Modelling Multivariate Autoregressive Conditional Heteroskedasticity with the Double Smooth Transition Conditional Correlation GARCH Model

Annastiina Silvennoinen*

School of Finance and Economics, University of Technology Sydney, Australia

Timo Teräsvirta†

CREATES, School of Economics and Management, University of Aarhus, Denmark

and

Department of Economic Statistics, Stockholm School of Economics, Sweden

August 2008

Abstract

In this paper we propose a multivariate GARCH model with a time-varying conditional correlation structure. The new Double Smooth Transition Conditional Correlation GARCH model extends the Smooth Transition Conditional Correlation GARCH model of Silvennoinen and Teräsvirta (2005) by including another variable according to which the correlations change smoothly between states of constant correlations. A Lagrange multiplier test is derived to test the constancy of correlations against the DSTCC–GARCH model, and another one to test for another transition in the STCC–GARCH framework. In addition, other specification tests, with the aim of aiding the model building procedure, are considered. Analytical expressions for the test statistics and the required derivatives are provided. Applying the model to the stock and bond futures data, we discover that the correlation pattern between them has dramatically changed around the turn of the century. The model is applied also to a selection of world stock indices, and we find evidence for an increasing degree of integration in the capital markets.

JEL classification: C12; C32; C51; C52; G1

Key words: Multivariate GARCH; Constant conditional correlation; Dynamic conditional correlation; Return comovement; Variable correlation GARCH model; Volatility model evaluation

*e-mail: annastiina.silvennoinen@uts.edu.au

†e-mail: tterasvirta@econ.au.dk

Acknowledgements. This research has been supported by the Jan Wallander and Tom Hedelius Foundation, Grant No. J03–41 and P2005–0033:1, the Danish National Research Foundation, OP Bank Group Research Foundation, The Foundation for Promoting Finnish Equity Markets, and and Yrjö Jahnsson Foundation. We would like to thank Pierre Giot for valuable comments, and Robert Engle for providing the data for the first part of the empirical section. Part of the work was done when the first author was visiting CORE whose kind hospitality is gratefully acknowledged. Our special thanks go to Luc Bauwens for making the visit possible. The visit was financially supported by a Marie Curie Fellowship. Material from this paper has been presented at the Conference on Multivariate Modelling in Finance and Risk Management, Sandbjerg Manor, Denmark, June 2006, the Econometric Society Australasian Meeting, Alice Springs, July 2006, Conference on Multivariate Volatility Modelling, Faro, Portugal, October 2007, and in seminars at Université catholique de Louvain, Louvain-la-Neuve, Cracow University of Economics, WISE, Xiamen University, and Ente Einaudi, Rome. Comments of participants in these occasions are gratefully acknowledged. The responsibility for any errors and shortcomings in this paper remains ours.

1 Introduction

Multivariate financial time series have been subject to many modelling proposals incorporating conditional heteroskedasticity, originally introduced by Engle (1982) in a univariate context. For reviews, the reader is referred to recent surveys on multivariate GARCH models by Bauwens, Laurent, and Rombouts (2006) and Silvennoinen and Teräsvirta (2008). One may model the time-varying covariances directly. Examples of this are VEC and BEKK models, as well as factor GARCH ones, all discussed in these two surveys. Alternatively, one may model the conditional correlations. The simplest approach builds on the assumption that the correlations are time-invariant. Although the Constant Conditional Correlation (CCC) GARCH model of Bollerslev (1990) is attractive to the practitioner due to its interpretable parameters and easy estimation, its fundamental assumption that correlations remain constant over time has often been found unrealistic. In order to remedy this problem, Tse and Tsui (2002) and Engle (2002) introduced models with dynamic conditional correlations called the VC–GARCH and the DCC–GARCH model, respectively, that impose GARCH-type structure on the correlations. By construction, these models have the property that the variation in correlations is mainly due to the size and the sign of the shock of the previous time period.

An interesting model combining aspects from both the CCC–GARCH and the DCC–GARCH has been suggested by Pelletier (2006). The author introduces a regime switching correlation structure driven by an unobserved state variable following a first-order Markov chain. The regime switching model asserts that the correlations remain constant in each regime and the change between the states is abrupt and governed by transition probabilities. Thus the factors affecting the correlations remain latent and are not observed.

In a recent paper, Silvennoinen and Teräsvirta (2005) introduced the Smooth Transition Conditional Correlation (STCC) GARCH model.[1] In this model the correlations vary smoothly between two extreme states of constant correlations and the dynamics are driven by an observable transition variable. The transition variable can be chosen by the modeller, and the model combined with tests of constant correlations constitutes a useful tool for modellers interested in characterizing the dynamic structure of the correlations. This paper extends the STCC–GARCH model into one that allows variation in conditional correlations to be controlled by two observable transition variables instead of only one. This makes it possible, for example, to nest the Berben and Jansen (2005a) model with time as the transition variable in this general double-transition model.

It has become a widely accepted feature of financial data that volatile periods in financial markets are related to an increase in correlations among assets. However, as pointed out by Boyer, Gibson, and Loretan (1999) and Longin and Solnik (2001), in many studies this hypothesis is not investigated properly and the reported results may be misleading. In fact, the latter authors report evidence that in international markets correlations are not related to market volatility as measured in large absolute returns, but only to large negative returns, or to the market trend. Our modelling framework allows the researcher to easily explore such possibilities by first testing the relevance of a model with a transition variable corresponding to the hypothesis to be tested and, in case of rejection, estimating the model to find out the direction of change in correlations controlled by that variable; see Silvennoinen and Teräsvirta (2005) for an example.

The paper is organized as follows. In Section 2 the new DSTCC–GARCH model is introduced and its estimation discussed. Section 3 gives the testing procedures and Section 4 reports simulation experiments on the tests. In Section 5 we compare the estimated correlations from

[1] A bivariate special case of the STCC–GARCH model was coincidentally introduced in Berben and Jansen (2005a).

our new model, the DCC–GARCH, and the semiparametric MGARCH one, using S&P 500 index and long term bond futures data. In Section 6 we apply our model to a set of five international stock market indices, namely French CAC 40, German DAX, FTSE 100 from UK, Hong Kong Hang Seng, and Japanese Nikkei 225 from December 1990 until April 2006. Finally, Section 7 concludes. The detailed derivations of the tests can be found in the Appendix.

2 The Double Smooth Transition Conditional Correlation GARCH model

2.1 The general multivariate GARCH model

Consider the following stochastic N-dimensional vector process with the standard representation

$$\boldsymbol{y}_t = E\left[\boldsymbol{y}_t \mid \mathcal{F}_{t-1}\right] + \boldsymbol{\varepsilon}_t \qquad t = 1, 2, \ldots, T \tag{1}$$

where $\mathcal{F}_{t-1}$ is the sigma-field generated by all the information until time $t-1$. Each of the univariate error processes has the specification

$$\varepsilon_{it} = h_{it}^{1/2} z_{it}$$

where the errors z_{it} form a sequence of independent random variables with mean zero and variance one, for each $i = 1, \ldots, N$. The conditional variance h_{it} follows a univariate GARCH process, for example that of Glosten, Jagannathan, and Runkle (1993)

$$h_{it} = \alpha_{i0} + \sum_{j=1}^{q} \alpha_{ij}\varepsilon_{i,t-j}^2 + \sum_{j=1}^{p} \beta_{ij} h_{i,t-j} + \sum_{j=1}^{q} \delta_{ij}(\varepsilon_{i,t-j}^-)^2 \tag{2}$$

where $\varepsilon_{it}^- = \min(0, \varepsilon_{it})$,with the non-negativity and stationarity restrictions imposed. For $\delta_{ij} \neq 0$ the model allows the conditional volatility to respond asymmetrically to positive and negative shocks. The results in this paper are derived using (2) with $p = q = 1$ to account for the conditional heteroskedasticity. It is straightforward to modify them to allow for a higher-order or some other type of GARCH process. The conditional covariances of the vector $\boldsymbol{z}_t$ are given by

$$E\left[\boldsymbol{z}_t \boldsymbol{z}_t' \mid \mathcal{F}_{t-1}\right] \quad = \quad \boldsymbol{P}_t. \tag{3}$$

Furthermore, the standardized errors $\boldsymbol{\eta}_t = \boldsymbol{P}_t^{-1/2}\boldsymbol{z}_t \sim iid(\mathbf{0}, \boldsymbol{I}_N)$. Since z_{it} has unit variance for all i, $\boldsymbol{P}_t = [\rho_{ij,t}]$ is the conditional correlation matrix for the $\boldsymbol{\varepsilon}_t$ whose elements $\rho_{ij,t}$ are allowed to be time-varying for $i \neq j$. It will, however, be assumed that $\boldsymbol{P}_t \in \mathcal{F}_{t-1}$.

The conditional covariance matrix $\boldsymbol{H}_t = \boldsymbol{S}_t \boldsymbol{P}_t \boldsymbol{S}_t$, where $\boldsymbol{P}_t$ is the conditional correlation matrix as in equation (3), and $\boldsymbol{S}_t = diag(h_{1t}^{1/2}, .., h_{Nt}^{1/2})$ with elements defined in (2), is positive definite whenever the correlation matrix $\boldsymbol{P}_t$ is positive definite.

2.2 Smooth transitions in conditional correlations

The idea of introducing smooth transition in the conditional correlations is discussed in detail in Silvennoinen and Teräsvirta (2005) where a simple structure with one type of transition between

two states of constant correlations is introduced. Specifically, the STCC–GARCH model defines the time-varying correlation structure as

$$\boldsymbol{P}_t = (1 - G_t)\boldsymbol{P}_{(1)} + G_t\boldsymbol{P}_{(2)} \tag{4}$$

where the transition function $G_t = G(s_t; \gamma, c)$ is the logistic function

$$G_t = \left(1 + e^{-\gamma(s_t - c)}\right)^{-1}, \quad \gamma > 0 \tag{5}$$

that is bounded between zero and one. Furthermore, $\boldsymbol{P}_{(1)}$ and $\boldsymbol{P}_{(2)}$ represent the two extreme states of correlations between which the conditional correlations can vary over time according to the transition variable s_t. The two parameters in (5), γ and c, define the speed and location of the transition. When the transition variable has values less than c, the correlations are closer to the state defined by $\boldsymbol{P}_{(1)}$ than the one defined by $\boldsymbol{P}_{(2)}$. For $s_t > c$, the situation is the opposite. The parameter γ controls the smoothness of the transition between the two states. The closer γ is to zero the slower the transition. As $\gamma \to \infty$, the transition function eventually becomes a step function. The positive definiteness of $\boldsymbol{P}_t$ at each point in time is ensured by the requirement that the two correlation matrices $\boldsymbol{P}_{(1)}$ and $\boldsymbol{P}_{(2)}$ are positive definite. By defining the transition variable to be the calendar time, $s_t = t/T$, one arrives at the Time-Varying Smooth Transition Conditional Correlation (TVCC) GARCH model. A bivariate version of this special case was introduced by Berben and Jansen (2005a).

We extend the original STCC–GARCH model by allowing the conditional correlations to vary according to two transition variables. The time-varying correlation structure in the Double Smooth Transition Conditional Correlation (DSTCC) GARCH model is imposed through the following equations:

$$\begin{aligned} \boldsymbol{P}_t &= (1 - G_{1t})\boldsymbol{P}_{(1)t} + G_{1t}\boldsymbol{P}_{(2)t} \\ \boldsymbol{P}_{(i)t} &= (1 - G_{2t})\boldsymbol{P}_{(i1)} + G_{2t}\boldsymbol{P}_{(i2)}, \quad i = 1, 2 \end{aligned} \tag{6}$$

where the transition functions are the logistic functions

$$G_{it} = \left(1 + e^{-\gamma_i(s_{it} - c_i)}\right)^{-1}, \quad \gamma_i > 0, \quad i = 1, 2 \tag{7}$$

and s_{it}, $i = 1, 2$, are transition variables that can be either stochastic or deterministic. The correlation matrix $\boldsymbol{P}_t$ is thus a convex combination of four positive definite matrices, $\boldsymbol{P}_{(11)}$, $\boldsymbol{P}_{(12)}$, $\boldsymbol{P}_{(21)}$, and $\boldsymbol{P}_{(22)}$, each of which defines an extreme state of constant correlations. The positive definiteness of $\boldsymbol{P}_t$ at each point in time follows from the positive definiteness of these four matrices. In (7) the parameters γ_i and c_i determine the speed and the location of the transition i, $i = 1, 2$. The transition variables are chosen by the modeller. As in the STCC–GARCH model, the values of these variables are assumed to be known at time t. Possible choices are for instance functions of lagged elements of $\boldsymbol{y}_t$, or exogenous variables. When applying the model to stock return series one could consider functions of market indices or business cycle indicators, or simply time. If one of the transition variables is time, say $s_{2t} = t/T$, the model with correlation dynamics (6) is the Time-Varying Smooth Transition Conditional Correlation (TVSTCC) GARCH model. In this case it may be illustrative to write (6) as

$$\boldsymbol{P}_t = (1 - G_{2t})\left((1 - G_{1t})\boldsymbol{P}_{(11)} + G_{1t}\boldsymbol{P}_{(21)}\right) + G_{2t}\left((1 - G_{1t})\boldsymbol{P}_{(12)} + G_{1t}\boldsymbol{P}_{(22)}\right). \tag{8}$$

The role of the correlation matrices describing the constant states is easily seen from (8). At the beginning of the sample the correlations vary smoothly between the states defined by $\boldsymbol{P}_{(11)}$

and $\boldsymbol{P}_{(21)}$: when $s_{1t} < c_1$, the correlations are closer to the state in $\boldsymbol{P}_{(11)}$ than $\boldsymbol{P}_{(21)}$ whereas when $s_{1t} > c_1$, the situation is the opposite. As time evolves the correlations in $\boldsymbol{P}_{(11)}$ and $\boldsymbol{P}_{(21)}$ transform smoothly to the ones in $\boldsymbol{P}_{(12)}$ and $\boldsymbol{P}_{(22)}$, respectively. Therefore, at the end of the sample, s_{1t} shifts the correlations between these two matrices.

The specification (6) describes the correlation structure of the DSTCC–GARCH model in its fully general form. Imposing certain restrictions on the correlations give rise to numerous special cases; those will be discussed in Section 3. One restricted version, however, is worth discussing in detail. The effects of the two transition variables can be independent in a sense that the time-variation of the correlations due to one of the transition variables does not depend on the value of the other transition variable. This condition can be expressed as

$$\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(12)} = \boldsymbol{P}_{(21)} - \boldsymbol{P}_{(22)} \tag{9a}$$

or, equivalently,

$$\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(21)} = \boldsymbol{P}_{(12)} - \boldsymbol{P}_{(22)}. \tag{9b}$$

In terms of equation (8) these conditions imply that on the right-hand side of this equation the matrices with coefficients $\pm G_{1t}G_{2t}$ are eliminated. Furthermore, from (9a) and (9b) it follows that the difference between the extreme states described by one of the transition variables remains constant across all values of the other transition variable. In this case the dynamic conditional correlations of the DSTCC–GARCH model become

$$\boldsymbol{P}_t = (1 - G_{1t} - G_{2t})\boldsymbol{P}_{(11)} + G_{1t}\boldsymbol{P}_{(21)} + G_{2t}\boldsymbol{P}_{(12)}. \tag{10}$$

This parsimonious specification may prove useful when dealing with large systems because one only has to estimate three correlation matrices instead of four in an unrestricted DSTCC–GARCH model.

2.3 Estimation of the DSTCC–GARCH model

For maximum likelihood estimation of parameters we assume joint conditional normality of the errors:

$$\boldsymbol{z}_t \mid \mathcal{F}_{t-1} \sim N\left(\boldsymbol{0}, \boldsymbol{P}_t\right).$$

Denoting by $\boldsymbol{\theta}$ the vector of all the parameters in the model, the log-likelihood for observation t is

$$l_t\left(\boldsymbol{\theta}\right) = -\frac{N}{2}\log\left(2\pi\right) - \frac{1}{2}\sum_{i=1}^{N}\log h_{it} - \frac{1}{2}\log\left|\boldsymbol{P}_t\right| - \frac{1}{2}\boldsymbol{z}_t'\boldsymbol{P}_t^{-1}\boldsymbol{z}_t, \quad t = 1, \ldots, T \tag{11}$$

and maximizing $\sum_{t=1}^{T} l_t(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ yields the maximum likelihood estimator $\hat{\boldsymbol{\theta}}_T$.

For inference we assume that the asymptotic distribution of the ML-estimator is normal, that is,

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right) \xrightarrow{d} N\left(\boldsymbol{0}, \mathfrak{I}^{-1}(\boldsymbol{\theta}_0)\right)$$

where $\boldsymbol{\theta}_0$ is the true parameter and $\mathfrak{I}(\boldsymbol{\theta}_0)$ the population information matrix evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Asymptotic properties of the ML-estimator are not known. The model is strongly nonlinear, and asymptotic normality has not been proven even for univariate smooth transition GARCH models yet. For the latest results, see Meitz and Saikkonen (2008: a, b) who derive conditions for the stability and ergodicity of a class of models including the smooth transition GARCH model.

Nevertheless, the simulation results in this paper do not counteract our normality assumption, and as we consider our model useful we have to proceed without formal proofs.

In order to increase numerical efficiency of the estimation, maximization of the log-likelihood is carried out iteratively by concentrating the likelihood in each round by dividing the parameters into three sets: the GARCH, the correlation, and the transition function parameters. The log-likelihood is maximized with respect to one set at the time, keeping the other parameters fixed at their previously estimated values. The convergence is obtained once the estimated values cannot be improved upon when compared with the ones from the preceding iteration. As mentioned in Section 2.2, the transition between the extreme states becomes more rapid and the transition function eventually becomes a step function as $\gamma \to \infty$. When γ has reached a value large enough, no increment will change the shape of the transition function. The likelihood function becomes flat with respect to that parameter and numerical optimizers have difficulties in converging. Therefore, one may want to set an upper limit for γ, whose value naturally depends on the transition variable in question. Plotting a graph of the transition function can be useful in deciding such a limit. It should be noted that, if this limit is reached, the resulting estimates for the remaining parameters are conditional on this value of γ. Furthermore, it should be pointed out that estimation requires care. The log-likelihood may have several local maxima, so estimation should be initiated from a set of different starting-values, and the maxima thus obtained compared with each other before settling for final estimates. All computations in this paper have been performed using Ox, version 4.02, see Doornik (2002), and our own source code.

Before estimating an STCC–GARCH or a DSTCC–GARCH model, however, it is necessary to test the hypothesis that the conditional correlations are constant. The reason for this is that some of the parameters of the alternative model are not identified if the true model has constant conditional correlations. Estimating an STCC–GARCH or a DSTCC–GARCH model without first testing the constancy hypothesis could thus lead to inconsistent parameter estimates. The same is true if one wishes to increase the number of transitions in an already estimated STCC–GARCH model. Testing procedures will be discussed in the next section.

3 Hypothesis testing

3.1 Testing for smooth transitions in conditional correlations

Parametric modelling of the dynamic behaviour of conditional correlation must begin with testing constancy of the correlations. Neglected variation in parameters leads to a misspecified likelihood and thus to invalid asymptotic inference. Tse (2000), Bera and Kim (2002), Engle and Sheppard (2001), and Silvennoinen and Teräsvirta (2005) have proposed tests for this purpose. Tse (2000) derives a Lagrange multiplier (LM) test where the alternative model imposes ARCH-type dynamics on the conditional correlations. Bera and Kim (2002) discuss testing the hypothesis of no parameter variation using the information matrix test of White (1982). The test of Engle and Sheppard (2001) is based on the fact that the standardized residuals $\boldsymbol{\eta}_t$ should be iid both in time and across the series if the model is correctly specified. This test, however, is not only a test of constant correlations but a general misspecification test as it cannot distinguish between misspecified conditional correlations and conditional heteroskedasticity in the univariate residual series.

The approach of Silvennoinen and Teräsvirta (2005) differs from the others in that the test is conditioned on a particular transition variable and in effect tests whether that particular factor affects conditional correlations between the variables. A failure to reject the constancy

of correlations is thus interpreted as evidence that this transition variable is not informative about possible time-variation of the correlations. A non-rejection thus does not indicate that the correlations are constant, but the test may of course be carried out for a set of different transition variables. But then, a rejection of the null hypothesis does provide evidence of nonconstancy of the conditional correlations and may be taken to imply that the transition variable in question carries information about the time-varying structure of the correlations.

After fitting an STCC–GARCH model to the data one may wish to see whether or not the transition variable of the model is the only factor that affects the conditional correlations over time. In the present framework this means that there may be another factor whose effect on correlations cannot be ignored. For instance, the unconditional correlations may vary as a function of time, in which case a second transition depending directly on time together with the previous one would provide a better description of the correlation dynamics than the STCC–GARCH model does. A linear function of time would indicate a monotonic relationship between calendar time and correlations, whereas introducing higher-order polynomials or nonlinear functions would allow that structure to capture more complicated patterns in time-varying correlations.

An indication of the importance of a second transition variable can be obtained by testing the constancy of correlations against an STCC–GARCH model in which the correlations are functions of to this particular transition variable. The next step is to estimate the STCC–GARCH model with the transition variable against which the strongest rejection of constancy is obtained, and proceed by testing this model against the DSTCC–GARCH one.

The null hypothesis in testing for another transition is $\gamma_2 = 0$ in (6) and (7). The problem of unidentified parameters under the null hypothesis is circumvented by linearizing the form of dynamic correlations under the alternative model. This is done by a Taylor approximation of the second transition function, G_{2t}, around $\gamma_2 = 0$; see Luukkonen, Saikkonen, and Teräsvirta (1988) for this idea. Replacing the transition function in (6) by the approximation, the dynamic conditional correlations become

$$\boldsymbol{P}_t^* \quad = \quad (1 - G_{1t})\boldsymbol{P}_{(1)}^* + G_{1t}\boldsymbol{P}_{(2)}^* + s_{2t}\boldsymbol{P}_{(3)}^* + \boldsymbol{R} \tag{12}$$

where the remainder $\boldsymbol{R}$ is the error due to the linearization. Note that under the null hypothesis $\boldsymbol{R} = \mathbf{0}_{N\times N}$, so the remainder does not affect the asymptotic null distribution of the LM–test statistic. Note also that when $G_{1t} \equiv 0$ so that the correlations are constant under the null hypothesis, the test collapses into the correlation constancy test in Silvennoinen and Teräsvirta (2005). Details of the linearization and the transformed dynamic correlations in (12) can be found in the Appendix. The auxiliary null hypothesis can now be stated as $vecl\boldsymbol{P}_{(3)}^* = \mathbf{0}_{N(N-1)/2\times 1}$, where $vecl(\cdot)$ is an operator that stacks the columns of the strict lower triangular part of its argument square matrix. Under the null hypothesis,

$$\boldsymbol{P}_t^* \quad = \quad (1 - G_{1t})\boldsymbol{P}_{(1)}^* + G_{1t}\boldsymbol{P}_{(2)}^*. \tag{13}$$

Constructing the Lagrange multiplier test yields the statistic and its asymptotic null distribution in the usual way. The test statistic is

$$T^{-1}\left(\sum_{t=1}^{T}\frac{\partial l_t(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\rho}_{(3)}^{*\prime}}\right)\left[\hat{\mathfrak{I}}_T(\hat{\boldsymbol{\theta}})\right]^{-1}_{(\boldsymbol{\rho}_{(3)}^*,\boldsymbol{\rho}_{(3)}^*)}\left(\sum_{t=1}^{T}\frac{\partial l_t(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\rho}_{(3)}^*}\right) \overset{\mathrm{a}}{\sim} \chi^2_{N(N-1)/2}. \tag{14}$$

The detailed form of (14) can be found in the Appendix.

It should be pointed out that even if constancy is rejected against an STCC–GARCH model for both s_{1t} and s_{2t}, the test for another transition after estimating this model for one of the two

may not be able to reject the null hypothesis. This may be the case when both variables contain similar information about the correlations, whereby adding a second transition will not improve the model. Because estimation of an STCC–GARCH model can sometimes be a computationally demanding task, some idea of suitable transition variables may be obtained by testing constancy of correlations directly against the DSTCC–GARCH model. The null hypothesis is $\gamma_1 = \gamma_2 = 0$ in (6) and (7). To circumvent the problem with unidentified parameters under the null, both transition functions, G_{1t} and G_{2t}, in (6) are linearized around $\gamma_1 = 0$ and $\gamma_2 = 0$, respectively, as discussed above. The linearized dynamic correlations then become

$$\boldsymbol{P}_t^* \quad = \quad \boldsymbol{P}_{(1)}^* + s_{1t}\boldsymbol{P}_{(2)}^* + s_{2t}\boldsymbol{P}_{(3)}^* + s_{1t}s_{2t}\boldsymbol{P}_{(4)}^* + \boldsymbol{R} \tag{15}$$

where $\boldsymbol{R}$ again holds the approximation error. The auxiliary null hypothesis based on the transformed dynamic correlations is now $vecl\boldsymbol{P}_{(2)}^* = vecl\boldsymbol{P}_{(3)}^* = vecl\boldsymbol{P}_{(4)}^* = \mathbf{0}_{N(N-1)/2\times 1}$ under which the conditional correlations are constant, $\boldsymbol{P}_t^* = \boldsymbol{P}_{(1)}^*$. The Lagrange multiplier test statistic and its asymptotic null distribution are constructed as usual:

$$T^{-1}\left(\sum_{t=1}^{T}\frac{\partial l_t(\hat{\boldsymbol{\theta}})}{\partial(\boldsymbol{\rho}_{(2)}^{*\prime},\boldsymbol{\rho}_{(3)}^{*\prime},\boldsymbol{\rho}_{(4)}^{*\prime})}\right)\left[\hat{\mathfrak{I}}_T(\hat{\boldsymbol{\theta}})\right]^{-1}_{(\boldsymbol{\rho}_{(2-4)}^*,\boldsymbol{\rho}_{(2-4)}^*)}\left(\sum_{t=1}^{T}\frac{\partial l_t(\hat{\boldsymbol{\theta}})}{\partial(\boldsymbol{\rho}_{(2)}^{*\prime},\boldsymbol{\rho}_{(3)}^{*\prime},\boldsymbol{\rho}_{(4)}^{*\prime})^{\prime}}\right)\overset{\mathrm{a}}{\sim}\chi^2_{3N(N-1)/2}. \tag{16}$$

The detailed form of (16) can be found in the Appendix.

As discussed in Section 2.2, the full DSTCC–GARCH model is simplified when the effects of the two transition variables are independent. This restricted version can be used as a more parsimonious alternative than the DSTCC–GARCH model when testing constancy. For instance, if one of the transition variables, say s_{2t}, is time, one can test constancy against an alternative where the variation controlled by s_{1t} does not depend on time, that is, the differences between the two extremes states are equal during the whole sample period. The constancy of correlations is tested by testing $\gamma_1 = \gamma_2 = 0$ in (10) and now the linearized equation is simply a special case of (15) such that $\boldsymbol{P}_{(4)}^* = \mathbf{0}$. The auxiliary null hypothesis is $vecl\boldsymbol{P}_{(2)}^* = vecl\boldsymbol{P}_{(3)}^* = \mathbf{0}_{N(N-1)/2\times 1}$ under which the conditional correlations are constant: $\boldsymbol{P}_t^* = \boldsymbol{P}_{(1)}^*$. The Lagrange multiplier test statistic and its asymptotic null distribution are the following:

$$T^{-1}\left(\sum_{t=1}^{T}\frac{\partial l_t(\hat{\boldsymbol{\theta}})}{\partial(\boldsymbol{\rho}_{(2)}^{*\prime},\boldsymbol{\rho}_{(3)}^{*\prime})}\right)\left[\hat{\mathfrak{I}}_T(\hat{\boldsymbol{\theta}})\right]^{-1}_{(\boldsymbol{\rho}_{(2-3)}^*,\boldsymbol{\rho}_{(2-3)}^*)}\left(\sum_{t=1}^{T}\frac{\partial l_t(\hat{\boldsymbol{\theta}})}{\partial(\boldsymbol{\rho}_{(2)}^{*\prime},\boldsymbol{\rho}_{(3)}^{*\prime})^{\prime}}\right)\overset{\mathrm{a}}{\sim}\chi^2_{N(N-1)}. \tag{17}$$

Statistic (17) is a special case of (16) and its detailed form can be found in the Appendix.

If the tests fail to reject the null hypothesis or if the rejection is not particularly strong, the reason for this can be that some correlations are constant. If sufficiently many but not all correlations are constant according to one of the transition variables or both, the tests may not be sufficiently powerful to reject the null hypothesis. In these cases the power of the tests can be increased by modifying the alternative model. For instance, one may want to test constancy of correlations against a DSTCC–GARCH model in which some correlations are constant with respect to one of the transition variables, or both. Similarly, an STCC–GARCH model containing constant correlations can be tested against a DSTCC–GARCH model with constancy restrictions. These tests are straightforward extensions of the tests already discussed, see the Appendix for details. A test of constant correlations against an STCC–GARCH model containing constant correlations is discussed in Silvennoinen and Teräsvirta (2005).

When it comes to 'fine-tuning' of the model, i.e., when the model under the null hypothesis has the same number of transitions as under the alternative and the modeller is focused on potential constancy of some of the correlations controlled by one of the transition variables or

both, tests of partial constancy can also be built on the Wald principle. This is quite practical because after estimating the alternative model, several restrictions can be tested at the same time without re-estimation. In our experience, when it comes to restricting correlations to be constant, the specification search beginning with a general model and restricting correlations generally yields the same final model as would a bottoms-up approach beginning with a restricted model and testing for additional time-varying correlations. This conclusion has been reached using both simulated series and observed returns. The only difference between these two approaches seems to be that the final model is obtained faster with the former than with the latter. Restricting some of the correlations to be constant decreases the number of parameters to be estimated, which is convenient especially in large models.

4 Size simulations

Empirical size of the LM–type test of STCC–GARCH models against DSTCC–GARCH ones are investigated by simulation. The observations are generated from a bivariate first-order STCC–GARCH model. The transition variable is generated from an exogenous process: $s_t = h_{et}^{1/2} z_{et}$, where h_{et} has a GARCH(1,1) structure and $z_{et} \sim nid(0,1)$. The STCC–GARCH model is tested against a DSTCC–GARCH model where the correlations also vary as a function of time, in which case the alternative model is the TVSTCC–GARCH model. The parameter values in each of the individual GARCH equations are chosen such that they resemble results often found in fitting GARCH(1,1) models to financial return series. Thus,

$$\begin{aligned} h_{1t} &= 0.01 + 0.04\varepsilon_{1,t-1}^2 + 0.94 h_{1,t-1} \\ h_{2t} &= 0.03 + 0.05\varepsilon_{2,t-1}^2 + 0.92 h_{2,t-1} \\ h_{et} &= 0.005 + 0.03 s_{t-1}^2 + 0.96 h_{et,t-1}. \end{aligned}$$

We conduct three experiments where $\rho_{(1)} = 0$, and $\rho_{(2)} = 1/3$, $1/2$, $2/3$. The location parameter $c = 0$. We consider two choices for the value of the slope parameter γ. The first one, $\gamma = 5$, represents a rather slow transition: about 75% of the correlations lie genuinely between $\rho_{(1)}$ and $\rho_{(2)}$, and the remaining 25% take one of the extreme values. The other choice is $\gamma = 20$, so the ratios are now interchanged: only 25% of the correlations are different from $\rho_{(1)}$ or $\rho_{(2)}$. The sample sizes are $T = 1000$ and $T = 2500$. Considering longer time series was found unnecessary because the results suggested that the empirical size is close to the nominal one at these sample sizes already. The results in Table 1 are based on 5000 replications.

We carry out another size simulation experiment in which we test the CCC–GARCH model directly against the TVSTCC–GARCH model. In the latter model, one transition is controlled by an exogenous GARCH(1,1) process and the other one by time. Specifically, the model under the null is a bivariate CCC–GARCH(1,1) model where the GARCH processes are h_{1t} and h_{2t} from the previous study. For the constant correlation between the series we use four values: 0, 1/3, 1/2, and 2/3. These four experiments are performed using samples of sizes $T = 1000$ and $T = 2500$, with 5000 replications. The results in Table 2 indicate that the test does not suffer from size distortions and can thus be applied without further adjustments. This was also the case for the test of constant correlations against the STCC–GARCH model, reported in Silvennoinen and Teräsvirta (2005).

5 Estimated correlations from different MGARCH models

In this section we compare the estimated correlations from a selection of multivariate GARCH models by fitting them to the same data set. In order to keep the comparison transparent, we only consider bivariate models. Our observations are the daily returns of S&P 500 index futures and 10-year bond futures from January 1990 to August 2003. This data set has been analyzed by Engle and Colacito (2006).[2] There is no consensus in the literature about how stock and long term bond returns are related. Historically, the long-run correlations have been assumed constant, an assumption that has led to contradicting conclusions because evidence for both positive and negative correlation has been found over the years (short-run correlations have been found to be affected, among other things, by news announcements). From a theoretical point of view, the long-run correlation between the two should be state-dependent, driven by macroeconomic factors such as growth, inflation, and interest rates. The way the correlations respond to these factors may, however, change over time.

For this reason it is interesting to see what the correlations between the two asset returns obtained from different models are and how they fluctuate over time. For the comparison we use three MGARCH models that model the univariate volatilities as first-order GARCH models and differ mainly by their way of defining the correlations. One of them is naturally the DSTCC–GARCH one. Another is the DCC–GARCH model of Engle (2002) which has gained popularity partly due to the small number of parameters related to the time-varying conditional correlations. In this model, the correlation dynamics are inherited from the past covariances among the returns. The third model is the semiparametric SPCC–GARCH model of Hafner, van Dijk, and Franses (2005). In their model the conditional correlations are estimated nonparametrically through a kernel smoother. The volatility of the returns of the S&P 500 index futures responds asymmetrically to past shocks and is therefore modelled as GJR–GARCH(1,1). The bond returns exhibit no such asymmetry and hence a GARCH(1,1) model is used for modelling its volatility. The focus of reporting results will be on conditional correlations implied by the estimated models. For conciseness, we do not present the parameter estimates for the models.

Relying on the tests in Section 3 and in Silvennoinen and Teräsvirta (2005) we selected relevant transition variables for the DSTCC–GARCH model. Candidates for transition variables were constructed from the past observations, prices or returns. Exogenous alternatives included interest rates, in levels or as changes over periods of different lengths.

Out of the multitude of variables, however, The Chicago Board Options Exchange volatility index (VIX) that represents the market expectations of 30-day volatility turned out to have the best performance. The VIX is constructed using the implied volatilities of a wide range of S&P 500 index options. It is a commonly used measure of market risk and is for this reason often referred to as the 'investor fear gauge'. The values of the index exceeding 30 are generally associated with a large amount of volatility, due to investors' fear or uncertainty, whereas the values below 20 indicate less stressful, even complacent, times in the markets. A graph of the VIX appears in Figure 2. Calendar time seemed to be another well-performing transition variable. Table 3 presents the results from testing constant correlations hypothesis against the STCC–GARCH model where the transition variable is one-day lag of VIX, the TVCC–GARCH model, and the DSTCC–GARCH model with both lagged VIX and time as transition variables. These tests result in clear rejection of the correlations being constant. The tests for an additional transition, both when using STCC–GARCH model with VIX or TVCC–GARCH model as null

[2] The data set in Engle and Colacito (2006) begins in August 1988, but our sample starts from January 1990 because we also use the time series for a volatility index that is available only from that date onwards.

models, also reject the constant correlations hypothesis and thus shows support for the DSTCC–GARCH model. As a result, the first-order TVSTCC–GARCH model was fitted to the bivariate data.

The semiparametric SPCC–GARCH model also requires a choice of an indicator variable. Because the previous test results indicated that the volatility index is informative about the dynamics of the correlations, we chose the one-day lag of VIX as the transition variable in this model as well. The SPCC–GARCH model was estimated using a standard kernel smoother with an optimal fixed bandwidth, see Pagan and Ullah (1999, Sections 2.4.2 and 2.7) for discussion on the choice of constant bandwidth.

The correlations implied by each of the models are presented in Figure 1. A rough inspection of them shows that the general shape of the graphs appears similar for every model. The conditional correlations fluctuate around 0.4 at the beginning of the sample. Towards the end of the 1990's, however, the fluctuations increase while the correlations tend to decrease, reaching negative values as well. Despite these similarities, there are some drastic differences between the models. One of them can be observed in the second half of 1990 and early 1991, a period that, according to the VIX, reflects market distress. In the SPCC–GARCH model the correlations decrease, whereas the ones implied by the TVSTCC– and DCC–GARCH models rise sharply. During 1997, right from the beginning of the year, the correlations from the SPCC–GARCH model start to fluctuate towards a lower level. For the other two models they first remain at a fairly constant level before a rather dramatic decrease due to the Asian crisis much later that year. The correlations from the DCC–GARCH model display a general downward tendency starting in 1999, whereas a similar pattern cannot be found in the ones from the TVSTCC– and SPCC–GARCH models. The SPCC–GARCH model produces correlations that are fluctuating quite rapidly, and the range of variation is wider than it is for the correlations from the other two models. It is worth noting that during the second half of the observation period the models agree in the way the correlations change when the markets grow turbulent: periods of high uncertainty are matched with decreasing correlations.

The log-likelihood values for the estimated TVSTCC–, SPCC–, and DCC–GARCH models are -5944, -6007, and -6118, respectively. Because the models are bivariate, none of them are penalized for an excessive number of parameters. The order of preference thus remains the same, independent of whether Akaike's or the Bayesian information criterion or just the pure likelihood value is used for ranking the models (note that the semiparametric model is in principle favoured in rankings based on AIC or BIC due to the nonparametric correlation estimates).

Conditional correlations are not observable and their behaviour is often conjectured and reasoned by economic assumptions. For instance, it is often stated that the correlations among asset returns increase during times of distress. In this study, however, we are able to establish the link between correlations and market uncertainty. The test results indicate that the volatility index carries information about the correlation movements, and the estimated TVSTCC– and SPCC–GARCH models reveal the direction of the movement: high uncertainty pushes correlations towards lower values than they would be during calm periods.

The correlation estimates from the TVSTCC–GARCH model reveal an interesting detail in the general investment behaviour. In the turbulent (according to the VIX) period covering the second half of 1990 and early 1991, the estimated correlations are *higher* than during the calm period that ends with the Asian crisis in 1997. This behaviour is reversed after the Asian crisis: increased uncertainty reflected as an increase in the VIX results in *lower* and strongly negative correlations.

6 Correlations between world market indices

Correlations are especially relevant to risk management and finding efficient hedging positions for portfolios. Inaccurate estimates of correlations put the performance of hedging operations at risk. In the preceding section, the focus was on the correlations between the returns of two different types of assets, stock and bond futures. We now turn to the situation in which an investor is investing in stocks, and his hedging strategy is to diversify the portfolio internationally. A potential problem is that, due to globalization, the financial markets around the world may have become increasingly integrated, which can weaken the protection of investors' portfolios against local or national crises.

Consequently, the focus will be on the correlation dynamics among world stock indices. The interest lies in revealing potential risks to internationally diversified portfolios, posed by increasing integration of the world markets. The five major indices considered are the French CAC 40, the German DAX, FTSE 100 from the UK, the Hong Kong Hang Seng index (HSI), and the Japanese Nikkei 225 (NKY). We use weekly observations recorded as the closing price of the current week from the beginning of December 1990 to the end of April 2006, 804 observations in all. Weekly observations are preferred to daily ones because the aggregation over time is likely to weaken the effect of different opening hours of the markets around the world. Martens and Poon (2001) discussed the problem of distinguishing contemporaneous correlation from a spillover effect and provided evidence of downward bias in estimated correlations in the presence of nonsynchronous markets.

The returns are calculated as differenced log prices. Descriptive statistics of the return series are reported in Table 4.

It is often found, see for instance Lin, Engle, and Ito (1994), de Santis and Gerard (1997), Longin and Solnik (2001), Chesnay and Jondeau (2001), and Cappiello, Engle, and Sheppard (2006), that the correlations between stock indices behave differently in times of distress from what they do during periods of tranquillity. It is therefore of interest to study how the general level of uncertainty or market turbulence affects the correlation dynamics between the stock indices. In order to do this, we choose our first transition variable to be the one-week lag of the VIX already employed in the previous section. But then, even here the general level of conditional correlations can change over time. In order to allow for this effect our second transition variable in the DSTCC–GARCH model will again be time.

Tests of constant correlations against smooth transition over time as well as against correlations that vary according to the lagged VIX result in rejecting the null hypothesis in the full five-variate model. When testing constancy against the TVSTCC–GARCH model, the rejection of the null model is very strong (the p–value equals 4×10^{-31}). It appears that both time itself and the volatility index convey information about the process causing the conditional correlations to fluctuate over time. As Silvennoinen and Teräsvirta (2005) showed, valuable information of the behaviour of the correlations can be extracted from studying submodels. For this reason we study bivariate models of stock returns before considering the full five-variate model.

It is also of interest to assess the potential importance of including an asymmetry component in the variance equations. To this end, we also estimated the full five-variate model with constant correlations imposing the restriction $\delta_{ij} = 0$ in (2). The test against the TVSTCC–GARCH model still rejects the null of constant correlations as strongly as before (p-value equals 4×10^{-31}). Although there appears to be little difference in the resulting correlations from the two models with regard to the dynamics due to the VIX and time, we shall compare the results from using only symmetric GARCH equations to the ones obtained from GJR–GARCH equations.

As already mentioned, studying submodels may provide information that can get lost in

higher dimensional models. Table 5 contains p–values of the bivariate tests of constant correlations against the TVCC–GARCH model, the STCC–GARCH model where the transition variable is the lagged VIX, and the TVSTCC–GARCH model. The univariate GARCH models restrict the asymmetry parameter δ to be equal to zero. Time clearly appears to be an indicator of change in correlations: the null hypothesis is rejected at the 1% significance level for every bivariate combination of the indices except for the combination of FTSE and Nikkei ($p = 0.0184$). The volatility index seems to be a substantially weaker indicator of change than time. When VIX is the transition variable, the test rejects at the 1% significance level only in four out of the ten cases, although yet another p–value does remain close to 0.01. When constancy is tested directly against TVSTCC–GARCH model the rejections are very strong, as is seen from the fourth column of Table 5. The estimated time-varying bivariate conditional correlations from TVSTCC–GARCH models for the five models that reject the test for additional transition and from TVCC–GARCH models for the remaining five models, are plotted in Figure 3. Note that one of them is the FTSE – Nikkei model. Although constancy of correlations was not rejected at the 1% level against VIX, the TVCC-GARCH model for this pair of returns is rejected at this level ($p = 0.0021$), so the TVSTCC-GARCH model is fitted to these series as well.

The results of the corresponding tests, and estimation results from the TVCC–GARCH models that have GJR–GARCH model in their volatilities, can be found in Table 6. The outcomes are practically identical to the previous ones. Because all tests except one reject constancy of correlations in favour of variation in time at the 1% level and the remaining p-value is close ($p = 0.0139$), we estimate the TVCC–GARCH model for each pair of series and then test for another transition. The correlation estimates for each of these bivariate models along with estimates for the smooth transition parameters c and γ are reported in the last four columns of Tables 5 and 6. In some of the models the parameter estimate of γ reaches its upper bound (500) and thus defines the transition as being nearly a break. The p–values resulting from testing for another transition appear in the fifth column of Tables 5 and 6. Now some differences between the GARCH and GJR-GARCH -based models begin to emerge. For the latter models, the TVCC–GARCH model is not rejected against the TVSTCC–GARCH alternative. For the former, one half of the tests favour the VIX as the second transition variable. This indicates that modelling the asymmetric behaviour in the volatilities accounts for some of the information that would otherwise have been included in the model through the VIX as a transition variable in the matrix of time-varying correlations.

From now on, the focus will be on the GJR-GARCH -based models. The results in Table 6 indicate that VIX may not be an appropriate transition variable for the full five-variate model, because it does not affect all pairs of correlations. Nevertheless, VIX should not be abandoned altogether for the reason that it may still be a useful indicator somewhere in the model. This means that some of the correlations may be constant with respect to VIX whereas others are not, while all of them vary as a function of time. This possibility can be investigated by partial constancy tests. We shall come back to this later on after estimating the five-variate TVCC–GARCH model. Because this model restricts both the location and the shape of the smooth transition over time to be the same for all indices in the full five-variate model, we have to find out whether or not imposing that kind of a restriction is plausible. This is done by comparing the estimated bivariate TVCC–GARCH models in Table 6. Their time-varying bivariate conditional correlations are plotted in Figure 4. They are positive and increase during the observation period. The transitions between the three European indices, CAC, DAX, and FTSE, on one hand and the Asian market indices, HSI and NKY, on the other, are very rapid. With HSI the transition occurs around the introduction of the euro, whereas for NKY the change takes place almost five years later. On the contrary, the correlation between HSI and NKY returns increases quite

slowly.

The introduction of euro does not seem to have any special impact on the correlation between CAC and DAX returns. It increases slowly but steadily and exceeds 0.9 at the end of the period. The correlation between FTSE and these two Euroland indices stays roughly around 0.6 before climbing up to about 0.85 around 2000 or 2001.

In order to check how well the bivariate models fit together, we construct the five-dimensional correlation matrix out of the correlations estimated from the bivariate models. It turns out to be positive definite for all t. There is no guarantee, however, that a high-dimensional correlation matrix constructed from bivariate correlations from DSTCC-GARCH models should be positive definite. That it happens to be the case here may be due to the fact that the correlations change roughly at the same time. Although they do not change equally rapidly, the mid-point of the change seems to lie around the turn of the century.

The complete five-variate TVSTCC–GARCH model for these five return series models does have a positive definite correlation matrix by definition if the four correlation matrices in (6) are positive definite. We shall now consider this model.

Within this framework the rejection of constancy of correlations when using time as the transition variable is much stronger ($p = 1 \times 10^{-34}$) than it is when the lagged VIX is the transition variable ($p = 0.0003$). Consequently, we first fit a TVCC–GARCH model to our data and then test for another transition. This test rejects the null model ($p = 0.0002$), and we proceed by estimating the full five-variate TVSTCC–GARCH model. As the bivariate tests already suggest, some of the correlation estimates do not differ significantly from each other at the beginning of the sample between the states defined by $\boldsymbol{P}_{(11)}$ and $\boldsymbol{P}_{(21)}$, or at the end of the sample between the ones indicated by $\boldsymbol{P}_{(12)}$ and $\boldsymbol{P}_{(22)}$. We test the TVCC–GARCH model against TVSTCC–GARCH models in which some correlations are constant in the VIX dimension. In order to save space, the results of these tests are not reported here.

For the years before the 1990's, the return correlations have been reported to increase during periods of distress from the levels that they maintain during calm periods, a phenomenon that had become global with the increase of financial market integration, see for instance Lin, Engle, and Ito (1994) and de Santis and Gerard (1997). Our observation period starts in December 1990, and our estimation results suggest that the behaviour of the correlations, when it comes to turbulent and calm periods, may have changed from what it was before the 1990's. The estimated transition due to VIX, when it occurs, is rather abrupt, and one may thus speak about high and low volatility regimes. The former regime prevails when VIX exceeds $\hat{c}_1 = 22.9$, which is the case in about 22% of the observations. As we shall shortly discuss, there is evidence that some of the correlations that do respond to VIX actually behave in a way diametrically opposite to the one observed before 1990.

The parameter estimates of the final TVSTCC–GARCH model can be found in Table 7. For illustration, the estimated correlations are plotted in Figure 5. As may be expected from previous results, the correlations seem to have increased at the turn of the century. The estimated midpoint of the transition, 0.59, points at the spring of 1999. This is in agreement with the results from the bivariate models and also accords with the findings of Cappiello, Engle, and Sheppard (2006). These authors found a structural break around January 1999 that meant an increase in the correlations from their previous unconditional levels. Their estimated break date coincides with the introduction of euro.

Figure 5 shows how the correlations between Nikkei and the European index returns are linked to the overall turbulence of the markets. As VIX climbs to a high level towards the end of the 1990's and fluctuates there early this century, the correlations tend to reach the low level that prevailed before the uncertainty of the late 1990's. But then, the range of the correlations

as a function of VIX increases over time, because a low value of VIX means an increasingly high correlation. After 2004 when the markets calm down, the correlations between Nikkei on one hand and CAC, DAX, and FTSE on the other, stabilize to a level between 0.6 and 0.7.

It is also interesting to note the dynamics of the correlations between Hang Seng and DAX as well as Hang Seng and FTSE. As VIX increases, then, contrary to the previous case, the correlations do the same. Between HSI and DAX this pattern becomes evident during the latter half of the sample. Hang Seng and FTSE apparently have this behaviour over the whole sample, as can be seen from the estimated correlations during the turbulent times of the early 1990's and the ones around the turn of the century. A similar pattern can be seen in correlations between the CAC and DAX returns during the first half of the sample. However, increased uncertainty has a completely opposite effect on the correlations after the introduction of the euro: the correlations tend to be lower during the turbulent times than they become when the markets calm down again.

Finally, we have a look at the correlations between the returns of FTSE and those of both CAC and DAX. Although the extreme levels of correlations as a function of VIX are barely significantly different from each other, an interesting pattern can be discerned. Early in the sample the correlations between the FTSE and CAC returns decrease with increasing market uncertainty, whereas the ones between the FTSE and DAX increase. However, by 1999 these differences are no longer there, and the correlations settle on a high but constant level and no longer respond to market volatility.

7 Conclusions

In this paper we extend the Smooth Transition Conditional Correlation (STCC) GARCH model of Silvennoinen and Teräsvirta (2005). The new model, the Double Smooth Transition Conditional Correlation (DSTCC) GARCH model allows time-variation in the conditional correlations to be controlled by two transition variables instead of only one. A useful choice for one of the transition variables is simply time, in which case the model also accounts for a change in unconditional correlations over time. This is a very appealing property because in applications of GARCH models the number of observations is often quite large. The time series may be, for example, daily returns and consist of several years of data. It is not reasonable to simply assume that the correlations remain constant over years and in fact, as shown in the empirical application, and also in that of Berben and Jansen (2005a, b), this is generally not the case.

We also complement the battery of specification and misspecification tests in Silvennoinen and Teräsvirta (2005). We derive LM–tests for testing constancy of correlations against the DSTCC–GARCH model and testing whether another transition is required, i.e. testing STCC–GARCH model against DSTCC–GARCH model. We also discuss the implementation of partial constancy restrictions into the tests above. This becomes especially relevant when the number of variables in the model is large, because the tests offer an opportunity to reduce the otherwise large number of parameters to be estimated.

We estimate three different MGARCH models to a data set consisting of the S&P 500 index and long term bond futures. While the overall tendency in the conditional correlations is relatively similar in all of the models, there are some differences that could yield different conclusions, depending on the application at hand. Therefore, it might be advisable not to rely on estimation results from a single model only. Comparing ones from several models can reveal properties of the correlation dynamics and help choosing a model. Some models may yield estimated correlations that are fairly smooth while others can fluctuate largely. Depending on

what the model is used for, one model may be seen more suitable than the others; for instance, smoothly behaving correlations do not require as frequent updating in an application as would the erratic ones.

We apply the DSTCC–GARCH model to a set of world stock indices from Europe and Asia. As discussed in Longin and Solnik (2001), the market trend affects correlations more than volatility. We use the CBOE volatility index as one transition variable to account for both uncertainty and volatility on the markets. The other transition variable has been time which allows the level of unconditional correlations to change over time. We find a clear upward shift in the level of unconditional correlations around the turn of the century. This change is significant both within and across the two geographical areas, Europe and Asia. The volatility index seems to carry some information about the time-varying correlations. This is more evident in Europe at the beginning of the observation period, and between Europe and Asia towards the end of the sample. In the latter, we found that the correlations behave in an opposite way to the documented tendency of correlations to increase during increased market volatility. Especially the Japanese markets seem to drop the level of correlations with the European markets whenever the markets grow uncertain. The extension of the original STCC–GARCH model clearly proves useful because of its ability to describe the effects of two different transition variables, in our applications market uncertainty and time, on conditional correlations.

Figures

Figure 1: Conditional correlations implied by the estimated models: DCC–GARCH, TVSTCC–GARCH, and SPCC–GARCH. The daily observation period is from January 1990 to August 2003.

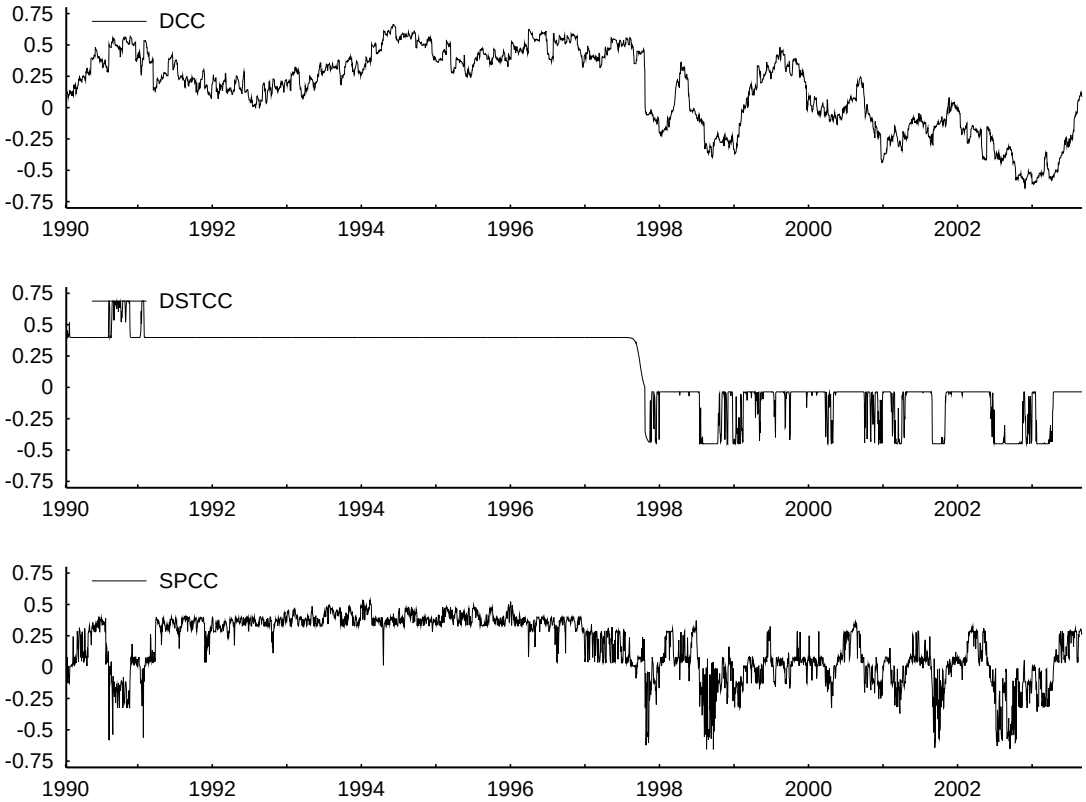


Figure 2: The daily volatility index VIX. The daily observation period is from January 1990 to August 2003.

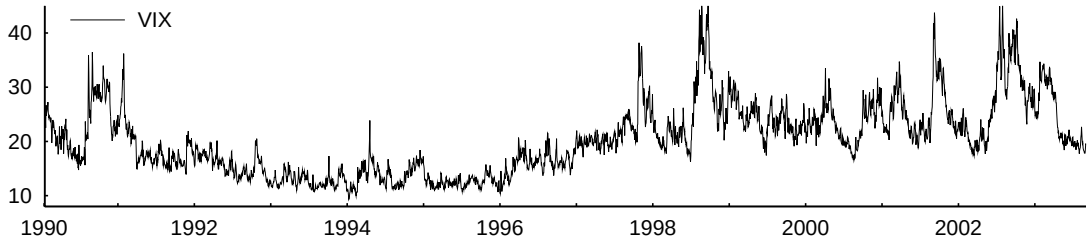


Figure 3: Estimated time-varying correlations in the bivariate TVSTCC–GARCH model where univariate volatilities are modelled as GARCH. The weekly observation period is from the beginning of December 1990 to end of April 2006.

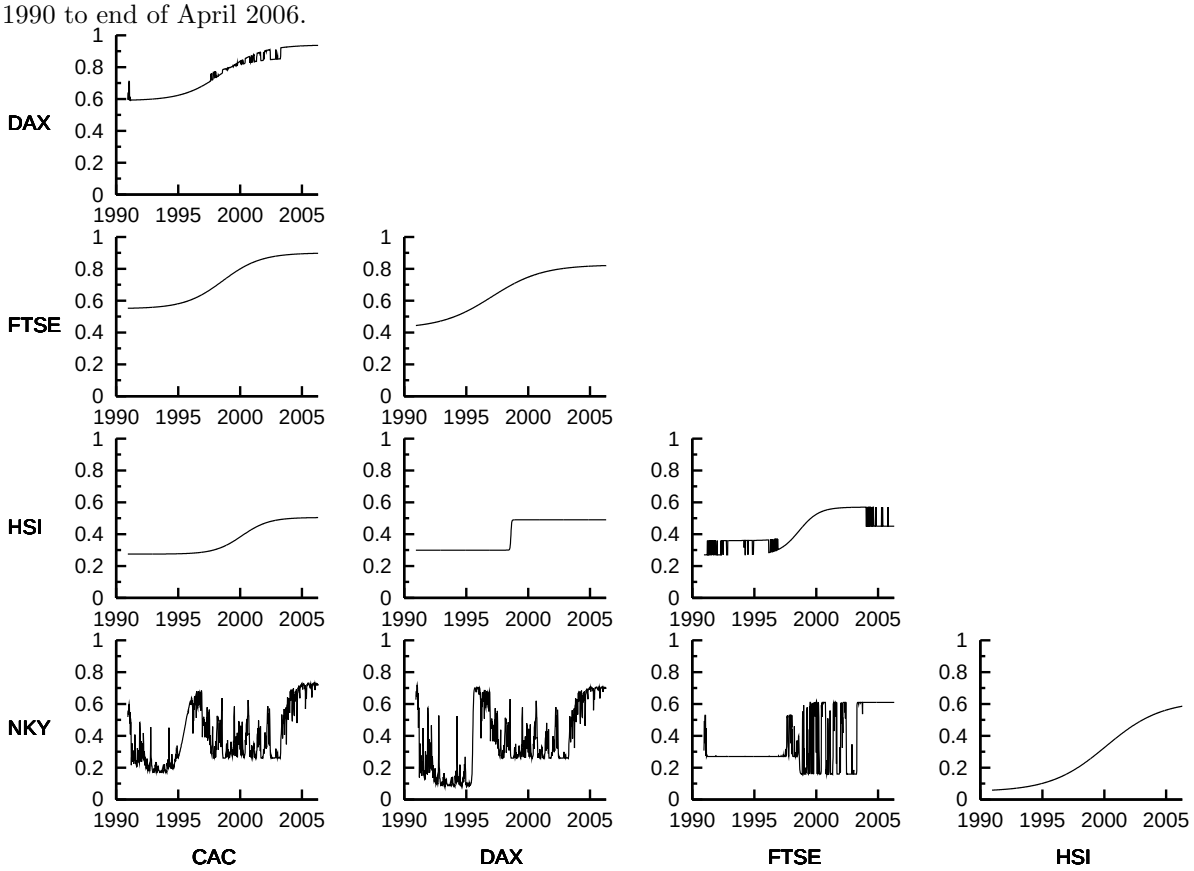


Figure 4: Estimated time-varying correlations in the bivariate TVCC–GARCH models where univariate volatilities are modelled as GJR–GARCH. The weekly observation period is from the beginning of December 1990 to end of April 2006.

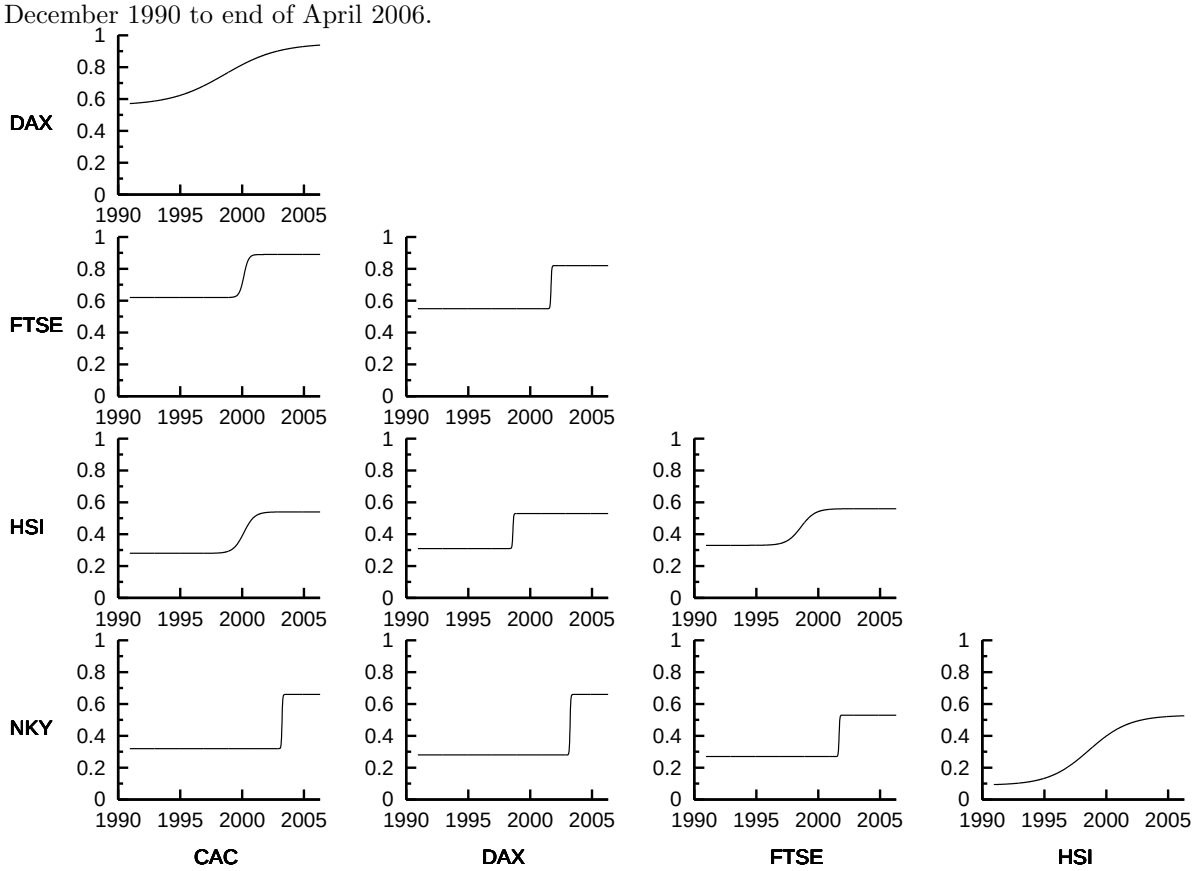
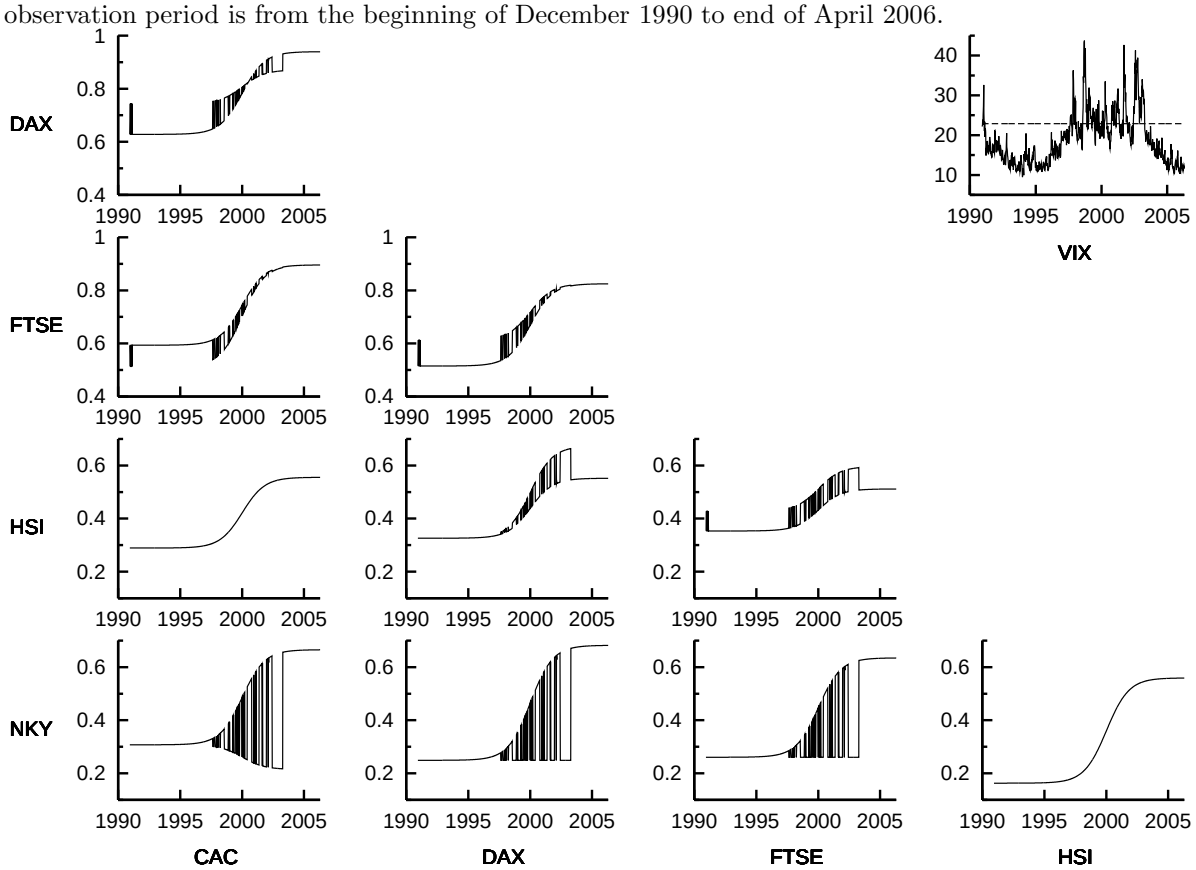


Figure 5: Estimated time-varying correlations in the five-variate TVSTCC–GARCH model. The weekly observation period is from the beginning of December 1990 to end of April 2006.



Tables

		$\rho_1 = 0,\ \rho_2 = 1/3$		$\rho_1 = 0,\ \rho_2 = 1/2$		$\rho_1 = 0,\ \rho_2 = 2/3$	
	nominal size	$\gamma = 5$	$\gamma = 20$	$\gamma = 5$	$\gamma = 20$	$\gamma = 5$	$\gamma = 20$
	1%	0.0104	0.0082	0.0082	0.0098	0.0072	0.0100
$T = 1000$	5%	0.0482	0.0418	0.0472	0.0500	0.0432	0.0502
	10%	0.0944	0.0884	0.0970	0.0996	0.0912	0.1016
	1%	0.0098	0.0106	0.0106	0.0120	0.0074	0.0108
$T = 2500$	5%	0.0518	0.0512	0.0498	0.0472	0.0456	0.0462
	10%	0.0958	0.1020	0.1018	0.1010	0.0964	0.1012

Table 1: Empirical size of the test of the STCC–GARCH model against an STCC–GARCH model with an additional transition for sample sizes 1000 and 2500 and for three choices of correlations for the extreme states; 5000 replications.

	nominal size	$\rho = 0$	$\rho = 1/3$	$\rho = 1/2$	$\rho = 2/3$
	1%	0.0116	0.0122	0.0138	0.0144
$T = 1000$	5%	0.0522	0.0538	0.0542	0.0616
	10%	0.1016	0.1054	0.1050	0.1156
	1%	0.0102	0.0136	0.0146	0.0146
$T = 2500$	5%	0.0466	0.0550	0.0588	0.0538
	10%	0.0950	0.1068	0.1132	0.1072

Table 2: Empirical size of the test of the CCC–GARCH model against an STCC–GARCH model with two transitions for sample sizes 1000 and 2500 and for three choices of correlations for the extreme states; 5000 replications.

Estimated model	CCC		STCC s_t =VIX		TVCC $s_t = t/T$		TVSTCC s_{1t} =VIX, $s_{2t} = t/T$	
Correlations	ρ	0.1460 (0.0194)	$\rho_{(1)}$	0.5372 (0.0864)	$\rho_{(1)}$	0.4134 (0.0194)	$\rho_{(11)}$	0.3979 (0.0202)
			$\rho_{(2)}$	−0.3242 (0.0790)	$\rho_{(2)}$	−0.1613 (0.0249)	$\rho_{(12)}$	−0.0356 (0.0348)
							$\rho_{(21)}$	0.6894 (0.0584)
							$\rho_{(22)}$	−0.4497 (0.0395)
Transition parameters			c	21.93 (0.97)	c	0.57 (0.00)	c_1	27.44 (0.49)
			γ	0.23 (0.07)	γ	500 (—)	γ_1	1.94 (1.09)
							c_2	0.57 (0.00)
							γ_2	500 (—)
		p-value		p-value		p-value		
Transition variable(s) in the test	VIX	3×10^{-49}	t/T	8×10^{-15}	VIX	0.0005		
	t/T	1×10^{-47}						
	VIX, t/T	6×10^{-74}						
Loglikelihood	-6124		-6020		-5978		-5944	

Table 3: Estimated correlations (standard errors in parentheses) and the results (p-values) from tests of constant correlations against an STCC–GARCH model with single or double transition, and tests for additional transition.

	min	max	mean	st.dev	skewness	kurtosis
CAC	-12.27	10.89	0.0000	2.7337	-0.1272	4.1330
DAX	-14.26	12.71	0.0000	2.9717	-0.2849	5.1952
FTSE	-8.99	9.94	0.0000	2.0680	-0.0944	4.8674
HSI	-20.14	13.70	0.0000	3.4723	-0.4302	5.9599
NKY	-11.26	11.08	0.0000	2.8602	0.0436	3.9162

Table 4: Descriptive statistics of the return series.

Estimated model	CCC	CCC	CCC	TVCC	$\rho_{(1)}$	$\rho_{(2)}$	c	γ
Transition variable in the test	t/T	VIX_{t-1}	VIX_{t-1} and t/T	VIX_{t-1}				
CAC–DAX	1×10^{-24}	0.0005	5×10^{-23}	0.0049	0.5475 (0.0962)	0.9505 (0.0340)	0.48 (0.09)	6.42 (3.30)
CAC–FTSE	1×10^{-16}	0.0084	2×10^{-15}	0.3435	0.6267 (0.0258)	0.8818 (0.0117)	0.63 (0.01)	74.66 (2.10)
CAC–HSI	0.0008	0.0132	0.0016	0.1583	0.2961 (0.0409)	0.5335 (0.0375)	0.57 (0.00)	500 (—)
CAC–NKY	0.0042	0.0608	0.0002	0.0018	0.2919 (0.0379)	0.5501 (0.0411)	0.68 (0.00)	500 (—)
DAX–FTSE	3×10^{-11}	0.0078	7×10^{-10}	0.9708	0.5103 (0.0509)	0.8179 (0.0247)	0.53 (0.07)	9.77 (2.75)
DAX–HSI	0.0031	0.0899	0.0049	0.3126	0.3120 (0.0430)	0.5334 (0.0354)	0.51 (0.00)	500 (—)
DAX–NKY	0.0004	0.1186	3×10^{-5}	0.0031	0.2566 (0.0392)	0.5557 (0.0401)	0.68 (0.01)	500 (—)
FTSE–HSI	0.0027	0.0021	0.0018	0.0047	0.2401 (0.1275)	0.5303 (0.0443)	0.31 (0.18)	8.84 (2.22)
FTSE–NKY	0.0184	0.0580	0.0032	0.0021	0.2701 (0.0385)	0.5207 (0.0435)	0.68 (0.01)	500 (—)
HSI–NKY	9×10^{-7}	0.0752	1×10^{-5}	0.5452	0.0851 (0.0904)	0.5433 (0.0579)	0.50 (0.09)	9.33 (4.23)

Table 5: Results (p–values) from bivariate tests of constant correlations against an STCC–GARCH model with single or double transition, and bivariate tests of another transition in TVCC–GARCH model. The last four columns report the estimation results for each of the bivariate TVCC–GARCH model. The univariate volatilities are modelled as GARCH for both stocks and bonds. The standard errors are given in parentheses.

Estimated model	CCC	CCC	CCC	TVCC	$\rho_{(1)}$	$\rho_{(2)}$	c	γ
Transition variable in the test	t/T	VIX_{t-1}	VIX_{t-1} and t/T	VIX_{t-1}				
CAC–DAX	1×10^{-24}	0.0002	2×10^{-23}	0.1313	0.5616 (0.0762)	0.9492 (0.0287)	0.51 (0.07)	7.02 (3.22)
CAC–FTSE	7×10^{-17}	0.0077	5×10^{-16}	0.6910	0.6160 (0.0259)	0.8905 (0.0110)	0.63 (0.01)	82.80 (2.10)
CAC–HSI	0.0004	0.0104	0.0010	0.1606	0.2825 (0.0536)	0.5428 (0.0423)	0.55 (0.05)	36.07 (35.37)
CAC–NKY	0.0061	0.0996	0.0009	0.7874	0.3167 (0.0340)	0.6641 (0.0462)	0.82 (0.01)	500 (—)
DAX–FTSE	2×10^{-11}	0.0060	3×10^{-10}	0.3801	0.5546 (0.0289)	0.8205 (0.0171)	0.65 (0.00)	500 (—)
DAX–HSI	0.0013	0.0651	0.0020	0.2538	0.3051 (0.0432)	0.5335 (0.0352)	0.51 (0.00)	500 (—)
DAX–NKY	0.0004	0.1878	0.0001	0.6658	0.2843 (0.0352)	0.6604 (0.0426)	0.81 (0.01)	500 (—)
FTSE–HSI	0.0005	0.0010	0.0003	0.0336	0.3262 (0.0483)	0.5619 (0.0387)	0.50 (0.05)	26.99 (2.09)
FTSE–NKY	0.0139	0.1388	0.0092	0.0294	0.2659 (0.0381)	0.5321 (0.0439)	0.68 (0.01)	500 (—)
HSI–NKY	3×10^{-6}	0.0585	4×10^{-5}	0.8052	0.0927 (0.0793)	0.5270 (0.0555)	0.49 (0.09)	9.43 (2.40)

Table 6: Results (p–values) from bivariate tests of constant correlations against an STCC–GARCH model with single or double transition, and bivariate tests of another transition in TVCC–GARCH model. The last four columns report the estimation results for each of the bivariate TVCC–GARCH model. The univariate volatilities are modelled as GJR–GARCH. The standard errors are given in parentheses.

GARCH-parameters:

	CAC	DAX	FTSE	HSI	NKY
α_0	0.1685 (0.0448)	0.1799 (0.0506)	0.1287 (0.0349)	0.2582 (0.1059)	0.6148 (0.1299)
α_1	0.0277 (0.0117)	0.0175 (0.0122)	0.0074 (0.0106)	0.0628 (0.0185)	0.0106 (0.0250)
β_1	0.9236 (0.0129)	0.9276 (0.0139)	0.9236 (0.0150)	0.8945 (0.0218)	0.8683 (0.0498)
δ_1	0.0432 (0.0165)	0.0590 (0.0182)	0.0941 (0.0208)	0.0388 (0.0297)	0.1042 (0.0409)

s_{1t}

Correlation parameters:

$\boldsymbol{P}_{(21)}$

	CAC	DAX	FTSE	HSI
DAX	0.7437 (0.0455)			
FTSE	0.5136 (0.0370)	0.6130 (0.0426)		
HSI	0.2891^{r_1} (0.0425)	0.3261^{r_1} (0.0443)	0.4282 (0.0463)	
NKY	0.3075^{r_1} (0.0422)	$0.2491^{r_1,r_2}$ (0.0391)	$0.2606^{r_1,r_2}$ (0.0374)	0.1630^{r_1} (0.0471)

$\boldsymbol{P}_{(22)}$

	CAC	DAX	FTSE	HSI
DAX	0.8709 (0.0078)			
FTSE	0.8960^{r_1} (0.0119)	0.8250^{r_1} (0.0183)		
HSI	0.5555^{r_1} (0.0438)	0.6730 (0.0523)	0.5963 (0.0513)	
NKY	0.2146 (0.0540)	0.2491^{r_2} (0.0391)	0.2606^{r_2} (0.0374)	0.5594^{r_1} (0.0465)

$\boldsymbol{P}_{(11)}$

	CAC	DAX	FTSE	HSI
DAX	0.6280 (0.0301)			
FTSE	0.5934 (0.0303)	0.5151 (0.0371)		
HSI	0.2891^{r_1} (0.0425)	0.3261^{r_1} (0.0443)	0.3533 (0.0411)	
NKY	0.3075^{r_1} (0.0422)	0.2491^{r_1} (0.0391)	0.2606^{r_1} (0.0374)	0.1630^{r_1} (0.0471)

$\boldsymbol{P}_{(12)}$

	CAC	DAX	FTSE	HSI
DAX	0.9396 (0.0078)			
FTSE	0.8960^{r_1} (0.0119)	0.8250^{r_1} (0.0183)		
HSI	0.5555^{r_1} (0.0438)	0.5520 (0.0441)	0.5116 (0.0492)	
NKY	0.6657 (0.0386)	0.6823 (0.0359)	0.6350 (0.0425)	0.5594^{r_1} (0.0465)

s_{2t}

Transition parameters:

	G_{1t}
c_1	22.91 (0.05)
γ_1	67.57 (14.35)

	G_{2t}
c_2	0.59 (0.02)
γ_2	17.62 (2.16)

Table 7: Estimation results for the TVSTCC–GARCH model. The transition variable s_{1t} is one week lag of the volatility index (VIX) and s_{2t} is time in percentage. The correlation matrices $\boldsymbol{P}_{(11)}$ and $\boldsymbol{P}_{(21)}$ refer to the extreme states according to the VIX at the beginning of the sample, and similarly matrices $\boldsymbol{P}_{(12)}$ and $\boldsymbol{P}_{(22)}$ refer to those at the end of the sample. The parameters indicated by a superscript r_1 are restricted constant with respect to the transition variable s_{1t} (vertical direction in the Table), and the parameters with a superscript r_2 are restricted constant with respect to the transitin variable s_{2t} (time-dimension, horizontal direction in the Table). The transition functions G_{1t} and G_{2t} are functions of s_{1t} and s_{2t}, respectively. The standard errors are given in parentheses.

Appendix

Construction of the auxiliary null hypothesis 1: Test for another transition

The null hypothesis for the test for another transition is $\gamma_2 = 0$ in (6). When the null hypothesis is true, some of the parameters in the model cannot be identified. This problem is circumvented following Luukkonen, Saikkonen, and Teräsvirta (1988). Linearizing the transition function G_{2t} by a first-order Taylor approximation around $\gamma_2 = 0$ yields

$$G_{2t} \quad \tilde{=} \quad 1/2 + 1/4\gamma_2(s_{2t} - c_2) + \boldsymbol{R} \tag{18}$$

where $\boldsymbol{R}$ is the remainder that equals zero when the null hypothesis is valid. Thus, ignoring $\boldsymbol{R}$ and inserting (18) into (6) the dynamic correlations become

$$\begin{aligned} \boldsymbol{P}_t^* \quad = \quad & (1 - G_{1t})(1/2 - 1/4\gamma_2(s_{2t} - c_2))\boldsymbol{P}_{(11)} + (1 - G_{1t})(1/2 + 1/4\gamma_2(s_{2t} - c_2))\boldsymbol{P}_{(12)} \\ & + G_{1t}(1/2 - 1/4\gamma_2(s_{2t} - c_2))\boldsymbol{P}_{(21)} + G_{1t}(1/2 + 1/4\gamma_2(s_{2t} - c_2))\boldsymbol{P}_{(21)}. \end{aligned}$$

Rearranging the terms yields

$$\boldsymbol{P}_t^* \quad = \quad (1 - G_{1t})\boldsymbol{P}_{(1)}^* + G_{1t}\boldsymbol{P}_{(2)}^* + s_{2t}\boldsymbol{P}_{(3)}^* \tag{19}$$

where

$$\begin{aligned} \boldsymbol{P}_{(1)}^* \quad &= \quad 1/2(\boldsymbol{P}_{(11)} + \boldsymbol{P}_{(12)}) + 1/4c_2\gamma_2(\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(12)}) \\ \boldsymbol{P}_{(2)}^* \quad &= \quad 1/2(\boldsymbol{P}_{(21)} + \boldsymbol{P}_{(22)}) + 1/4c_2\gamma_2(\boldsymbol{P}_{(21)} - \boldsymbol{P}_{(22)}) \\ \boldsymbol{P}_{(3)}^* \quad &= \quad -1/4(1 - G_{1t})\gamma_2(\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(12)}) - 1/4G_{1t}\gamma_2(\boldsymbol{P}_{(21)} - \boldsymbol{P}_{(22)}). \end{aligned}$$

Under H_0, $\gamma_2 = 0$ and hence

$$\begin{aligned} \boldsymbol{P}_{(1)}^* \quad &= \quad 1/2(\boldsymbol{P}_{(11)} + \boldsymbol{P}_{(12)}) \\ \boldsymbol{P}_{(2)}^* \quad &= \quad 1/2(\boldsymbol{P}_{(21)} + \boldsymbol{P}_{(22)}) \\ \boldsymbol{P}_{(3)}^* \quad &= \quad \boldsymbol{0}_{N\times N} \end{aligned}$$

and the model collapses to the STCC–GARCH model with correlations varying according to s_{1t}. The auxiliary null hypothesis is therefore

$$H_0^{aux} \quad : \quad vecl\boldsymbol{P}_{(3)}^* = \boldsymbol{0}_{N(N-1)/2\times 1}$$

in (19).[3] The LM–test of this auxiliary null hypothesis is carried out in the usual way and the test statistic is χ^2 distributed with $N(N-1)/2$ degrees of freedom. The construction of the LM–test is discussed in a later subsection.

Construction of the auxiliary null hypothesis 2: Test of constant correlations against an STCC–GARCH model with two transitions

The constancy of correlations hypothesis is equivalent to $\gamma_1 = \gamma_2 = 0$ in (6). The problem with unidentified parameters under the null is avoided by linearizing both transition functions, G_{1t} and G_{2t}, by first-order Taylor expansions around $\gamma_1 = 0$ and $\gamma_2 = 0$, respectively. This yields

$$G_{it} \quad \tilde{=} \quad 1/2 + 1/4\gamma_i(s_{it} - c_i) + \boldsymbol{R}_i, \qquad i = 1, 2. \tag{20}$$

G_{it} does not carry any approximation error under the null. Replacing the transition functions in (6) with the equations (20) gives

$$\begin{aligned} \boldsymbol{P}_t^* \quad = \quad & (1/2 - 1/4\gamma_1(s_{1t} - c_1))(1/2 - 1/4\gamma_2(s_{2t} - c_2))\boldsymbol{P}_{(11)} \\ & + (1/2 - 1/4\gamma_1(s_{1t} - c_1))(1/2 + 1/4\gamma_2(s_{2t} - c_2))\boldsymbol{P}_{(12)} \\ & + (1/2 + 1/4\gamma_1(s_{1t} - c_1))(1/2 - 1/4\gamma_2(s_{2t} - c_2))\boldsymbol{P}_{(21)} \\ & + (1/2 + 1/4\gamma_1(s_{1t} - c_1))(1/2 + 1/4\gamma_2(s_{2t} - c_2))\boldsymbol{P}_{(22)}. \end{aligned}$$

[3] The notation $vecl\boldsymbol{P}$ is used to denote the vec-operator applied to the strictly lower triangular part of the square matrix $\boldsymbol{P}$.

Rearranging the terms gives

$$\boldsymbol{P}_t^* = \boldsymbol{P}_{(1)}^* + s_{1t}\boldsymbol{P}_{(2)}^* + s_{2t}\boldsymbol{P}_{(3)}^* + s_{1t}s_{2t}\boldsymbol{P}_{(4)}^* \tag{21}$$

where

$$\begin{aligned}
\boldsymbol{P}_{(1)}^* &= 1/4(\boldsymbol{P}_{(11)} + \boldsymbol{P}_{(12)} + \boldsymbol{P}_{(21)} + \boldsymbol{P}_{(22)}) + 1/8c_1\gamma_1(\boldsymbol{P}_{(11)} + \boldsymbol{P}_{(12)} - \boldsymbol{P}_{(21)} - \boldsymbol{P}_{(22)}) \\
&\quad +1/8c_2\gamma_2(\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(12)} + \boldsymbol{P}_{(21)} - \boldsymbol{P}_{(22)}) + 1/16c_1c_2\gamma_1\gamma_2(\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(12)} - \boldsymbol{P}_{(21)} + \boldsymbol{P}_{(22)}) \\
\boldsymbol{P}_{(2)}^* &= -1/8\gamma_1(\boldsymbol{P}_{(11)} + \boldsymbol{P}_{(12)} - \boldsymbol{P}_{(21)} - \boldsymbol{P}_{(22)}) - 1/16c_2\gamma_1\gamma_2(\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(12)} - \boldsymbol{P}_{(21)} + \boldsymbol{P}_{(22)}) \\
\boldsymbol{P}_{(3)}^* &= -1/8\gamma_2(\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(12)} + \boldsymbol{P}_{(21)} - \boldsymbol{P}_{(22)}) - 1/16c_1\gamma_1\gamma_2(\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(12)} - \boldsymbol{P}_{(21)} + \boldsymbol{P}_{(22)}) \\
\boldsymbol{P}_{(4)}^* &= 1/16\gamma_1\gamma_2(\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(12)} - \boldsymbol{P}_{(21)} + \boldsymbol{P}_{(22)}).
\end{aligned}$$

Under H_0, $\gamma_1 = \gamma_2 = 0$ and hence

$$\begin{aligned}
\boldsymbol{P}_{(1)}^* &= 1/4(\boldsymbol{P}_{(11)} + \boldsymbol{P}_{(12)} + \boldsymbol{P}_{(21)} + \boldsymbol{P}_{(22)}) \\
\boldsymbol{P}_{(2)}^* &= \boldsymbol{0}_{N\times N} \\
\boldsymbol{P}_{(3)}^* &= \boldsymbol{0}_{N\times N} \\
\boldsymbol{P}_{(4)}^* &= \boldsymbol{0}_{N\times N}.
\end{aligned}$$

Therefore, the auxiliary null hypothesis is stated as

$$H_0^{aux} \quad : \quad vecl\boldsymbol{P}_{(2)}^* = vecl\boldsymbol{P}_{(3)}^* = vecl\boldsymbol{P}_{(4)}^* = \boldsymbol{0}_{N(N-1)/2\times 1}$$

in (21). The test statistic for the LM–test for this auxiliary null hypothesis is χ^2 distributed with $3N(N-1)/2$ degrees of freedom. The details of the construction of the LM–test are discussed in a later subsection.

Construction of the auxiliary null hypothesis 3: Test of constant correlations against an STCC–GARCH model with two transitions, transition variables are independent

The constancy of correlations hypothesis is imposed by setting $\gamma_1 = \gamma_2 = 0$ in (10). The problem with unidentified parameters under the null is avoided by linearizing both transition functions, G_{1t} and G_{2t}, by first-order Taylor expansions around $\gamma_1 = 0$ and $\gamma_2 = 0$, respectively. Replacing the transition functions in (10) by the linearized ones (20) gives

$$\begin{aligned}
\boldsymbol{P}_t^* &= (-1/4\gamma_1(s_{1t} - c_1) - 1/4\gamma_2(s_{2t} - c_2))\boldsymbol{P}_{(11)} \\
&\quad +(1/2 + 1/4\gamma_1(s_{1t} - c_1))\boldsymbol{P}_{(21)} \\
&\quad +(1/2 + 1/4\gamma_2(s_{2t} - c_2))\boldsymbol{P}_{(12)}.
\end{aligned}$$

Rearranging the terms gives

$$\boldsymbol{P}_t^* = \boldsymbol{P}_{(1)}^* + s_{1t}\boldsymbol{P}_{(2)}^* + s_{2t}\boldsymbol{P}_{(3)}^* \tag{22}$$

where

$$\begin{aligned}
\boldsymbol{P}_{(1)}^* &= 1/2(\boldsymbol{P}_{(12)} + \boldsymbol{P}_{(21)}) + 1/4c_1\gamma_1(\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(21)}) + 1/4c_2\gamma_2(\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(12)}) \\
\boldsymbol{P}_{(2)}^* &= 1/4\gamma_1(\boldsymbol{P}_{(21)} - \boldsymbol{P}_{(11)}) \\
\boldsymbol{P}_{(3)}^* &= 1/4\gamma_2(\boldsymbol{P}_{(12)} - \boldsymbol{P}_{(11)}).
\end{aligned}$$

Under H_0, $\gamma_1 = \gamma_2 = 0$ and hence

$$\begin{aligned}
\boldsymbol{P}_{(1)}^* &= 1/2(\boldsymbol{P}_{(12)} + \boldsymbol{P}_{(21)}) \\
\boldsymbol{P}_{(2)}^* &= \boldsymbol{0}_{N\times N} \\
\boldsymbol{P}_{(3)}^* &= \boldsymbol{0}_{N\times N}.
\end{aligned}$$

Therefore, the auxiliary null hypothesis is stated as

$$H_0^{aux} \quad : \quad vecl\boldsymbol{P}_{(2)}^* = vecl\boldsymbol{P}_{(3)}^* = \boldsymbol{0}_{N(N-1)/2\times 1}$$

in (22). The test statistic for the LM–test for this auxiliary null hypothesis is χ^2 distributed with $N(N-1)$ degrees of freedom. The details of the construction of the LM–test are discussed in a later subsection.

Construction of LM(/Wald)–statistic

Let $\boldsymbol{\theta}_0$ be the vector of true parameters. Under suitable assumptions and regularity conditions,

$$\sqrt{T}^{-1}\frac{\partial l\left(\boldsymbol{\theta}_0\right)}{\partial\boldsymbol{\theta}}\xrightarrow{d}N\left(0,\Im\left(\boldsymbol{\theta}_0\right)\right). \tag{23}$$

To derive LM–statistics of the null hypothesis consider the following quadratic form:

$$T^{-1}\frac{\partial l\left(\boldsymbol{\theta}_0\right)}{\partial\boldsymbol{\theta}'}\Im\left(\boldsymbol{\theta}_0\right)^{-1}\frac{\partial l\left(\boldsymbol{\theta}_0\right)}{\partial\boldsymbol{\theta}}=T^{-1}\left(\sum_{t=1}^{T}\frac{\partial l_t\left(\boldsymbol{\theta}_0\right)}{\partial\boldsymbol{\theta}'}\right)\Im\left(\boldsymbol{\theta}_0\right)^{-1}\left(\sum_{t=1}^{T}\frac{\partial l_t\left(\boldsymbol{\theta}_0\right)}{\partial\boldsymbol{\theta}}\right)$$

and evaluate it at the maximum likelihood estimators under the restriction of the null hypothesis. The limiting information matrix $\Im(\boldsymbol{\theta}_0)$ is replaced by the consistent estimator

$$\hat{\Im}_T\left(\boldsymbol{\theta}_0\right)=T^{-1}\sum_{t=1}^{T}E\left[\frac{\partial l_t\left(\boldsymbol{\theta}_0\right)}{\partial\boldsymbol{\theta}}\frac{\partial l_t\left(\boldsymbol{\theta}_0\right)}{\partial\boldsymbol{\theta}'}\mid\mathcal{F}_{t-1}\right]. \tag{24}$$

The following derivations are straightforward implications of the definitions and elementary rules of matrix algebra. Results in Anderson (2003) and Lütkepohl (1996) are heavily relied upon. Note that the following results are derived placing the restriction $\delta_{ij}=0$ in (2). It is straightforward extension to modify the results to account for the asymmetric effects. For the modifications needed, the reader is asked to refer to the appendix in Silvennoinen and Teräsvirta (2005).

Test of constant conditional correlations against a DSTCC–GARCH model

The model under the null is the CCC–GARCH model. The alternative model is an STCC–GARCH model with two transitions where the correlations are controlled by the transition variables s_{1t} and s_{2t}. Under the null, the linearized time-varying correlation matrix is $\boldsymbol{P}_t^*=\boldsymbol{P}_{(1)}^*+s_{1t}\boldsymbol{P}_{(2)}^*+s_{2t}\boldsymbol{P}_{(3)}^*+s_{1t}s_{2t}\boldsymbol{P}_{(4)}^*$ as defined in (21). To construct the test statistic we introduce some simplifying notation. Let $\boldsymbol{\omega}_i=(\alpha_{i0},\alpha_i,\beta_i)'$, $i=1,\ldots,N$, denote the parameter vectors of the GARCH equations, and $\boldsymbol{\rho}^*=(\boldsymbol{\rho}_{(1)}^{*\prime},\ldots,\boldsymbol{\rho}_{(4)}^{*\prime})'$, where $\boldsymbol{\rho}_{(j)}^*=vecl\boldsymbol{P}_{(j)}^*$, $j=1,\ldots,4$, are the vectors holding all unique off-diagonal elements in the four matrices $\boldsymbol{P}_{(1)}^*,\ldots,\boldsymbol{P}_{(4)}^*$, respectively. Let $\boldsymbol{\theta}=(\boldsymbol{\omega}_1',\ldots,\boldsymbol{\omega}_N',\boldsymbol{\rho}^{*\prime})'$ be the full parameter vector and $\boldsymbol{\theta}_0$ the corresponding vector of true parameters under the null. Furthermore, let $\boldsymbol{v}_{it}=(1,\varepsilon_{it}^2,h_{it})'$, $i=1,\ldots,N$, and $\boldsymbol{v}_{\boldsymbol{\rho}^*t}=(1,s_{1t},s_{2t},s_{1t}s_{2t})'$. Symbols $\otimes$ and $\odot$ represent the Kronecker and Hadamard products of two matrices, respectively. Let $\boldsymbol{e}_i$ be an $N\times 1$ vector of zeros with ith element equal to one and $\mathbf{1}_n$ be an $n\times n$ matrix of ones. The identity matrix $\boldsymbol{I}$ is of size N unless otherwise indicated by a subscript.

Consider the log-likelihood function for observation t as defined in (11) with linearized time-varying correlation matrix:

$$l_t\left(\boldsymbol{\theta}\right)=-\frac{N}{2}\log\left(2\pi\right)-\frac{1}{2}\sum_{i=1}^{N}\log\left(h_{it}\right)-\frac{1}{2}\log\left|\boldsymbol{P}_t^*\right|-\frac{1}{2}\boldsymbol{z}_t'\boldsymbol{P}_t^{*-1}\boldsymbol{z}_t.$$

The first order derivatives of the log-likelihood function with respect to the GARCH and correlation parameters are

$$\begin{aligned}\frac{\partial l_t\left(\boldsymbol{\theta}\right)}{\partial\boldsymbol{\omega}_i}&=-\frac{1}{2h_{it}}\frac{\partial h_{it}}{\partial\boldsymbol{\omega}_i}\left\{1-z_{it}\boldsymbol{e}_i'\boldsymbol{P}_t^{*-1}\boldsymbol{z}_t\right\},\quad i=1,\ldots,N\\ \frac{\partial l_t\left(\boldsymbol{\theta}\right)}{\partial\boldsymbol{\rho}^*}&=-\frac{1}{2}\frac{\partial\left(vec\boldsymbol{P}_t^*\right)'}{\partial\boldsymbol{\rho}^*}\left\{vec\boldsymbol{P}_t^{*-1}-\left(\boldsymbol{P}_t^{*-1}\otimes\boldsymbol{P}_t^{*-1}\right)\left(\boldsymbol{z}_t\otimes\boldsymbol{z}_t\right)\right\}\end{aligned}$$

where

$$\begin{aligned}\frac{\partial h_{it}}{\partial\boldsymbol{\omega}_i}&=\boldsymbol{v}_{i,t-1}+\beta_i\frac{\partial h_{i,t-1}}{\partial\boldsymbol{\omega}_i},\quad i=1,\ldots,N\\ \frac{\partial\left(vec\boldsymbol{P}_t^*\right)'}{\partial\boldsymbol{\rho}^*}&=\boldsymbol{v}_{\boldsymbol{\rho}^*t}\otimes\boldsymbol{U}'.\end{aligned}$$

The matrix $\boldsymbol{U}$ is an $N^2 \times \frac{N(N-1)}{2}$ matrix of zeros and ones, whose columns are defined as

$$\left[vec\left(\boldsymbol{e}_i\boldsymbol{e}_j' + \boldsymbol{e}_j\boldsymbol{e}_i'\right)\right]_{i=1,\ldots N-1, j=i+1,\ldots,N}$$

and the columns appear in the same order from left to right as the indices in $vecl\boldsymbol{P}_t$. Under the null hypothesis $\boldsymbol{\rho}^*_{(2)} = \boldsymbol{\rho}^*_{(3)} = \boldsymbol{\rho}^*_{(4)} = 0$, and thus the derivatives at the true parameter values under the null can be written as

$$\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\omega}_i} = -\frac{1}{2h_{it}}\frac{\partial h_{it}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\omega}_i}\left\{1 - z_{it}\boldsymbol{e}_i'\boldsymbol{P}_{(1)}^{*-1}\boldsymbol{z}_t\right\}, \quad i = 1,\ldots,N \tag{25}$$

$$\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\rho}^*} = -\frac{1}{2}\frac{\partial\left(vec\boldsymbol{P}_t^*(\boldsymbol{\theta}_0)\right)'}{\partial \boldsymbol{\rho}^*}\left\{vec\boldsymbol{P}_{(1)}^{*-1} - \left(\boldsymbol{P}_{(1)}^{*-1}\otimes\boldsymbol{P}_{(1)}^{*-1}\right)(\boldsymbol{z}_t\otimes\boldsymbol{z}_t)\right\}. \tag{26}$$

Taking conditional expectations of the cross products of (25) and (26) yields, for $i,j = 1,\ldots,N$,

$$\begin{aligned}
E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\omega}_i}\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\omega}_i'}\right] &= \frac{1}{4h_{it}^2}\frac{\partial h_{it}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\omega}_i}\frac{\partial h_{it}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\omega}_i'}\left(1 + \boldsymbol{e}_i'\boldsymbol{P}_{(1)}^{*-1}\boldsymbol{e}_i\right)\\
E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\omega}_i}\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\omega}_j'}\right] &= \frac{1}{4h_{it}h_{jt}}\frac{\partial h_{it}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\omega}_i}\frac{\partial h_{jt}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\omega}_j'}\left(\rho_{1,ij}^*\boldsymbol{e}_i'\boldsymbol{P}_{(1)}^{*-1}\boldsymbol{e}_j\right), \quad i \neq j\\
E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\rho}^*}\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\rho}^{*\prime}}\right] &= \frac{1}{4}\frac{\partial\left(vec\boldsymbol{P}_t^*(\boldsymbol{\theta}_0)\right)'}{\partial \boldsymbol{\rho}^*}\left(\boldsymbol{P}_{(1)}^{*-1}\otimes\boldsymbol{P}_{(1)}^{*-1} + \left(\boldsymbol{P}_{(1)}^{*-1}\otimes\boldsymbol{I}\right)\boldsymbol{K}\left(\boldsymbol{P}_{(1)}^{*-1}\otimes\boldsymbol{I}\right)\right)\frac{\partial vec\boldsymbol{P}_t^*(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\rho}^{*\prime}}\\
E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\omega}_i}\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\rho}^{*\prime}}\right] &= \frac{1}{4h_{it}}\frac{\partial h_{it}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\omega}_i}\left(\boldsymbol{e}_i'\boldsymbol{P}_{(1)}^{*-1}\otimes\boldsymbol{e}_i' + \boldsymbol{e}_i'\otimes\boldsymbol{e}_i'\boldsymbol{P}_{(1)}^{*-1}\right)\frac{\partial vec\boldsymbol{P}_t^*(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\rho}^{*\prime}}
\end{aligned} \tag{27}$$

where

$$\boldsymbol{K} = \begin{bmatrix}\boldsymbol{e}_1\boldsymbol{e}_1' & \cdots & \boldsymbol{e}_N\boldsymbol{e}_1'\\ \vdots & \ddots & \vdots\\ \boldsymbol{e}_1\boldsymbol{e}_N' & \cdots & \boldsymbol{e}_N\boldsymbol{e}_N'\end{bmatrix}. \tag{28}$$

For the derivation of the expressions (27), see Silvennoinen and Teräsvirta (2005).

The estimator of the information matrix is obtained by making use of the submatrices in (27). For a more compact expression, let $\boldsymbol{x}_t = (\boldsymbol{x}_{1t}',\ldots,\boldsymbol{x}_{Nt}')'$ where $\boldsymbol{x}_{it} = -\frac{1}{2h_{it}}\frac{\partial h_{it}}{\partial \boldsymbol{\omega}_i}$, and let $\boldsymbol{x}_{\boldsymbol{\rho}^*t} = -\frac{1}{2}\frac{\partial(vec\boldsymbol{P}_t^*)'}{\partial \boldsymbol{\rho}^*}$, and let $\boldsymbol{x}_{it}^0$, $i = 1,\ldots,N, \boldsymbol{\rho}^*$, denote the corresponding expressions evaluated at the true values under the null hypothesis. Setting

$$\begin{aligned}
\boldsymbol{M}_1 &= T^{-1}\sum_{t=1}^{T}\boldsymbol{x}_t^0\boldsymbol{x}_t^{0\prime}\odot\left(\left(\boldsymbol{I} + \boldsymbol{P}_{(1)}^*\odot\boldsymbol{P}_{(1)}^{*-1}\right)\otimes\boldsymbol{1}_3\right)\\
\boldsymbol{M}_2 &= T^{-1}\sum_{t=1}^{T}\begin{bmatrix}\boldsymbol{x}_{1t}^0 & & 0\\ & \ddots & \\ 0 & & \boldsymbol{x}_{Nt}^0\end{bmatrix}\begin{bmatrix}\boldsymbol{e}_1'\boldsymbol{P}_{(1)}^{*-1}\otimes\boldsymbol{e}_1' + \boldsymbol{e}_1'\otimes\boldsymbol{e}_1'\boldsymbol{P}_{(1)}^{*-1}\\ \vdots\\ \boldsymbol{e}_N'\boldsymbol{P}_{(1)}^{*-1}\otimes\boldsymbol{e}_N' + \boldsymbol{e}_N'\otimes\boldsymbol{e}_N'\boldsymbol{P}_{(1)}^{*-1}\end{bmatrix}\boldsymbol{x}_{\boldsymbol{\rho}^*t}^{0\prime}\\
\boldsymbol{M}_3 &= T^{-1}\sum_{t=1}^{T}\boldsymbol{x}_{\boldsymbol{\rho}^*t}^0\left(\boldsymbol{P}_{(1)}^{*-1}\otimes\boldsymbol{P}_{(1)}^{*-1} + \left(\boldsymbol{P}_{(1)}^{*-1}\otimes\boldsymbol{I}\right)\boldsymbol{K}\left(\boldsymbol{P}_{(1)}^{*-1}\otimes\boldsymbol{I}\right)\right)\boldsymbol{x}_{\boldsymbol{\rho}^*t}^{0\prime}
\end{aligned}$$

the information matrix $\mathfrak{I}(\boldsymbol{\theta}_0)$ is approximated by

$$\begin{aligned}
\hat{\mathfrak{I}}_T(\boldsymbol{\theta}_0) &= T^{-1}\sum_{t=1}^{T}E\left[\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}}\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'} \mid \mathcal{F}_{t-1}\right]\\
&= \begin{bmatrix}\boldsymbol{M}_1 & \boldsymbol{M}_2\\ \boldsymbol{M}_2' & \boldsymbol{M}_3\end{bmatrix}.
\end{aligned}$$

The block corresponding to the correlation parameters of the inverse of $\hat{\mathfrak{I}}_T(\boldsymbol{\theta}_0)$ can be calculated as

$$\left(\boldsymbol{M}_3 - \boldsymbol{M}_2'\boldsymbol{M}_1^{-1}\boldsymbol{M}_2\right)^{-1}$$

from where the south-east $\frac{3N(N-1)}{2}\times\frac{3N(N-1)}{2}$ block corresponding to $\boldsymbol{\rho}^*_{(2)}$, $\boldsymbol{\rho}^*_{(3)}$, and $\boldsymbol{\rho}^*_{(4)}$ can be extracted. Replacing the true unknown values with maximum likelihood estimators, the test statistic simplifies to

$$T^{-1}\left(\sum_{t=1}^{T}\frac{\partial l_t(\hat{\boldsymbol{\theta}})}{\partial(\boldsymbol{\rho}_{(2)}^{*\prime},\boldsymbol{\rho}_{(3)}^{*\prime},\boldsymbol{\rho}_{(4)}^{*\prime})}\right)\left[\hat{\mathfrak{I}}_T(\hat{\boldsymbol{\theta}})\right]^{-1}_{(\boldsymbol{\rho}_{(2-4)}^*,\boldsymbol{\rho}_{(2-4)}^*)}\left(\sum_{t=1}^{T}\frac{\partial l_t(\hat{\boldsymbol{\theta}})}{\partial(\boldsymbol{\rho}_{(2)}^{*\prime},\boldsymbol{\rho}_{(3)}^{*\prime},\boldsymbol{\rho}_{(4)}^{*\prime})'}\right) \tag{29}$$

where $[\hat{\mathfrak{I}}_T(\hat{\boldsymbol{\theta}})]^{-1}_{(\boldsymbol{\rho}^*_{(2-4)},\boldsymbol{\rho}^*_{(2-4)})}$ is the block of the inverse of $\hat{\mathfrak{I}}_T$ corresponding to those correlation parameters that are set to zero under the null. It follows from (23) and consistency and asymptotic normality of ML estimators that the statistic (29) has an asymptotic $\chi^2_{\frac{3N(N-1)}{2}}$ distribution when the null hypothesis is valid.

Test of constant conditional correlations against a partially constant DSTCC–GARCH model

The test of constant correlations of previous subsection is not affected unless one or more of the parameters are restricted to be constant according to one of the transition variables in both extreme states described by the other transition variable. In those cases certain parameters in the linearized time-varying correlation matrix $\boldsymbol{P}^*_t$ in (21) are set to zero. Let there be k pairs of correlation parameters that, under the alternative hypothesis, are restricted to be constant with respect to the transition variable s_{t1} in both extreme states described by s_{t2}. That is, there are k pairs of restrictions as follows:

$$\rho_{(11)ij} = \rho_{(21)ij} \quad \text{and} \quad \rho_{(12)ij} = \rho_{(22)ij}, \quad i > j$$

where $\rho_{(mn)ij}$ is the ij-element of the correlation matrix $\boldsymbol{P}_{(mn)}$ in (6). In the linearized correlation matrix $\boldsymbol{P}^*_t$, the k elements in each of the matrices $\boldsymbol{P}^*_{(2)}$ and $\boldsymbol{P}^*_{(4)}$ corresponding to these restrictions are set to zero. Similarly, let there be l pairs of correlation parameters that, under the alternative hypothesis, are restricted to be constant with respect to s_{t2} in both extreme states described by s_{t1}. The pairs of restrictions are then

$$\rho_{(11)ij} = \rho_{(12)ij} \quad \text{and} \quad \rho_{(21)ij} = \rho_{(22)ij}, \quad i > j.$$

In the linearized correlation matrix $\boldsymbol{P}^*_t$ the l elements in each of the matrices $\boldsymbol{P}^*_{(3)}$ and $\boldsymbol{P}^*_{(4)}$ corresponding to these restrictions are set to zero. The vector of correlation parameters $\boldsymbol{\rho}^*$ is formed as before but the elements corresponding to the restrictions, i.e., the elements that were set to zero, are excluded. Furthermore,

$$\frac{\partial (vec\boldsymbol{P}^*_t)'}{\partial \boldsymbol{\rho}^*}$$

is defined as before, but with m (m equals $2k + 2l$ less the number of possibly overlapping restrictions) rows deleted so that the remaining rows correspond to the elements in $\boldsymbol{\rho}^*$. The same rows are also deleted from $\boldsymbol{x}_{\rho^* t}$. With these modifications the test statistic is as in (29) above, and its asymptotic distribution under the null hypothesis is $\chi^2_{\frac{3N(N-1)}{2}-m}$.

Test of constant conditional correlations against a DSTCC–GARCH model whose transition variables are independent

The model under the null is the CCC–GARCH model. The alternative model is an STCC–GARCH model with two transitions where the correlations are varying according to the transition variables s_{1t} and s_{2t} and the restriction $\boldsymbol{P}_{(11)} - \boldsymbol{P}_{(12)} = \boldsymbol{P}_{(21)} - \boldsymbol{P}_{(22)}$ holds. Under the null, the linearized time-varying correlation matrix is $\boldsymbol{P}^*_t = \boldsymbol{P}^*_{(1)} + s_{1t}\boldsymbol{P}^*_{(2)} + s_{2t}\boldsymbol{P}^*_{(3)}$ as defined in (22). The statistic is constructed as in the case of testing constancy of correlations against DSTCC–GARCH model but with following modifications: Let $\boldsymbol{\rho}^* = (\boldsymbol{\rho}^{*\prime}_{(1)}, \boldsymbol{\rho}^{*\prime}_{(2)}, \boldsymbol{\rho}^{*\prime}_{(3)})'$, where $\boldsymbol{\rho}^*_{(j)} = vecl\boldsymbol{P}^*_{(j)}$, $j = 1, 2, 3$, are the vectors holding all the unique off-diagonal elements in the four matrices $\boldsymbol{P}^*_{(1)}, \boldsymbol{P}^*_{(2)}, \boldsymbol{P}^*_{(3)}$, respectively. Furthermore, define $\boldsymbol{v}_{\boldsymbol{\rho}^* t} = (1, s_{1t}, s_{2t})'$. With these changes the test statistic is as constructed as before, and the block corresponding to the correlation parameters of the inverse of $\hat{\mathfrak{I}}_T(\boldsymbol{\theta}_0)$ can be calculated as

$$\left(\boldsymbol{M}_3 - \boldsymbol{M}_2'\boldsymbol{M}_1^{-1}\boldsymbol{M}_2\right)^{-1}$$

from where the south-east $N(N-1)\times N(N-1)$ block corresponding to $\boldsymbol{\rho}^*_{(2)}$ and $\boldsymbol{\rho}^*_{(3)}$ can be extracted. Replacing the true unknown values with maximum likelihood estimators, the test statistic simplifies to

$$T^{-1}\left(\sum_{t=1}^{T}\frac{\partial l_t(\hat{\boldsymbol{\theta}})}{\partial(\boldsymbol{\rho}^{*\prime}_{(2)},\boldsymbol{\rho}^{*\prime}_{(3)})}\right)\left[\hat{\mathfrak{I}}_T(\hat{\boldsymbol{\theta}})\right]^{-1}_{(\boldsymbol{\rho}^*_{(2-3)},\boldsymbol{\rho}^*_{(2-3)})}\left(\sum_{t=1}^{T}\frac{\partial l_t(\hat{\boldsymbol{\theta}})}{\partial(\boldsymbol{\rho}^{*\prime}_{(2)},\boldsymbol{\rho}^{*\prime}_{(3)})'}\right) \tag{30}$$

where $[\hat{\mathfrak{I}}_T(\hat{\boldsymbol{\theta}})]^{-1}_{(\boldsymbol{\rho}^*_{(2-3)},\boldsymbol{\rho}^*_{(2-3)})}$ is the block of the inverse of $\hat{\mathfrak{I}}_T$ corresponding to those correlation parameters that are set to zero under the null. It follows from (23) and consistency and asymptotic normality of ML estimators that the statistic (30) has an asymptotic $\chi^2_{N(N-1)}$ distribution when the null hypothesis is valid.

Testing for the additional transition in the STCC–GARCH model

The model under the null is an STCC–GARCH model where the correlations are varying according to the transition variable s_{1t}. The transition that we wish to test for is a function of s_{2t}. Under the null, the linearized time-varying correlation matrix is $\boldsymbol{P}^*_t=(1-G_{1t})\boldsymbol{P}^*_{(1)}+G_{1t}\boldsymbol{P}^*_{(2)}+s_{2t}\boldsymbol{P}^*_{(3)}$ as defined in (19). The notation is as in the previous subsection with the following modifications: Let $\boldsymbol{\rho}^*=(\boldsymbol{\rho}^{*\prime}_{(1)},\boldsymbol{\rho}^{*\prime}_{(2)},\boldsymbol{\rho}^{*\prime}_{(3)})'$ where $\boldsymbol{\rho}^*_{(j)}=vecl\boldsymbol{P}^*_{(j)}$, $i=1,\ldots,3$, and $\boldsymbol{\varphi}=(c_1,\gamma_1)'$. Let $\boldsymbol{\theta}=(\boldsymbol{\omega}'_1,\ldots,\boldsymbol{\omega}'_N,\boldsymbol{\rho}^{*\prime},\boldsymbol{\varphi}')'$ be the full parameter vector, and $\boldsymbol{\theta}_0$ the corresponding vector of the true parameters under the null. Let $\boldsymbol{v}_{\boldsymbol{\rho}^*t}=(1-G_{1t},G_{1t},s_{2t})'$, and let $\boldsymbol{v}_{\boldsymbol{\varphi}t}=(\gamma_1,c_1-s_{1t})'$.

The first order derivatives of the log-likelihood function with respect to the GARCH, correlation, and transition parameters are

$$\begin{aligned}
\frac{\partial l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\omega}_i} &= -\frac{1}{2h_{it}}\frac{\partial h_{it}}{\partial\boldsymbol{\omega}_i}\left\{1-z_{it}\boldsymbol{e}'_i\boldsymbol{P}^{*-1}_t\boldsymbol{z}_t\right\},\quad i=1,\ldots,N\\
\frac{\partial l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\rho}^*} &= -\frac{1}{2}\frac{\partial(vec\boldsymbol{P}^*_t)'}{\partial\boldsymbol{\rho}^*}\left\{vec\boldsymbol{P}^{*-1}_t-\left(\boldsymbol{P}^{*-1}_t\otimes\boldsymbol{P}^{*-1}_t\right)(\boldsymbol{z}_t\otimes\boldsymbol{z}_t)\right\}\\
\frac{\partial l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\varphi}} &= -\frac{1}{2}\frac{\partial(vec\boldsymbol{P}^*_t)'}{\partial\boldsymbol{\varphi}}\left\{vec\boldsymbol{P}^{*-1}_t-\left(\boldsymbol{P}^{*-1}_t\otimes\boldsymbol{P}^{*-1}_t\right)(\boldsymbol{z}_t\otimes\boldsymbol{z}_t)\right\}
\end{aligned}$$

where

$$\begin{aligned}
\frac{\partial h_{it}}{\partial\boldsymbol{\omega}_i} &= \boldsymbol{v}_{i,t-1}+\beta_i\frac{\partial h_{i,t-1}}{\partial\boldsymbol{\omega}_i},\quad i=1,\ldots,N\\
\frac{\partial(vec\boldsymbol{P}^*_t)'}{\partial\boldsymbol{\rho}^*} &= \boldsymbol{v}_{\boldsymbol{\rho}^*t}\otimes\boldsymbol{U}'\\
\frac{\partial(vec\boldsymbol{P}^*_t)'}{\partial\boldsymbol{\varphi}} &= \boldsymbol{v}_{\boldsymbol{\varphi}t}(1-G_{1t})G_{1t}vec(\boldsymbol{P}^*_{(1)}-\boldsymbol{P}^*_{(2)})'.
\end{aligned}$$

Evaluating the score at the true parameters under the null and taking conditional expectations of the cross products of the first-order derivatives gives

$$\begin{aligned}
E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\omega}_i}\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\omega}'_i}\right] &= \frac{1}{4h^2_{it}}\frac{\partial h_{it}(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\omega}_i}\frac{\partial h_{it}(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\omega}'_i}\left(1+\boldsymbol{e}'_i\boldsymbol{P}^{*0-1}_t\boldsymbol{e}_i\right)\\
E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\omega}_i}\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\omega}'_j}\right] &= \frac{1}{4h_{it}h_{jt}}\frac{\partial h_{it}(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\omega}_i}\frac{\partial h_{jt}(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\omega}'_j}\left(\rho^{*0}_{t,ij}\boldsymbol{e}'_i\boldsymbol{P}^{*0-1}_t\boldsymbol{e}_j\right),\quad i\neq j\\
E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\omega}_i}\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\rho}^{*\prime}}\right] &= \frac{1}{4h_{it}}\frac{\partial h_{it}(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\omega}_i}\left(\boldsymbol{e}'_i\boldsymbol{P}^{*0-1}_t\otimes\boldsymbol{e}'_i+\boldsymbol{e}'_i\otimes\boldsymbol{e}'_i\boldsymbol{P}^{*0-1}_t\right)\frac{\partial vec\boldsymbol{P}^*_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\rho}^{*\prime}}\\
E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\omega}_i}\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\varphi}'}\right] &= \frac{1}{4h_{it}}\frac{\partial h_{it}(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\omega}_i}\left(\boldsymbol{e}'_i\boldsymbol{P}^{*0-1}_t\otimes\boldsymbol{e}'_i+\boldsymbol{e}'_i\otimes\boldsymbol{e}'_i\boldsymbol{P}^{*0-1}_t\right)\frac{\partial vec\boldsymbol{P}^*_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\varphi}'}\\
E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\rho}^*}\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\rho}^{*\prime}}\right] &= \frac{1}{4}\frac{\partial(vec\boldsymbol{P}^*_t(\boldsymbol{\theta}_0))'}{\partial\boldsymbol{\rho}^*}\left(\boldsymbol{P}^{*0-1}_t\otimes\boldsymbol{P}^{*0-1}_t+\left(\boldsymbol{P}^{*0-1}_t\otimes\boldsymbol{I}\right)\boldsymbol{K}\left(\boldsymbol{P}^{*0-1}_t\otimes\boldsymbol{I}\right)\right)\frac{\partial vec\boldsymbol{P}^*_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\rho}^{*\prime}}\\
E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\rho}^*}\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\varphi}'}\right] &= \frac{1}{4}\frac{\partial(vec\boldsymbol{P}^*_t(\boldsymbol{\theta}_0))'}{\partial\boldsymbol{\rho}^*}\left(\boldsymbol{P}^{*t-1}_t\otimes\boldsymbol{P}^{*0-1}_t+\left(\boldsymbol{P}^{*0-1}_t\otimes\boldsymbol{I}\right)\boldsymbol{K}\left(\boldsymbol{P}^{*0-1}_t\otimes\boldsymbol{I}\right)\right)\frac{\partial vec\boldsymbol{P}^*_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\varphi}'}\\
E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\varphi}}\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\varphi}'}\right] &= \frac{1}{4}\frac{\partial(vec\boldsymbol{P}^*_t(\boldsymbol{\theta}_0))'}{\partial\boldsymbol{\varphi}}\left(\boldsymbol{P}^{*0-1}_t\otimes\boldsymbol{P}^{*0-1}_t+\left(\boldsymbol{P}^{*0-1}_t\otimes\boldsymbol{I}\right)\boldsymbol{K}\left(\boldsymbol{P}^{*0-1}_t\otimes\boldsymbol{I}\right)\right)\frac{\partial vec\boldsymbol{P}^*_t(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\varphi}'}
\end{aligned} \tag{31}$$

where $\boldsymbol{K}$ is defined as before.

The estimator of the information matrix is obtained by using the submatrices in (31). To make the expression more compact, let $\boldsymbol{x}_t = (\boldsymbol{x}'_{1t}, \ldots, \boldsymbol{x}'_{Nt})'$ where $\boldsymbol{x}_{it} = -\frac{1}{2h_{it}}\frac{\partial h_{it}}{\partial \omega_i}$. Furthermore, let $\boldsymbol{x}_{\boldsymbol{\rho}^* t} = -\frac{1}{2}\frac{\partial (vec\boldsymbol{P}_t^*)'}{\partial \boldsymbol{\rho}^*}$, and $\boldsymbol{x}_{\boldsymbol{\varphi} t} = -\frac{1}{2}\frac{\partial (vec\boldsymbol{P}_t^*)'}{\partial \boldsymbol{\varphi}}$. Finally, let $\boldsymbol{x}_{it}^0$, $i = 1, \ldots, N, \boldsymbol{\rho}, \boldsymbol{\varphi}$, denote the corresponding expressions evaluated at the true parameters under the null. Setting

$$
\begin{aligned}
\boldsymbol{M}_1 &= T^{-1}\sum_{t=1}^{T} \boldsymbol{x}_t^0 \boldsymbol{x}_t^{0\prime} \odot \left(\left(\boldsymbol{I} + \boldsymbol{P}_t^{*0} \odot \boldsymbol{P}_t^{*0-1}\right) \otimes \boldsymbol{1}_3\right) \\
\boldsymbol{M}_2 &= T^{-1}\sum_{t=1}^{T} \begin{bmatrix} \boldsymbol{x}_{1t}^0 & & 0 \\ & \ddots & \\ 0 & & \boldsymbol{x}_{Nt}^0 \end{bmatrix} \begin{bmatrix} \boldsymbol{e}_1'\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{e}_1' + \boldsymbol{e}_1' \otimes \boldsymbol{e}_1'\boldsymbol{P}_t^{*0-1} \\ \vdots \\ \boldsymbol{e}_N'\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{e}_N' + \boldsymbol{e}_N' \otimes \boldsymbol{e}_N'\boldsymbol{P}_t^{*0-1} \end{bmatrix} \boldsymbol{x}_{\boldsymbol{\rho}^* t}^{0\prime} \\
\boldsymbol{M}_3 &= T^{-1}\sum_{t=1}^{T} \boldsymbol{x}_{\boldsymbol{\rho}^* t}^0 \left(\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{P}_t^{*0-1} + \left(\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{I}\right) \boldsymbol{K} \left(\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{I}\right)\right) \boldsymbol{x}_{\boldsymbol{\rho}^* t}^{0\prime} \\
\boldsymbol{M}_4 &= T^{-1}\sum_{t=1}^{T} \begin{bmatrix} \boldsymbol{x}_{1t}^0 & & 0 \\ & \ddots & \\ 0 & & \boldsymbol{x}_{Nt}^0 \end{bmatrix} \begin{bmatrix} \boldsymbol{e}_1'\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{e}_1' + \boldsymbol{e}_1' \otimes \boldsymbol{e}_1'\boldsymbol{P}_t^{*0-1} \\ \vdots \\ \boldsymbol{e}_N'\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{e}_N' + \boldsymbol{e}_N' \otimes \boldsymbol{e}_N'\boldsymbol{P}_t^{*0-1} \end{bmatrix} \boldsymbol{x}_{\boldsymbol{\varphi} t}^{0\prime} \\
\boldsymbol{M}_5 &= T^{-1}\sum_{t=1}^{T} \boldsymbol{x}_{\boldsymbol{\rho}^* t}^0 \left(\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{P}_t^{*0-1} + \left(\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{I}\right) \boldsymbol{K} \left(\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{I}\right)\right) \boldsymbol{x}_{\boldsymbol{\varphi} t}^{0\prime} \\
\boldsymbol{M}_6 &= T^{-1}\sum_{t=1}^{T} \boldsymbol{x}_{\boldsymbol{\varphi} t}^0 \left(\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{P}_t^{*0-1} + \left(\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{I}\right) \boldsymbol{K} \left(\boldsymbol{P}_t^{*0-1} \otimes \boldsymbol{I}\right)\right) \boldsymbol{x}_{\boldsymbol{\varphi} t}^{0\prime}
\end{aligned}
$$

the information matrix $\mathfrak{I}(\boldsymbol{\theta}_0)$ is approximated by

$$
\begin{aligned}
\hat{\mathfrak{I}}_T\left(\boldsymbol{\theta}_0\right) &= T^{-1}\sum_{t=1}^{T} E\left[\frac{\partial l_t\left(\boldsymbol{\theta}_0\right)}{\partial \boldsymbol{\theta}} \frac{\partial l_t\left(\boldsymbol{\theta}_0\right)}{\partial \boldsymbol{\theta}'} \mid \mathcal{F}_{t-1}\right] \\
&= \begin{bmatrix} \boldsymbol{M}_1 & \boldsymbol{M}_2 & \boldsymbol{M}_4 \\ \boldsymbol{M}_2' & \boldsymbol{M}_3 & \boldsymbol{M}_5 \\ \boldsymbol{M}_4' & \boldsymbol{M}_5' & \boldsymbol{M}_6 \end{bmatrix}.
\end{aligned}
$$

The block of the inverse of $\hat{\mathfrak{I}}_T(\boldsymbol{\theta}_0)$ corresponding to the correlation and transition parameters is given by

$$
\left(\begin{bmatrix} \boldsymbol{M}_3 & \boldsymbol{M}_5 \\ \boldsymbol{M}_5' & \boldsymbol{M}_6 \end{bmatrix} - \begin{bmatrix} \boldsymbol{M}_2' \\ \boldsymbol{M}_4' \end{bmatrix} \boldsymbol{M}_1^{-1} \begin{bmatrix} \boldsymbol{M}_2 & \boldsymbol{M}_4 \end{bmatrix}\right)^{-1}
$$

from where the $\frac{N(N-1)}{2} \times \frac{N(N-1)}{2}$ block corresponding to $\boldsymbol{\rho}_{(3)}^*$ can be extracted. Replacing the true unknown parameter values with their maximum likelihood estimators, the test statistic simplifies to

$$
T^{-1}\left(\sum_{t=1}^{T} \frac{\partial l_t(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\rho}_{(3)}^{*\prime}}\right) \left[\hat{\mathfrak{I}}_T(\hat{\boldsymbol{\theta}})\right]^{-1}_{(\boldsymbol{\rho}_{(3)}^*, \boldsymbol{\rho}_{(3)}^*)} \left(\sum_{t=1}^{T} \frac{\partial l_t(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\rho}_{(3)}^*}\right) \tag{32}
$$

where $[\hat{\mathfrak{I}}_T(\boldsymbol{\theta}_0)]^{-1}_{(\boldsymbol{\rho}_{(3)}^*, \boldsymbol{\rho}_{(3)}^*)}$ is the block of the inverse of $\hat{\mathfrak{I}}_T$ corresponding to those correlation parameters that are set to zero under the null. (32) has an asymptotic $\chi^2_{\frac{N(N-1)}{2}}$ distribution when the null is true.

Test of the partially constant STCC–GARCH model against a partially constant DSTCC–GARCH model

When testing the hypothesis that some of the correlation parameters are constant according to the transition variable s_{2t} in both extreme states described by variable s_{1t}, the following modifications need to be done to the testing procedure of the previous subsection: Let there be k pairs of correlation restrictions in the alternative hypothesis of the form

$$
\rho_{(11)ij} = \rho_{(12)ij} \quad \text{and} \quad \rho_{(21)ij} = \rho_{(22)ij}, \quad i > j.
$$

In the linearized correlation matrix $\boldsymbol{P}_t^*$ the k elements in matrix $\boldsymbol{P}_{(3)}^*$ corresponding to these restrictions are set to zero, and the vector $\boldsymbol{\rho}^*$ is defined as before but excluding the elements that have been set to zero. Furthermore,

$$\frac{\partial(vec\boldsymbol{P}_t^*)'}{\partial\boldsymbol{\rho}^*}$$

is defined as before, but with the corresponding k rows deleted. The same rows are also deleted from $\boldsymbol{x}_{\rho^* t}$.

When restricting some of the correlation parameters constant according to the transition variable s_{1t} in both extreme states described by variable s_{2t} the test is as defined in the previous subsection with the following modifications: Let there be l pairs of correlations of the form

$$\rho_{(11)ij} = \rho_{(21)ij} \quad \text{and} \quad \rho_{(12)ij} = \rho_{(22)ij}, \quad i > j$$

in both null and alternative hypothesis. In the linearized correlation matrix $\boldsymbol{P}_t^*$ the l elements in matrix $\boldsymbol{P}_{(2)}^*$ corresponding to these restrictions are set to zero, and the vector $\boldsymbol{\rho}^*$ is defined as before but excluding the elements that have been set to zero. Furthermore, when forming

$$\frac{\partial(vec\boldsymbol{P}_t^*)'}{\partial\boldsymbol{\rho}^*}$$

the first $\frac{N(N-1)}{2}$ rows are multiplied by 1 instead of $1 - G_{1t}$, and from the next $\frac{N(N-1)}{2}$ rows, l rows corresponding to the restricted correlations are deleted. The same rows are also deleted from $\boldsymbol{x}_{\rho^* t}$.

These two specifications for partial constancy can be combined, and the test statistic is as defined in the previous subsection with the modifications described above. The asymptotic distribution needs to be adjusted for degrees of freedom to equal the number of restrictions that are tested.

References

Anderson, T. W. (2003): *An Introduction to Multivariate Statistical Analysis*. Wiley, New York, 3rd edn.

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

Bera, A. K., and S. Kim (2002): "Testing constancy of correlation and other specifications of the BGARCH model with an application to international equity returns," *Journal of Empirical Finance*, 9, 171–195.

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

——— (2005b): "Bond market and stock market integration in Europe," DNB Working Paper No. 60.

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.

Boyer, B. H., M. S. Gibson, and M. Loretan (1999): "Pitfalls in tests for changes in correlations," Board of Governors of the Federal Reserve System, International Finance Discussion Paper No. 597.

Cappiello, L., R. F. Engle, and K. Sheppard (2006): "Asymmetric dynamics in the correlations of global equity and bond returns," *Journal of Financial Econometrics*, 4, 537–572.

Chesnay, F., and E. Jondeau (2001): "Does correlation between stock returns really increase during turbulent periods?," *Economic Notes*, 30, 53–80.

de Santis, G., and B. Gerard (1997): "International asset pricing and portfolio diversification with time-varying risk," *Journal of Finance*, 52, 1881–1912.

Doornik, J. A. (2002): *Object-Oriented Matrix Programming Using Ox*. Timberlake Consultants Press, 3rd edn., see also www.doornik.com.

Engle, R. F. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, 50, 987–1006.

——— (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 R. Colacito (2006): "Testing and valuing dynamic correlations for asset allocation," *Journal of Business and Economic Statistics*, 24, 238–253.

Engle, R. F., and K. Sheppard (2001): "Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH," NBER Working Paper 8554.

Glosten, L. W., R. Jagannathan, and D. E. Runkle (1993): "On the relation between the expected value and the volatility of the nominal excess return on stocks," *Journal of Finance*, 48, 1779–1801.

Hafner, C. M., D. van Dijk, and P. H. Franses (2005): "Semi-parametric modelling of correlation dynamics," in *Advances in Econometrics*, ed. by T. Fomby, C. Hill, and D. Terrell, vol. 20/A, pp. 59–103. Amsterdam: Elsevier Sciences.

Lin, W.-L., R. F. Engle, and T. Ito (1994): "Do bulls and bears move across borders? International transmission of stock returns and volatility," *The Review of Financial Studies*, 7, 507–538.

Longin, F., and B. Solnik (2001): "Extreme correlation of international equity markets," *Journal of Finance*, 56, 649–676.

Lütkepohl, H. (1996): *Handbook of Matrices*. John Wiley & Sons, Chichester.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988): "Testing linearity against smooth transition autoregressive models," *Biometrika*, 75, 491–499.

Martens, M., and S.-H. Poon (2001): "Returns synchronization and daily correlation dynamics between international stock markets," *Journal of Banking and Finance*, 25, 1805–1827.

Meitz, M., and P. Saikkonen (2008a): "Ergodicity, mixing, and existence of moments of a class of Markov models with applications to GARCH and ACD models," *Econometric Theory*, 24, 1291–1320.

——— (2008b): "Stability of nonlinear AR-GARCH models," *Journal of Time Series Analysis*, 29, 453–475.

Pagan, A., and A. Ullah (1999): *Nonparametric Econometrics*. Cambridge University Press.

Pelletier, D. (2006): "Regime switching for dynamic correlations," *Journal of Econometrics*, 131, 445–473.

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

——— (2008): "Multivariate GARCH models," in *Handbook of Financial Time Series*, ed. by T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch. Springer, New York.

Tse, Y. K. (2000): "A test for constant correlations in a multivariate GARCH model," *Journal of Econometrics*, 98, 107–127.

Tse, Y. K., and K. C. Tsui (2002): "A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations," *Journal of Business and Economic Statistics*, 20, 351–362.

White, H. (1982): "Maximum likelihood estimation of misspecified models," *Econometrica*, 50, 1–25.