and estimated assuming they are independent, the next step is to specify the time-varying correlations. The null hypothesis is that the model is a TVV–Constant Correlation GARCH model as in Bollerslev (1990), except that the GARCH equations are TVV equations. The alternative is the MTV model in which the correlations are time-varying and defined by (6) and (7). As discussed in Silvennoinen and Teräsvirta (2005, 2015), the MTV model is only identified under the alternative, which invalidates the standard asymptotic inference. The identification problem can be circumvented by approximating the transition function (7) by its Taylor expansion around the null hypothesis, H_0: $\gamma = 0$. The form of the expansion depends on the order of the exponent in (7).

The test can be constructed along the lines presented in the appendix of Silvennoinen and Teräsvirta (2005); available also in `http://econ.au.dk/research/research-centres/creates/research/research-papers/`

`supplementary-downloads`. See also Silvennoinen and Teräsvirta (2017a). To derive the test statistic, consider the first-order Taylor expansion of (10) around $\gamma = 0$ assuming $K = 2$. It has the following form:

$$\begin{aligned} G_t(t/T, \gamma, c) &= (1 + \exp\{-\gamma \prod_{k=1}^{K} (t/T - c_k)\})^{-1} \\ &= \frac{1}{2} + \frac{1}{4}(t/T - c_1)(t/T - c_2)\gamma + R_2(t/T; \gamma) \end{aligned} \tag{14}$$

where $R_2(t/T; \gamma)$ is the remainder. Using (14), (6) becomes

$$\begin{aligned} \boldsymbol{P}_t &= (\boldsymbol{P}_{(1)} - \boldsymbol{P}_{(2)})(\frac{1}{2} + \frac{\gamma c_1 c_2}{4}) + \boldsymbol{P}_{(2)} - (t/T)(\boldsymbol{P}_{(1)} - \boldsymbol{P}_{(2)})\frac{\gamma(c_1 + c_2)}{4} \\ &\quad + (t/T)^2(\boldsymbol{P}_{(1)} - \boldsymbol{P}_{(2)})\frac{\gamma}{4} + (\boldsymbol{P}_{(1)} - \boldsymbol{P}_{(2)})R_2(t/T; \gamma) \\ &= \boldsymbol{P}_{(A0)} + (t/T)\boldsymbol{P}_{(A1)} + (t/T)^2\boldsymbol{P}_{(A2)} + (\boldsymbol{P}_{(1)} - \boldsymbol{P}_{(2)})R_2(t/T; \gamma) \end{aligned}$$

where $\boldsymbol{P}_{(1)} \neq \boldsymbol{P}_{(2)}$. The main diagonals of $\boldsymbol{P}_{(A1)}$ and $\boldsymbol{P}_{(A2)}$ consist of zeroes. Letting $\boldsymbol{\rho}_A = (\boldsymbol{\rho}'_{A0}, \boldsymbol{\rho}'_{A1}, \boldsymbol{\rho}'_{A2})'$, where $\boldsymbol{\rho}'_{Ai} = \text{vecl}(\boldsymbol{P}_{(Ai)})$, $i = 0, 1, 2$, the new null hypothesis is H_0: $\boldsymbol{\rho}_{A1} = \boldsymbol{\rho}_{A2} = \mathbf{0}_{N(N-1)/2}$.

The log-likelihood of the auxiliary TVV–STCC–GARCH model for observation t assuming $K = 1$ equals

$$\begin{aligned} &\ln f_A(\boldsymbol{\zeta}_t | \boldsymbol{\theta}, \mathcal{F}_{t-1}) \\ &= -(1/2)\sum_{i=1}^{N} \ln g_{it} - (1/2)\ln|\boldsymbol{P}_{At}| - (1/2)\boldsymbol{\varepsilon}'_t\{\boldsymbol{S}_t \boldsymbol{P}_{At} \boldsymbol{S}_t\}^{-1}\boldsymbol{\varepsilon}_t \end{aligned} \tag{15}$$

where

$$\boldsymbol{P}_{At} = \boldsymbol{P}_{(A0)} + (t/T)\boldsymbol{P}_{(A1)} + (t/T)^2\boldsymbol{P}_{(A2)}$$

and $g_t = \delta_0 + \delta_1 G_1(t/T, \gamma_1, c_1)$; only one transition for notational simplicity. The sub-blocks of the score under H_0 become

$$\boldsymbol{s}_t(\boldsymbol{\theta}_{gi}) = \frac{1}{2g_{it}}\frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}}(\boldsymbol{e}'_i \boldsymbol{P}^{-1}_{(A0)} \boldsymbol{z}_t \boldsymbol{z}'_t \boldsymbol{e}_i - 1)$$

where $\boldsymbol{e}_i = (\mathbf{0}'_{i-1}, 1, \mathbf{0}'_{N-i})'$, $i = 1, ..., N$, and $\mathbf{0}_0$ is an empty set. The remaining blocks under H_0 equal

$$\begin{aligned} \boldsymbol{s}_t(\boldsymbol{\rho}_{Aj}) &= -(1/2)\frac{\partial \text{vec}(\boldsymbol{P}_{At})'}{\partial \boldsymbol{\rho}_{Aj}}\{\text{vec}(\boldsymbol{P}^{-1}_{(A0)}) - (\boldsymbol{P}^{-1}_{(A0)} \otimes \boldsymbol{P}^{-1}_{(A0)})\text{vec}(\boldsymbol{z}_t \boldsymbol{z}'_t)\} \\ &= -(t/T)^j \frac{\partial \text{vec}(\boldsymbol{P}_{(Aj)})'}{2\partial \boldsymbol{\rho}_{Aj}}\{\text{vec}(\boldsymbol{P}^{-1}_{(A0)}) - (\boldsymbol{P}^{-1}_{(A0)} \otimes \boldsymbol{P}^{-1}_{(A0)})\text{vec}(\boldsymbol{z}_t \boldsymbol{z}'_t)\} \end{aligned}$$

$j = 0, 1, 2$, where $\partial \text{vec}(\boldsymbol{P}_{(Aj)})'/\partial \boldsymbol{\rho}_{Aj}$ consists of zeroes and ones. Consequently, the $N^2 \times 3N(N-1)/2$ matrix $\partial \text{vec}(\boldsymbol{P}_{At})/\partial \boldsymbol{\rho}'_A$ equals

$$\frac{\partial \text{vec}(\boldsymbol{P}_{At})}{\partial \boldsymbol{\rho}'_A} = \left[\begin{array}{ccc} \frac{\partial \text{vec}(\boldsymbol{P}_{(A0)})}{\partial \boldsymbol{\rho}'_{A0}} & (t/T)\frac{\partial \text{vec}(\boldsymbol{P}_{(A1)})}{\partial \boldsymbol{\rho}'_{A1}} & (t/T)^2\frac{\partial \text{vec}(\boldsymbol{P}_{(A2)})}{\partial \boldsymbol{\rho}'_{A2}} \end{array} \right].$$

The information matrix for observation t under H_0 is quite similar to but simpler than the corresponding one in Silvennoinen and Teräsvirta (2017a). In order to give the matrix a proper expression, we need the commutation matrix $\boldsymbol{K}$, an $N^2 \times N^2$ matrix whose (i, j) block equals $\boldsymbol{e}_j\boldsymbol{e}'_i$, that is, $[\boldsymbol{K}]_{ij} = \boldsymbol{e}_j\boldsymbol{e}'_i$, see for example Lütkepohl (1996, pp. 115–118). Let g^0_{it} equal g_{it} evaluated under H_0 and, similarly, $\partial g^0_{it}/\partial \boldsymbol{\theta}_{gi} = \partial g_{it}/\partial \boldsymbol{\theta}_{gi}|_{\text{H}_0}$. The matrix is defined through the following lemma.

Lemma 2 *The expectations of the four blocks of the information matrix at (rescaled) time t/T under $H_0 : \boldsymbol{\rho}_{A1} = \boldsymbol{\rho}_{A2} = \boldsymbol{0}_{N(N-1)/2}$ are*

$$\boldsymbol{B}^0_t = \mathsf{E}\boldsymbol{s}_t(\boldsymbol{\theta}^0)\boldsymbol{s}'_t(\boldsymbol{\theta}^0) = \mathsf{E}\left[\begin{array}{cc} s_t(\boldsymbol{\theta}^0_g)\boldsymbol{s}'_t(\boldsymbol{\theta}^0_g) & s_t(\boldsymbol{\theta}^0_g)\boldsymbol{s}'_t(\boldsymbol{\rho}_A) \\ \boldsymbol{s}_t(\boldsymbol{\rho}_A)s'_t(\boldsymbol{\theta}^0_g) & \boldsymbol{s}_t(\boldsymbol{\rho}_A)s'_t(\boldsymbol{\rho}_A) \end{array} \right]$$

where the (i, j) sub-block of $\boldsymbol{B}_{11} = \mathsf{E}\boldsymbol{s}_t(\boldsymbol{\theta}^0_g)\boldsymbol{s}'_t(\boldsymbol{\theta}^0_g)$, $i \neq j$, equals

$$\begin{aligned} [\boldsymbol{B}_{11}]_{ij} &= \mathsf{E}\boldsymbol{s}_t(\boldsymbol{\theta}^0_{gi})\boldsymbol{s}'_t(\boldsymbol{\theta}^0_{gj}) \\ &= \frac{1}{4g^0_{it}g^0_{jt}}\frac{\partial g^0_{it}}{\partial \boldsymbol{\theta}_{gi}}\frac{\partial g^0_{jt}}{\partial \boldsymbol{\theta}'_{gj}}\boldsymbol{e}'_i\boldsymbol{P}^{-1}_{(A0)}\boldsymbol{e}_j\boldsymbol{e}'_i\boldsymbol{P}_{(A0)}\boldsymbol{e}_j. \end{aligned} \quad (16)$$

When $i = j$,

$$\begin{aligned} [\boldsymbol{B}_{11}]_{ii} &= \mathsf{E}\boldsymbol{s}_t(\boldsymbol{\theta}^0_{gi})\boldsymbol{s}'_t(\boldsymbol{\theta}^0_{gi}) \\ &= \frac{1}{4(g^0_{it})^2}\frac{\partial g^0_{it}}{\partial \boldsymbol{\theta}_{gi}}\frac{\partial g^0_{it}}{\partial \boldsymbol{\theta}'_{gi}}(1 + \boldsymbol{e}'_i\boldsymbol{P}^{-1}_{(A0)}\boldsymbol{e}_i). \end{aligned} \quad (17)$$

Furthermore, the (i, j) sub-block of $\mathsf{E}s_t(\boldsymbol{\theta}^0_g)\boldsymbol{s}'_t(\boldsymbol{\rho}_A)$ equals

$$\begin{aligned} [\boldsymbol{B}_{12}]_{ij} &= \mathsf{E}\boldsymbol{s}_t(\boldsymbol{\theta}^0_{gi})\boldsymbol{s}'_t(\boldsymbol{\rho}_{Aj}) \\ &= \frac{1}{4}(t/T)^j\frac{\partial g^0_{it}}{\partial \boldsymbol{\theta}_{gi}}\{(\boldsymbol{e}_i \otimes \boldsymbol{e}_i)'(\boldsymbol{P}^{-1}_{(A0)} \otimes \boldsymbol{I}_N) \\ &\quad +(\boldsymbol{e}_i \otimes \boldsymbol{e}_i)'(\boldsymbol{I}_N \otimes \boldsymbol{P}^{-1}_{(A0)}\}\frac{\partial vec(\boldsymbol{P}_{At})}{\partial \boldsymbol{\rho}'_A} \end{aligned}$$

$i = 1, ..., N$; $j = 0, 1, 2$, and

$$\begin{aligned} [\boldsymbol{B}_{22}]_{ij} &= \mathsf{E}\boldsymbol{s}_t(\boldsymbol{\rho}_{Ai})\boldsymbol{s}'_t(\boldsymbol{\rho}_{Aj}) \\ &= \frac{1}{4}(t/T)^{i+j}\frac{\partial vec(\boldsymbol{P}_{Ai})'}{\partial \boldsymbol{\rho}_A}\boldsymbol{M}_A\frac{\partial vec(\boldsymbol{P}_{Aj})}{\partial \boldsymbol{\rho}'_A} \end{aligned}$$

$i,j=0,1,2$, where

$$\boldsymbol{M}_A = \boldsymbol{P}_{(A0)}^{-1} \otimes \boldsymbol{P}_{(A0)}^{-1} + (\boldsymbol{P}_{(A0)}^{-1} \otimes \boldsymbol{I}_N)\boldsymbol{K}(\boldsymbol{P}_{(A0)}^{-1} \otimes \boldsymbol{I}_N).$$

Proof. See the appendix of Silvennoinen and Teräsvirta (2005), or Silvennoinen and Teräsvirta (2017a).

Note that the elements of the submatrix $\partial\text{vec}(\boldsymbol{P}_{At})/\partial\boldsymbol{\rho}_{A0}$ consist of zeroes and ones, so this matrix is time-invariant. The information matrix is defined in the next theorem:

Theorem 3 *The blocks of the information matrix of the log-likelihood* (15) *are*

$$[\boldsymbol{B}_{11}]_{ij} = \frac{1}{4}\int_0^1 \frac{1}{g_{ir}^0 g_{jr}^0}\frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}_{gi}}\frac{\partial g_{jr}^0}{\partial \boldsymbol{\theta}_{gj}'} dr \boldsymbol{e}_i' \boldsymbol{P}_{(A0)}^{-1}\boldsymbol{e}_j \boldsymbol{e}_i' \boldsymbol{P}_{(A0)}\boldsymbol{e}_j \tag{18}$$

for $i \neq j$, and

$$[\boldsymbol{B}_{11}]_{ii} = \frac{1}{4}\int_0^1 \frac{1}{(g_{ir}^0)^2}\frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}_{gi}}\frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}_{gi}'} dr (1 + \boldsymbol{e}_i' \boldsymbol{P}_{(A0)}^{-1}\boldsymbol{e}_i) \tag{19}$$

for $i=j$. Furthermore,

$$\begin{aligned}[\boldsymbol{B}_{12}]_{ij} &= \frac{1}{4}\int_0^1 r^j \frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}_{gi}} dr \{(\boldsymbol{e}_i \otimes \boldsymbol{e}_i)'(\boldsymbol{P}_{(A0)}^{-1} \otimes \boldsymbol{I}_N) \\ &\quad + (\boldsymbol{e}_i \otimes \boldsymbol{e}_i)'(\boldsymbol{I}_N \otimes \boldsymbol{P}_{(A0)}^{-1})\}\frac{\partial vec(\boldsymbol{P}_{Aj})'}{\partial \boldsymbol{\rho}_A}\end{aligned}$$

$i=1,...,N$; $j=0,1,2$, and

$$\begin{aligned}[\boldsymbol{B}_{22}]_{ij} &= \frac{1}{4}\int_0^1 r^{i+j} dr \frac{\partial vec(\boldsymbol{P}_{Ai})}{\partial \boldsymbol{\rho}_{Ai}'}\boldsymbol{M}_A \frac{\partial vec(\boldsymbol{P}_{Aj})'}{\partial \boldsymbol{\rho}_{Aj}} \\ &= \frac{1}{4(i+j+1)}\frac{\partial vec(\boldsymbol{P}_{Ai})}{\partial \boldsymbol{\rho}_{Ai}'}\boldsymbol{M}_A \frac{\partial vec(\boldsymbol{P}_{Aj})'}{\partial \boldsymbol{\rho}_{Aj}}.\end{aligned} \tag{20}$$

Proof. Similar to a corresponding proof in Silvennoinen and Teräsvirta (2017a) and therefore omitted.

In order to define the test statistic, let

$$\boldsymbol{B}_{12\cdot j} = ([\boldsymbol{B}_{12}']_{1j}, \ldots, [\boldsymbol{B}_{12}']_{Nj})$$

$j=0,1,2$,

$$\boldsymbol{B}_{11}^0 = \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12\cdot 0} \\ \boldsymbol{B}_{12\cdot 0}' & [\boldsymbol{B}_{22}]_{00} \end{bmatrix},$$

$$\boldsymbol{B}_{12} = \begin{bmatrix} \boldsymbol{B}_{12\cdot 1} & \boldsymbol{B}_{12\cdot 2} \\ [\boldsymbol{B}_{22}]_{01} & [\boldsymbol{B}_{22}]_{02} \end{bmatrix}$$

and

$$\boldsymbol{B}_{22} = \begin{bmatrix} [\boldsymbol{B}_{22}]_{11} & [\boldsymbol{B}_{22}]_{12} \\ [\boldsymbol{B}'_{22}]_{12} & [\boldsymbol{B}_{22}]_{22} \end{bmatrix}$$

Next define

$$\widehat{\boldsymbol{x}}_{jt} = -\frac{1}{2}\frac{\partial \text{vec}(\boldsymbol{P}_{Aj})'}{\partial \boldsymbol{\rho}_{Aj}}\{\text{vec}(\widehat{\boldsymbol{P}}^{-1}_{(A0)}) - (\widehat{\boldsymbol{P}}^{-1}_{(A0)} \otimes \widehat{\boldsymbol{P}}^{-1}_{(A0)})\text{vec}(\widehat{\boldsymbol{z}}_t\widehat{\boldsymbol{z}}'_t)\}$$

$j = 1, 2$, where $\widehat{\boldsymbol{z}}_t$ and $\widehat{\boldsymbol{P}}_{(A0)}$ equal $\boldsymbol{z}_t$ and $\boldsymbol{P}_{(A0)}$ estimated under H_0, respectively. The test statistic

$$\begin{aligned} LM_T &= T(\frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{x}}'_{1t}, \frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{x}}'_{2t}) \\ &\quad \times\{\boldsymbol{B}_{22} - \boldsymbol{B}'_{12}(\boldsymbol{B}^0_{11})^{-1}(\boldsymbol{B}_{12}\}^{-1}(\frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{x}}'_{1t}, \frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{x}}'_{2t})' \end{aligned} \quad (21)$$

has an asymptotic χ^2-distribution with $N(N-1)$ degrees of freedom when H_0 holds. To make the test statistic operational, the sub-blocks of the information matrix in (21) have to be replaced by consistent plug-in estimators.

The number of degrees of freedom in this test quickly becomes large with increasing N. One way of restricting this growth would be to assume that under the alternative only the eigenvalues of the correlation matrix are changing over time. The alternative would be a correlation matrix only if all correlations are identical, but an LM test can nevertheless be built on this assumption. Write the correlation matrix as $\boldsymbol{P}_t = \boldsymbol{Q}_t\boldsymbol{\Lambda}_t\boldsymbol{Q}'_t$, where $\boldsymbol{P}_t$ is defined as in (6), $\boldsymbol{\Lambda}_t$ is the matrix of eigenvalues and $\boldsymbol{Q}_t$ contains the corresponding eigenvectors. Simplify this by assuming $\boldsymbol{Q}_t = \boldsymbol{Q}$ and approximate $\boldsymbol{\Lambda}_t$, the eigenvalue matrix of (6), by $\boldsymbol{\Psi}_t = \sum_{k=0}^{K}\boldsymbol{\Psi}_k(t/T)^k$. Under the null hypothesis, $K = 0$. The resulting test statistic is derived and its small-sample properties studied in Silvennoinen and Teräsvirta (2017b).

5 Estimation of the MTV–STCC–GARCH model

After specifying the deterministic components of the model, both in GARCH equations and correlations, one can estimate the complete model with conditional heteroskedasticity included. The log-likelihood of the MTV–STCC–GARCH model has the form

$$\ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta},\mathcal{F}_{t-1}) = -(1/2)\sum_{i=1}^{N}\ln g_{it} - (1/2)\sum_{i=1}^{N}\ln h_{it} - (1/2)\ln|\boldsymbol{P}_t|$$
$$-(1/2)\boldsymbol{\varepsilon}_t'\{\boldsymbol{S}_t\boldsymbol{D}_t\boldsymbol{P}_t\boldsymbol{D}_t\boldsymbol{S}_t\}^{-1}\boldsymbol{\varepsilon}_t. \tag{22}$$

Since $\boldsymbol{D}_t\boldsymbol{z}_t = \boldsymbol{S}_t^{-1}\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}/g_{1t}^{1/2},...,\varepsilon_{Nt}/g_{Nt}^{1/2})'$, it is seen from (4) that the conditional variance components in (22) are

$$h_{it} = \alpha_{i0} + \alpha_{i1}\phi_{i,t-1}^2 + \kappa_{i1}I(\phi_{t-1}<0)\phi_{i,t-1}^2 + \beta_{i1}h_{i,t-1} \tag{23}$$

$i = 1,...,N$. We make the following assumptions, see Silvennoinen and Teräsvirta (2017a):

AN1. In (4), $\alpha_{i0} > 0$, $\alpha_{i0} + \kappa_{i1}/2 > 0$, $\beta_{i1} \geq 0$, and $\alpha_{i1} + \kappa_{i1}/2 + \beta_{i1} < 1$ for $i = 1,...,N$.

AN2. The parameter subspaces $\{\alpha_{i0} \times \kappa_i \times \alpha_i \times \beta_i\}$, $i = 1,...,N$, are compact, the whole space Θ_h is compact, and the true parameter value $\boldsymbol{\theta}_h^0$ is an interior point of Θ_h.

AN3. $\mathsf{E}\phi_{it}^4 < \infty$.

AN1 is the necessary and sufficient weak stationarity condition for the ith first-order GJR-GARCH equation. Conditions AN2 and AN3 are standard regularity conditions required for proving asymptotic normality of maximum likelihood estimators of $\boldsymbol{\theta}_{hi}$, $i = 1,...,N$.

These assumptions are sufficient for the maximum likelihood estimators of the GARCH parameters in single-equation GARCH models to be consistent and asymptotically normal. Rewrite (22) indicating the parameters in the relevant functions:

$$\begin{aligned}&\ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta},\mathcal{F}_{t-1})\\ =\ &-(1/2)\sum_{i=1}^{N}\ln g_{it}(\boldsymbol{\theta}_{gi}) - (1/2)\sum_{i=1}^{N}\ln h_{it}(\boldsymbol{\theta}_{hi}) - (1/2)\ln|\boldsymbol{P}_t(\boldsymbol{\theta}_P)|\\ &-(1/2)\boldsymbol{\varepsilon}_t'\{\boldsymbol{S}_t(\boldsymbol{\theta}_g)\boldsymbol{D}_t(\boldsymbol{\theta}_g,\boldsymbol{\theta}_h)\boldsymbol{P}_t(\boldsymbol{\theta}_P)\boldsymbol{D}_t(\boldsymbol{\theta}_g,\boldsymbol{\theta}_h)\boldsymbol{S}_t(\boldsymbol{\theta}_g)\}^{-1}\boldsymbol{\varepsilon}_t.\end{aligned}$$

The parameters are estimated in turn: first estimate $\boldsymbol{\theta}_{gi}$ to obtain starting-values to joint estimation of $\boldsymbol{\theta}_g$ and $\boldsymbol{\theta}_P$. This is done assuming $h_{it}(\boldsymbol{\theta}_{hi}) \equiv 1$, $i = 1,...,N$. Amado and Teräsvirta (2013) showed in the single-equation GJR–GARCH case that under regularity conditions the maximum likelihood estimator of $\boldsymbol{\theta}_{gi}$ is consistent and asymptotically normal. Silvennoinen and Teräsvirta (2017a) generalised this result to MTV models. That means that joint estimation of $\boldsymbol{\theta}_g$ and $\boldsymbol{\theta}_P$ by maximum likelihood produces consistent estimates of these parameter vectors. If $\widehat{\boldsymbol{\theta}}_g$ and $\widehat{\boldsymbol{\theta}}_P$ are consistent and Assumptions AN1, AN2 and AN3 hold, then by Theorem (3.3) of Song, Fan and Kalbfleisch (2005), the maximum likelihood estimator of $\boldsymbol{\theta}_h$ is consistent and asymptotically normal. After estimating $\boldsymbol{\theta}_h$, the parameter vectors $\boldsymbol{\theta}_g$ and $\boldsymbol{\theta}_P$ are re-estimated. Iteration continues until convergence. Song et al. (2005) showed that the final maximum likelihood estimator of $\boldsymbol{\theta}$ is consistent and asymptotically normal. A more detailed description of the maximisation by parts applied to the present situation will be provided, see also Silvennoinen and Teräsvirta (2017a).

6 Simulations of test statistics

Preliminary pilot simulations were carried out with a sample size of 1000 observations. This constitutes a small sample in relation to the recommended sample sizes in GARCH literature. The first experiment assumes $h_{it} = 1$ and $g_{it} = \delta_{i0}$ (constant) for all $i = 1, \ldots, N$. Letting the level of constant correlation vary from -0.75 to 0.75, and considering a range of system dimension from bivariate to $N = 20$, the test of constant correlations shows no size distortion. The same result is obtained when the conditional variances are correctly specified (but not necessarily constant).

The second experiment then investigates the impact of neglecting the conditional heteroskedasticity. In this case, the size of the test shows no size distortion for low and moderate levels of correlation. However, as correlation increases past 0.5, small oversizing is observed. The situation gets more severe the higher the general level of correlation in the system. This indicates that size correction may be needed, and the proposed steps follow the approach in Amado, Silvennoinen and Teräsvirta (2017).

References

Amado, C., Silvennoinen, A. and Teräsvirta, T.: 2017, Modelling and forecasting WIG20 daily returns, *Central European Journal of Economic Modelling and Econometrics* **9**, 173–200.

Amado, C. and Teräsvirta, T.: 2008, Modelling conditional and unconditional 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.: 2013, Modelling volatility by variance decomposition, *Journal of Econometrics* **175**, 153–165.

Amado, C. and Teräsvirta, T.: 2014, Conditional correlation models of autoregressive conditional heteroscedasticity with nonstationary GARCH equations, *Journal of Business and Economic Statistics* **32**, 69–87.

Amado, C. and Teräsvirta, T.: 2017, Specification and testing of multiplicative time-varying GARCH models with applications, *Econometric Reviews* **36**, 421–446.

Bauwens, L., Laurent, S. and Rombouts, J. V. K.: 2006, Multivariate GARCH models: A survey, *Journal of Applied Econometrics* **21**, 79–109.

Berben, R.-P. and Jansen, W. J.: 2005, Comovement in international equity markets: A sectoral view, *Journal of International Money and Finance* **24**, 832–857.

Bollerslev, T.: 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **31**, 307–327.

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.

Bollerslev, T., Engle, R. F. and Wooldridge, J. M.: 1988, A capital asset pricing model with time-varying covariances, *The Journal of Political Economy* **96**, 116–131.

Box, G. E. P. and Jenkins, G. M.: 1970, *Time Series Analysis, Forecasting And Control*, Holden-Day, San Francisco.

Brownlees, C. T. and Gallo, G. M.: 2010, Comparison of volatility measures: A risk management perspective, *Journal of Financial Econometrics* **8**, 29–56.

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 Rangel, J. G.: 2008, The spline-GARCH model for low-frequency volatility and its global macroeconomic causes, *Review of Financial Studies* **21**, 1187–1222.

Feng, Y.: 2004, Simultaneously modeling conditional heteroskedasticity and scale change, *Econometric Theory* **20**, 563–596.

Glosten, L. W., 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.

Godfrey, L. G.: 1988, *Misspecification Tests in Econometrics*, Cambridge University Press, Cambridge.

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.

Silvennoinen, A. and Teräsvirta, T.: 2005, Multivariate autoregressive conditional heteroskedasticity with smooth transitions in conditional correlations, SSE/EFI Working Paper Series in Economics and Finance No. 577.

Silvennoinen, A. and Teräsvirta, T.: 2009a, Modelling multivariate autoregressive conditional heteroskedasticity with the double smooth transition conditional correlation GARCH model, *Journal of Financial Econometrics* **7**, 373–411.

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

Silvennoinen, A. and Teräsvirta, T.: 2015, Modeling conditional correlations of asset returns: A smooth transition approach, *Econometric Reviews* **34**, 174–197.

Silvennoinen, A. and Teräsvirta, T.: 2016, Testing constancy of unconditional variance in volatility models by misspecification and specification tests, *Studies in Nonlinear Dynamics and Econometrics* **20**, 347–364.

Silvennoinen, A. and Teräsvirta, T.: 2017a, Consistency and asymptotic normality of maximum likelihood estimators of the multiplicative time-varying smooth transition correlation GARCH model, *Working paper*, CREATES Research Paper 2017, Aarhus University.

Silvennoinen, A. and Teräsvirta, T.: 2017b, A parsimonious test of constancy of a positive definite correlation matrix in a multivariate time-varying GARCH model, *work in progress*, Queensland University of Technology, Brisbane.

Song, P. X., Fan, Y. and Kalbfleisch, J. D.: 2005, Maximization by parts in likelihood inference, *Journal of the American Statistical Association* **100**, 1145–1158.

Teräsvirta, T.: 1994, Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association* **89**, 208–218.

Teräsvirta, T., Tjøstheim, D. and Granger, C. W. J.: 2010, *Modelling Nonlinear Economic Time Series*, Oxford University Press, Oxford.

Tse, Y. K. and Tsui, K. C.: 2002, A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations, *Journal of Business and 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.

Appendices

A Proofs

Proof of Lemma 1. The 'sample' information matrix with T observations equals

$$\frac{1}{4T}\sum_{t=1}^{T}\boldsymbol{B}_t = \frac{1}{4T}\sum_{t=1}^{T}\left[\begin{array}{cc} \boldsymbol{B}_{11t} & \boldsymbol{B}_{12t} \\ \boldsymbol{B}_{21t} & \boldsymbol{B}_{22t} \end{array}\right].$$

Consider the $(1,2)$ element of $\boldsymbol{B}_{11(T)} = (1/T)\sum_{t=1}^{T}\boldsymbol{B}_{11t}$:

$$[\boldsymbol{B}_{11(T)}]_{12} = (1/T)\sum_{t=1}^{T}(g_t^0)^{-2}G_1^0(t/T)$$

which is an average of T values of the logistic cumulative distribution function. Let $[Tr] = t$ be the integer closest to t. Then

$$\begin{aligned}(1/T)\sum_{t=1}^{T} G_1^0(t/T) &= (1/T)\sum_{t=1}^{T}\int_{t/T}^{(t+1)/T} (g_{[Tr]}^{*0})^{-2} G_1^0([Tr]/T)\mathrm{d}r \\ &= \int_{1/T}^{(T+1)/T} (g_{[Tr]}^{*0})^{-2} G_1^0([Tr]/T)\mathrm{d}r \to \int_0^1 (g_r^{*0})^{-2} G_1^0(r)\mathrm{d}r\end{aligned}$$

as $T \to \infty$. The other elements of $\boldsymbol{B}_{11} = \lim_{T\to\infty} \boldsymbol{B}_{11(T)}$ are obtained similarly. The results are:

$$\boldsymbol{B}_{11} = \frac{1}{4}\begin{bmatrix} \int_0^1 (g_r^0)^{-2}\mathrm{d}r & \int_0^1 (g_r^0)^{-2}G_1^0(r)\mathrm{d}r & \int_0^1 (g_r^0)^{-2}G_{1\gamma}^0(r)\mathrm{d}r & \int_0^1 (g_r^0)^{-2}G_{1c}^0(r)\mathrm{d}r \\ & \int_0^1 (g_r^0)^{-2}G_1^0(r)^2\mathrm{d}r & \int_0^1 (g_r^0)^{-2}G_1^0(r)G_{1\gamma}^0(r)\mathrm{d}r & \int_0^1 (g_r^0)^{-2}G_1^0(r)G_{1c}^0(r)\mathrm{d}r \\ & & \int_0^1 (g_r^0)^{-2}G_{1\gamma}^0(r)^2\mathrm{d}r & \int_0^1 (g_r^0)^{-2}G_{1\gamma}^0(r)G_{1c}^0(r)\mathrm{d}r \\ & & & \int_0^1 (g_r^0)^{-2}G_{1c}^0(r)^2\mathrm{d}r \end{bmatrix}$$

Accordingly,

$$\boldsymbol{B}_{12} = \frac{1}{4}\begin{bmatrix} \int_0^1 r(g_r^0)^{-2}\mathrm{d}r & \int_0^1 r^2(g_r^0)^{-2}\mathrm{d}r & \int_0^1 r^3(g_r^0)^{-2}\mathrm{d}r \\ \int_0^1 r(g_r^0)^{-2}G_1^0(r)\mathrm{d}r & \int_0^1 r^2(g_r^0)^{-2}G_1^0(r)\mathrm{d}r & \int_0^1 r^3(g_r^0)^{-2}G_1^0(r)\mathrm{d}r \\ \int_0^1 r(g_r^0)^{-2}G_{1\gamma}^0(r)\mathrm{d}r & \int_0^1 r^2(g_r^0)^{-2}G_{1\gamma}^0(r)\mathrm{d}r & \int_0^1 r^3(g_r^0)^{-2}G_{1\gamma}^0(r)\mathrm{d}r \\ \int_0^1 r(g_r^0)^{-2}G_{1c}^0(r)\mathrm{d}r & \int_0^1 r^2(g_r^0)^{-2}G_{1c}^0(r)\mathrm{d}r & \int_0^1 r^3(g_r^0)^{-2}G_{1c}^0(r)\mathrm{d}r \end{bmatrix}$$

and, finally,

$$\boldsymbol{B}_{22} = \lim_{T\to\infty}\frac{1}{4T}\sum_{t=1}^{T}(g_t^0)^{-2}\boldsymbol{t}_t\boldsymbol{t}_t' = \frac{1}{4}\begin{bmatrix} \int_0^1 r^2(g_r^0)^{-2}\mathrm{d}r & \int_0^1 r^3(g_r^0)^{-2}\mathrm{d}r & \int_0^1 r^4(g_r^0)^{-2}\mathrm{d}r \\ & \int_0^1 r^4(g_r^0)^{-2}\mathrm{d}r & \int_0^1 r^5(g_r^0)^{-2}\mathrm{d}r \\ & & \int_0^1 r^6(g_r^0)^{-2}\mathrm{d}r \end{bmatrix}.$$

This concludes the proof. ■