Consistency and asymptotic normality of maximum likelihood estimators of a multiplicative time-varying smooth transition correlation GARCH model

Annastiina Silvennoinen[*] and Timo Teräsvirta[†‡]

[*]NCER, Queensland University of Technology, Brisbane
[†]CREATES, Aarhus University
[‡]C.A.S.E., Humboldt-Universität zu Berlin

March 24, 2017

Abstract

A new multivariate volatility model that belongs to the family of conditional correlation GARCH models is introduced. The GARCH equations of this model contain a multiplicative deterministic component to describe long-run movements in volatility and, in addition, the correlations are deterministically time-varying. Parameters of the model are estimated jointly using maximum likelihood. Consistency and asymptotic normality of maximum likelihood estimators is proved. Numerical aspects of the estimation algorithm are discussed. A bivariate empirical example is provided.

1　Introduction

Forecasting volatility is important for investors and portfolio managers alike. Multivariate models of generalized autoregressive conditional heteroskedasticity (GARCH) provide a useful way of generating (short-run) forecasts for that purpose. Research on these models has been lively ever since Bollerslev, Engle and Wooldridge (1988) introduced the first multivariate model, and the number of applications has been steadily increasing over the years. Developments in that area have been summed up in surveys such as Bauwens, Laurent and Rombouts (2006) and Silvennoinen and Teräsvirta (2009b). It appears that the most popular class of multivariate GARCH models is the class of conditional correlation GARCH (CC-GARCH) models introduced by Bollerslev (1990). In the original model the conditional correlations were constant, hence the name Constant Conditional Correlation (CCC-GARCH) model, but this restriction has later been removed by parameterising the change in conditional correlations. The most frequently applied CC-GARCH model is the Dynamic CC (DCC-) GARCH model by Engle (2002). Tse and Tsui (2002) introduced a similar model which they called the Varying Correlation (VC-) GARCH model. Other CC-GARCH models worth mentioning include the Markov-Switching CC-GARCH model of Pelletier (2006), the Smooth Transition CC (STCC-) GARCH model, see Silvennoinen and Teräsvirta (2005, 2009a, 2015), the Dynamic Equicorrelation GARCH (DECO) model by Engle and Kelly (2012) and Multivariate GARCH with covariance breakdowns by Jin and Maheu (2016). Time-varying correlations may also be estimated nonparametrically as in Hafner, van Dijk and Franses (2006) or Long, Su and Ullah (2011).

Although many of these models are in widespread use, there is no asymptotic theory available for maximum likelihood (ML) estimators of the parametric models. While the GARCH equations can be estimated consistently and the ML estimators of their parameters have been shown to be normal under standard regularity conditions, corresponding results for the complete CC-GARCH models do not exist. Standard asymptotic inference is routinely applied in connection with dynamic CC-GARCH models but, as pointed out by Engle and Kelly (2012), 'a rigorous analysis of asymptotic theory for multivariate GARCH processes remains an important unanswered question'. The most general asymptotic results are limited to CCC-GARCH models and appeared in Ling and McAleer (2003). These authors considered a model that also contains a conditional mean of autoregressive moving-average (ARMA) type, and the conditional variance component is an Extended Conditional GARCH specification, see Jeantheau (1998), which allows spillover effects between the assets. This asymptotic theory makes it possible to test the constant conditional correlation assumption by Lagrange multiplier or score tests, see Tse (2000), Péguin-Feissolle and Sanhaji (2016) and Silvennoinen and Teräsvirta (2009a, 2015) for examples. Testing specification of the GARCH equations is possible as well, see for instance Bollerslev (1986), Li and Mak (1994), Lundbergh and Teräsvirta (2002), Nakatani and Teräsvirta (2009) and Pedersen (2017).

It is well known that standard (weakly) stationary GARCH models are not suitable for modelling long (daily) return series because of nonstationarity. This difficulty has been tackled by applying Integrated GARCH, see Engle and Bollerslev

(1986), and Fractionally Integrated GARCH (FIGARCH) models, Baillie, Bollerslev and Mikkelsen (1996), but few generalisations to multivariate GARCH exist in the literature. Conrad, Karanasos and Zeng (2011) defined the Multivariate Fractionally Integrated Asymmetric Power ARCH (FIAPARCH) model and applied bivariate and trivariate versions of it to various stock index return series. In that model, the conditional correlations are constant. Dark (2015) recently introduced a multivariate model whose GARCH equations are of FIGARCH type while the conditional correlations are dynamic as in Engle (2002) and asymmetric as well.

Another way of dealing with the aforementioned nonstationarity is to model it by extending the GARCH model by a multiplicative deterministic component. Early examples of this are Feng (2004) and van Bellegem and von Sachs (2004) who estimated the deterministic component nonparametrically. Engle and Rangel (2008) and Brownlees and Gallo (2010) used quadratic exponential splines and Mazur and Pipień (2012) applied the Fourier Flexible Form by Gallant (1981). Amado and Teräsvirta (2008, 2013, 2014b, 2017) used a linear combination of logistic functions of rescaled time, which resulted in the Multiplicative Time-Varying GARCH (MTV-GARCH) model. A multivariate GARCH model that contains a deterministic component was introduced in Hafner and Linton (2010), but their GARCH model was the so-called BEKK-GARCH by Engle and Kroner (1995). Recently, Amado and Teräsvirta (2014a) generalised the MTV-GARCH model to the vector case. In that MTV-CC-GARCH model, the conditional correlations were either constant as in Bollerslev (1990), or dynamic as in Engle (2002) or Tse and Tsui (2002). Needless to say, no asymptotic inference is available for that model either, although under regularity conditions ML estimators of parameters in the MTV-GARCH equations are consistent and asymptotically normal; see Amado and Teräsvirta (2013).

In this work, the multivariate MTV-CCC-GARCH model is generalised by defining the time-varying correlations as in STCC-GARCH but assuming that the transition variable is rescaled time. The main contribution of this paper is that ML estimators of the parameters of this MTV-TVC-GARCH model, TVC model for short, are shown to be consistent and asymptotically normal. This may be the first time when such a result is available for multivariate CC-GARCH models with time-varying correlations. Maximum likelihood estimates of parameters are fully efficient. Unlike the DCC- or VC-GARCH model that rely on GARCH-type conditional correlations and thus, at least in principle, describe clustering in conditional correlations, the TVC model is designed to characterise long-run movements in correlations. Estimation of parameters of the highly nonlinear TVC model is numerically demanding and requires a good optimisation algorithm.

The plan of the paper is as follows. The TVC model is presented in Section 2 and its log-likelihood in Section 3. Consistency of ML estimators when the conditional variance component is assumed constant is considered in Section 4. The information matrix for this restricted model is defined in Section 5 and asymptotic normality of the ML estimators studied in Section 6. After defining the relevant Hessian matrix in Section 7, ML estimation of the complete model including the GARCH component is discussed in Section 8 where asymptotic normality of the ML estimators is established. The estimation algorithm is described and some numerical aspects of estimation are highlighted in Section 9. Section 10 contains an empirical example.

Conclusions can be found in Section 11.

2 The model

The model considered in this work is the multivariate multiplicative time-varying Smooth Transition Conditional Correlation GARCH model with deterministically time-varying correlations, that is, the TVC model. The observable stochastic $N \times 1$ vector $\boldsymbol{\varepsilon}_t$ is decomposed in the customary fashion as

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

where $\boldsymbol{z}_t$ is a vector of independent random variables with $\mathsf{E}\boldsymbol{z}_t = \boldsymbol{0}$ and a positive definite deterministically varying covariance matrix $\mathrm{cov}(\boldsymbol{z}_t) = \boldsymbol{P}_t$. It follows that $\boldsymbol{P}_t^{-1/2} \boldsymbol{z}_t = \boldsymbol{\zeta}_t \sim \mathrm{iid}(\boldsymbol{0}, \boldsymbol{I}_N)$. Time may be replaced by a stationary random variable that is independent of $\boldsymbol{z}_t$. The diagonal $N \times N$ matrix $\boldsymbol{H}_t^{1/2} = \boldsymbol{S}_t \boldsymbol{D}_t$, where the deterministic matrix $\boldsymbol{S}_t = \mathrm{diag}(g_{1t}^{1/2}, ..., g_{Nt}^{1/2})$ has positive diagonal elements for all t, and $\boldsymbol{D}_t = \mathrm{diag}(h_{1t}^{1/2}, ..., h_{Nt}^{1/2})$ contains the conditional standard deviations of the elements of $\boldsymbol{S}_t^{-1}\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}/g_{1t}^{1/2}, ..., \varepsilon_{Nt}/g_{Nt}^{1/2})'$. Following Amado and Teräsvirta (2008, 2013, 2014a,b, 2017), the diagonal elements of $\boldsymbol{S}_t^2$ are defined as follows:

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

$i = 1, ..., N$, where the (generalised) logistic function

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

and $\boldsymbol{c}_{ij} = (c_{ij1}, ..., c_{ijK_{ij}})'$ such that $c_{ij1} \leq ... \leq c_{ijK_{ij}}$. Both $\gamma_{ij} > 0$ and $c_{ij1} \leq ... \leq c_{ijK_{ij}}$ are identification restrictions. The intercept in (2) is set to one for notational convenience. Although any known positive constant will do, in practice some choices are numerically better motivated than others. This will be further discussed in Section 9.

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

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

where $\boldsymbol{\phi}_t$ is assumed to have a standard stationary GARCH representation with the conditional covariance matrix $\mathsf{E}\{\boldsymbol{\phi}_t \boldsymbol{\phi}_t' | \mathcal{F}_{t-1}\} = \boldsymbol{D}_t \boldsymbol{P}_t \boldsymbol{D}_t$. It follows that the ith first-order GJR-GARCH(1,1,1) equation, see Glosten, Jagannathan and Runkle (1993), is

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

where $I(A)$ is an indicator function: $I(A) = 1$ when A occurs, zero otherwise. (Note that ε_t and ϕ_t have the same sign.) The standard GARCH(1,1) model is obtained

by setting $\kappa_{i1} = 0$. A higher-order structure is possible, although there do not seem to exist applications of the GJR-GARCH model of order greater than one.

In this work, $\boldsymbol{P}_t$ is defined as in smooth transition conditional correlation model of Silvennoinen and Teräsvirta (2005, 2015):

$$\boldsymbol{P}_t = G_t(t/T, \gamma, c)\boldsymbol{P}_{(1)} + \{1 - G_t(t/T, \gamma, c)\}\boldsymbol{P}_{(2)} \tag{5}$$

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

$$G_t(t/T, \gamma, c) = (1 + \exp\{-\gamma \prod_{k=1}^{K}(t/T - c_k)\})^{-1}, \ \gamma > 0 \tag{6}$$

and $c_1 \leq ... \leq c_K$. As a convex combination of $\boldsymbol{P}_{(1)}$ and $\boldsymbol{P}_{(2)}$, $\boldsymbol{P}_t$ is positive definite. The resulting model is a Multiplicative Time-Varying Smooth Transition Conditional Correlation GARCH (TVC) model. A bivariate STCC-GARCH model with time as the transition variable was first considered by Berben and Jansen (2005). A similar MTV-Conditional Correlation GARCH model but with different definitions for $\boldsymbol{P}_t$ was discussed in Amado and Teräsvirta (2014a).

3 The log-likelihood

3.1 Notation

Let $\boldsymbol{\theta}_g = (\boldsymbol{\theta}'_{g1}, ..., \boldsymbol{\theta}'_{gN})'$ be the parameter vector containing the parameters of $\boldsymbol{S}_t$. If $K_{ij} = 1$ in (3) for simplicity, then $\boldsymbol{\theta}_{gi} = (\boldsymbol{\delta}'_i, \boldsymbol{\gamma}'_i, \boldsymbol{c}'_i)'$, with $\boldsymbol{\delta}_i = (\delta_{i0}, \delta_{i1}, ..., \delta_{ir_i})'$, $\boldsymbol{\gamma}_i = (\gamma_{i1}, ..., \gamma_{ir_i})'$, and $\boldsymbol{c}_i = (c_{i1}, ..., c_{ir_i})'$, $i = 1, ..., N$. The subscript i refers to the ith diagonal element of $\boldsymbol{S}_t$. Analogously, let $\boldsymbol{\theta}_h = (\boldsymbol{\theta}'_{h1}, ..., \boldsymbol{\theta}'_{hN})'$ be the parameter vector containing the parameters of $\boldsymbol{D}_t$, where $\boldsymbol{\theta}_{hi}$ is the parameter vector of $h_{it}^{1/2}$, the ith diagonal element of $\boldsymbol{D}_t$. For example, if the model is a GJR-GARCH(1,1,1) model, then $\boldsymbol{\theta}_{hi} = (\alpha_{i0}, \alpha_{i1}, \kappa_{i1}, \beta_{i1})'$. The vector $\boldsymbol{\theta}_\rho = (\boldsymbol{\rho}'_1, \boldsymbol{\rho}'_2, \gamma, \boldsymbol{c}')'$, where $\boldsymbol{c} = (c_1, ..., c_K)'$ and $\boldsymbol{\rho}_i = \mathrm{vecl}(\boldsymbol{P}_{(i)})$, $i = 1, 2$, contains the parameters in $\boldsymbol{P}_t$. The whole parameter vector of the model is denoted as $\boldsymbol{\theta} = (\boldsymbol{\theta}'_g, \boldsymbol{\theta}'_h, \boldsymbol{\theta}'_\rho)'$ and, whenever needed, $\boldsymbol{\theta}^0 = ((\boldsymbol{\theta}^0_g)', (\boldsymbol{\theta}^0_h)', (\boldsymbol{\theta}^0_\rho)')'$ is the true parameter vector. The following notation is used: $\boldsymbol{S}_t = \boldsymbol{S}_t(\boldsymbol{\theta}_g)$, $\boldsymbol{D}_t = \boldsymbol{D}_t(\boldsymbol{\theta}_h, \boldsymbol{\theta}_g)$ and $\boldsymbol{P}_t = \boldsymbol{P}_t(\boldsymbol{\theta}_\rho)$. Also, $\boldsymbol{P}^0_t = \boldsymbol{P}_t(\boldsymbol{\theta}^0_\rho)$ is the true correlation matrix. That $\boldsymbol{D}_t$ is a function of both $\boldsymbol{\theta}_h$ and $\boldsymbol{\theta}_g$ is a consequence of the fact that

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

Furthermore, let $\boldsymbol{0}_k$ be a k-dimensional null vector, $\boldsymbol{0}_{k \times m}$ a $k \times m$ null matrix, $\boldsymbol{I}_k$ a $k \times k$ identity matrix, and $\boldsymbol{1}_k$ a k-dimensional vector of ones. Finally, the symbol $\overset{p}{\to}$ signifies convergence in probability, whereas $\overset{d}{\to}$ represents convergence in distribution.

3.2 Log-likelihood and score

Assuming normal and independent errors, the log-likelihood function of the model (the log conditional density of $\boldsymbol{\zeta}_t$) for observation t (ignoring the constant) becomes

$$
\begin{aligned}
&\ln f(\boldsymbol{\zeta}_t | \boldsymbol{\theta}, \mathcal{F}_{t-1}) \\
&= -(1/2) \ln |\boldsymbol{S}_t \boldsymbol{D}_t \boldsymbol{P}_t \boldsymbol{D}_t \boldsymbol{S}_t| - (1/2)\boldsymbol{\varepsilon}_t'\{\boldsymbol{S}_t \boldsymbol{D}_t \boldsymbol{P}_t \boldsymbol{D}_t \boldsymbol{S}_t\}^{-1}\boldsymbol{\varepsilon}_t \\
&= -\ln|\boldsymbol{S}_t| - \ln|\boldsymbol{D}_t| - (1/2)\ln|\boldsymbol{P}_t| - (1/2)\boldsymbol{\varepsilon}_t'\{\boldsymbol{S}_t \boldsymbol{D}_t \boldsymbol{P}_t \boldsymbol{D}_t \boldsymbol{S}_t\}^{-1}\boldsymbol{\varepsilon}_t \quad (7)
\end{aligned}
$$

where $\mathcal{F}_{t-1}$ contains the conditioning information. Since $\boldsymbol{S}_t$ and $\boldsymbol{D}_t$ are diagonal, one may write

$$
\begin{aligned}
\ln f(\boldsymbol{\zeta}_t | \boldsymbol{\theta}, \mathcal{F}_{t-1}) =\ & -(1/2)\sum_{i=1}^{N} \ln g_i(\boldsymbol{\theta}_{gi}; t/T) - (1/2)\sum_{i=1}^{N} \ln h_{it}(\boldsymbol{\theta}_{hi}, \boldsymbol{\theta}_{gi}) \\
& -(1/2)\ln|\boldsymbol{P}_t| - (1/2)\boldsymbol{\varepsilon}_t'\{\boldsymbol{S}_t \boldsymbol{D}_t \boldsymbol{P}_t \boldsymbol{D}_t \boldsymbol{S}_t\}^{-1}\boldsymbol{\varepsilon}_t. \quad (8)
\end{aligned}
$$

In order to consider asymptotic properties of the maximum likelihood estimator of $\boldsymbol{\theta}$, the focus will first be on the submodel in which $\boldsymbol{D}_t = \boldsymbol{I}_N$. Apart from the error term $\boldsymbol{\zeta}_t$, this is a completely deterministic model. Let $\boldsymbol{\theta}_1 = (\boldsymbol{\theta}_g', \boldsymbol{\theta}_\rho')'$ and $\boldsymbol{\theta}_1^0 = ((\boldsymbol{\theta}_g^0)', (\boldsymbol{\theta}_\rho^0)')'$. Assuming, for notational simplicity, that $\boldsymbol{c} = c$, the score of the log-likelihood function for observation t equals

$$
\frac{\partial}{\partial \boldsymbol{\theta}_1} \ln f(\boldsymbol{\zeta}_t | \boldsymbol{\theta}_1) = \boldsymbol{s}_t(\boldsymbol{\theta}_1) = (\boldsymbol{s}_t'(\boldsymbol{\theta}_g), \boldsymbol{s}_t'(\boldsymbol{\theta}_\rho))' \quad (9)
$$

where

$$
\begin{aligned}
\boldsymbol{s}_t(\boldsymbol{\theta}_g) &= (\boldsymbol{s}_t'(\boldsymbol{\theta}_{g1}), ..., \boldsymbol{s}_t'(\boldsymbol{\theta}_{gN}))' \quad &(10) \\
\boldsymbol{s}_t(\boldsymbol{\theta}_\rho) &= (\boldsymbol{s}_t'(\boldsymbol{\rho}_1), \boldsymbol{s}_t'(\boldsymbol{\rho}_2), s_t(\gamma), s_t(c))'. \quad &(11)
\end{aligned}
$$

The form of the sub-blocks in (10) and (11) is given in the following lemma:

Lemma 1 *The blocks (10) and (11) of (9) for observation t when $\boldsymbol{D}_t = \boldsymbol{I}_N$ in (8), have the following representations:*

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

where $\boldsymbol{e}_i = (\boldsymbol{0}_{i-1}', 1, \boldsymbol{0}_{N-i}')'$, $i = 1, ..., N$, and

$$
\boldsymbol{s}_t(\boldsymbol{\theta}_\rho) = -(1/2)\frac{\partial vec(\boldsymbol{P}_t)'}{\partial \boldsymbol{\theta}_\rho}\{vec(\boldsymbol{P}_t^{-1}) - (\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1})vec(\boldsymbol{z}_t \boldsymbol{z}_t')\}.
$$

Dividing $\frac{\partial vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\theta}_\rho'}$ into four sub-blocks as follows:

$$
\frac{\partial vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\theta}_\rho'} = \left[\frac{\partial vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\rho}_1'}, \frac{\partial vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\rho}_2'}, \frac{\partial vec(\boldsymbol{P}_t)}{\partial \gamma}, \frac{\partial vec(\boldsymbol{P}_t)}{\partial c}\right]
$$

the two $N^2 \times N(N-1)/2$ blocks are

$$\frac{\partial vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\rho}_1'} = (1 - G_t)\frac{\partial vec(\boldsymbol{P}_{(1)})}{\partial \boldsymbol{\rho}_1'} \tag{12}$$

$$\frac{\partial vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\rho}_2'} = G_t\frac{\partial vec(\boldsymbol{P}_{(2)})}{\partial \boldsymbol{\rho}_2'} \tag{13}$$

and two $N^2 \times 1$ vectors

$$\frac{\partial vec(\boldsymbol{P}_t)}{\partial \gamma} = G_t(1 - G_t)(t/T - c)vec(\boldsymbol{P}_{(2)} - \boldsymbol{P}_{(1)})$$

and

$$\frac{\partial vec(\boldsymbol{P}_t)}{\partial c} = -\gamma G_t(1 - G_t)vec(\boldsymbol{P}_{(2)} - \boldsymbol{P}_{(1)}).$$

Proof: See Appendix 12. Formulas (12) and (13) can be combined into an $N^2 \times N(N-1)$ matrix:

$$\frac{\partial \mathrm{vec}(\boldsymbol{P}_t)}{\partial \boldsymbol{\rho}'} = (1 - G_t)\frac{\partial \mathrm{vec}(\boldsymbol{P}_{(1)})}{\partial \boldsymbol{\rho}'} + G_t\frac{\partial \mathrm{vec}(\boldsymbol{P}_{(2)})}{\partial \boldsymbol{\rho}'}.$$

4   Consistency

Direct maximum likelihood (ML) estimation of $\boldsymbol{\theta}$ from (8) is numerically very difficult. To alleviate the computational burden it is preferable to divide the maximisation problem into parts instead of maximising the whole log-likelihood at once. This is useful from the numerical point of view, but maximisation by parts has an even more important role to play. Following Song, Fan and Kalbfleisch (2005) it allows us to derive asymptotic properties for the ML estimator of $\boldsymbol{\theta}$, when the two parts of each iteration consist of first estimating $\boldsymbol{\theta}_1$ and then estimating the GARCH parameter vector $\boldsymbol{\theta}_h$ conditionally on the obtained estimate of $\boldsymbol{\theta}_1$. For reasons that will become clear in Section 8, we begin by considering both consistency and asymptotic normality of the ML estimator $\widehat{\boldsymbol{\theta}}_1$ of $\boldsymbol{\theta}_1$ in this framework. Our results may be viewed as generalisations of the ones in Amado and Teräsvirta (2013) who considered ML estimation of a single multiplicative TV-GARCH equation.

As already mentioned, the focus will be on the restricted model with $\boldsymbol{D}_t = \boldsymbol{I}_N$. Under this restriction, the intercept in (2) is a free parameter, which leads to redefining g_{it} as

$$g_{it}(\boldsymbol{\delta}_i, \boldsymbol{\gamma}_i, \boldsymbol{c}_i; t/T) = \delta_{i0} + \sum_{j=1}^{r_i} \delta_{ij}G_{ij}(\gamma_{ij}, c_{ij}; t/T), \ i = 1, ..., N. \tag{14}$$

We make the following assumptions:

AG1. The parameter space $\Theta_1 = \{\Delta \times \Gamma \times C \times P\}$ is compact, where $\boldsymbol{\delta} \in \Delta$, $\boldsymbol{\gamma} \in \Gamma$, $\boldsymbol{c} \in C$, and $\boldsymbol{\rho} \in P$.

AG2. The log-likelihood has a unique maximum at the true parameter $\boldsymbol{\theta}_1^0$ which is an interior point of Θ_1.

AG3. The elements of $\boldsymbol{\delta} \in \Delta$ are restricted such that $\max_{j=1,...,r_i;i=0,1,...,N} |\delta_{ij}|$ $\leq M_\delta < \infty$ and $\inf_{\boldsymbol{\delta}_i \in \Delta} g_{it}(t/T, \boldsymbol{\delta}_i, \boldsymbol{\gamma}_i, \boldsymbol{c}_i) > 0$ for each $i = 1,...,N$. In particular, $\sum_{k=0}^{j} \delta_{ik} > 0$, $j = 0, 1, ..., r_i$, and, for identification reasons, $\delta_{ij} \neq 0$, for $j = 1, ..., r_i$ and $i = 1, ..., N$. For any pair (i, m), $i \neq m$, there exists at least one time-point t such that $g_{it} \neq g_{mt}$. Furthermore, the positive intercepts δ_{i0}, $i = 1, ..., N$, are assumed known, so that the corresponding $\alpha_{i0} > 0$, $i = 1, ..., N$, in (4) are free parameters.

AG4. The slope parameters $\gamma_{ij} > 0$ and $c_{i1} < ... < c_{ir_i}$, $j = 1, ..., r_i$, $i = 1, ..., N$.

AG5. The matrices $\boldsymbol{P}_{(1)}$ and $\boldsymbol{P}_{(2)}$ in (5) are positive definite correlation matrices for $\boldsymbol{\rho} \in P$, $\boldsymbol{P}_{(1)} \neq \boldsymbol{P}_{(2)}$. Furthermore, $\gamma > 0$ and $c_1 \leq ... \leq c_k$ in (6).

AG6. $\mathsf{E}\zeta_{it}^{2(2+\phi)} < \infty$ for some $\phi > 0$ and $i = 1, ..., N$.

The assumption AG4 is necessary for AG2 but is stated separately. It follows from AG5 that $\boldsymbol{P}_t$ is positive definite because it is a convex combination of $\boldsymbol{P}_{(1)}$ and $\boldsymbol{P}_{(2)}$. The inequalities $\gamma > 0$ and $c_1 \leq ... \leq c_k$ are required for uniqueness of (6). If $c_j = c_{j+1}$, then the product $(t/T - c_i)(t/T - c_{i+1}) = (t/T - c_i)^2$ in (6), i.e., one parameter less is estimated. AG6 is required in proving asymptotic normality in the case where the errors are not normal. The consistency result is formulated as follows:

Theorem 1 *Consider the multiplicative time-varying smooth transition conditional correlation GARCH model (1), (3), (5), (6) and (14), where $\boldsymbol{D}_t = \boldsymbol{I}_N$. Let $\widehat{\boldsymbol{\theta}}_1$ be the maximum likelihood estimator of $\boldsymbol{\theta}_1^0 = ((\boldsymbol{\theta}_g^0)', (\boldsymbol{\theta}_\rho^0)')'$. Suppose that the assumptions AG1-AG5 hold and that $\boldsymbol{\zeta}_t \sim iid(\boldsymbol{0}_N, \boldsymbol{I}_N)$. Then $\widehat{\boldsymbol{\theta}}_1$ is consistent for $\boldsymbol{\theta}_1^0$.*

Proof: See Appendix 12.

It should be noted, however, that when $\boldsymbol{D}_t = \boldsymbol{I}_N$, the term 'conditional correlation' is misleading. The time-varying correlations are unconditional because the transition variable controlling the change is time both in the GARCH (sic!) equations and correlations.

5 Information matrix and second partial derivatives of the likelihood

5.1 Information matrix

Let $\boldsymbol{S}_t^0$ be the true diagonal matrix, that is, $\boldsymbol{S}_t^0 = \boldsymbol{S}_t(\boldsymbol{\theta}_g^0)$, where $\boldsymbol{\theta}_g^0$ is the true parameter vector. Let $\boldsymbol{P}_t^0$ be the true covariance matrix, so $\boldsymbol{\varepsilon}_t$ is decomposed as $\boldsymbol{\varepsilon}_t = \boldsymbol{S}_t^0 (\boldsymbol{P}_t^0)^{1/2} \boldsymbol{\zeta}_t$, where $\boldsymbol{\zeta}_t \sim iid(\boldsymbol{0}_N, \boldsymbol{I}_N)$. The information matrix of the log-likelihood

$$L_T = (1/T) \sum_{t=1}^{T} \ln f(\boldsymbol{\zeta}_t | \boldsymbol{\theta}_1) \tag{15}$$

is given by the following result:

Lemma 2 *The information matrix of the log-likelihood (15) equals*

$$\boldsymbol{B}(\boldsymbol{\theta}_1^0) = \int_0^1 \mathsf{E}\boldsymbol{s}_r(\boldsymbol{\theta}_1^0)\boldsymbol{s}_r'(\boldsymbol{\theta}_1^0)\,dr$$

where the element of the score $s_r(\boldsymbol{\theta}_1^0)$ at $r = t/T$ is defined in Lemma 1.

Proof: See Appendix 12.

To illustrate the structure of the information matrix, we present the expressions of the expectations of its four blocks at time t. For this purpose, define the (N, N) commutation matrix $\boldsymbol{K}$: the $N \times N$ (i, j) block of $\boldsymbol{K}$ equals $\boldsymbol{e}_j\boldsymbol{e}_i'$, $i, j = 1, ..., N$, see for example Lütkepohl (1996, pp. 115–118). Furthermore, to simplify the notation, set $g_{it}^0 = g_{it}|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_1^0}$ and $\partial g_{it}^0/\partial\boldsymbol{\theta}_1 = \partial g_{it}/\partial\boldsymbol{\theta}_1|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_1^0}$, $i = 1, ..., N$. The expectations can be found in the following lemma:

Lemma 3 *The expectations of the four blocks of the information matrix at (rescaled) time t/T are*

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

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

$$\mathsf{E}\boldsymbol{s}_t(\boldsymbol{\theta}_{gi}^0)\boldsymbol{s}_t'(\boldsymbol{\theta}_{gj}^0) = \frac{1}{4g_{it}^0 g_{jt}^0}\frac{\partial g_{it}^0}{\partial\boldsymbol{\theta}_{gi}}\frac{\partial g_{jt}^0}{\partial\boldsymbol{\theta}_{gj}'}(\boldsymbol{e}_i'\boldsymbol{P}_t^0\boldsymbol{e}_j)(\boldsymbol{e}_i'(\boldsymbol{P}_t^0)^{-1}\boldsymbol{e}_j)$$

and for $i = j$, where

$$\mathsf{E}\boldsymbol{s}_t(\boldsymbol{\theta}_{gi}^0)\boldsymbol{s}_t'(\boldsymbol{\theta}_{gi}^0) = \frac{1}{4(g_{it}^0)^2}\frac{\partial g_{it}^0}{\partial\boldsymbol{\theta}_{gi}}\frac{\partial g_{it}^0}{\partial\boldsymbol{\theta}_{gi}'}(1 + \boldsymbol{e}_i'(\boldsymbol{P}_t^0)^{-1}\boldsymbol{e}_i)$$

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

$$\begin{aligned} \mathsf{E}\boldsymbol{s}_t(\boldsymbol{\theta}_{gi}^0)\boldsymbol{s}_t'(\boldsymbol{\theta}_\rho^0) &= \frac{1}{4}\frac{\partial g_{it}^0}{\partial\boldsymbol{\theta}_{gi}}\{(\boldsymbol{e}_i' \otimes \boldsymbol{e}_i')((\boldsymbol{P}_t^0)^{-1} \otimes \boldsymbol{I}_N) \\ &\quad + (\boldsymbol{e}_i' \otimes \boldsymbol{e}_i')(\boldsymbol{I}_N \otimes (\boldsymbol{P}_t^0)^{-1}\}\frac{\partial vec(\boldsymbol{P}_t^0)}{\partial\boldsymbol{\theta}_\rho'} \end{aligned}$$

$i = 1, ..., N$, and

$$\begin{aligned} \mathsf{E}\boldsymbol{s}_t(\boldsymbol{\theta}_\rho)\boldsymbol{s}_t'(\boldsymbol{\theta}_\rho) &= \frac{1}{4}\frac{\partial vec(\boldsymbol{P}_t^0)'}{\partial\boldsymbol{\theta}_\rho}\{(\boldsymbol{P}_t^0)^{-1} \otimes (\boldsymbol{P}_t^0)^{-1} \\ &\quad + ((\boldsymbol{P}_t^0)^{-1} \otimes \boldsymbol{I}_N)\boldsymbol{K}(\boldsymbol{I}_N \otimes (\boldsymbol{P}_t^0)^{-1})\}\frac{\partial vec(\boldsymbol{P}_t^0)}{\partial\boldsymbol{\theta}_\rho}. \end{aligned}$$

Proof: See Appendix 12.

5.2 Second partial derivatives of the log-likelihood

Determining the Hessian matrix of the log-likelihood requires the second derivatives of (7) with respect to $\boldsymbol{\theta}_g$ and $\boldsymbol{\theta}_\rho$. They are defined by the following four lemmas:

Lemma 4 *Consider the log-likelihood (7) with $\boldsymbol{D}_t = \boldsymbol{I}_N$. Its second partial derivatives with respect to $\boldsymbol{\theta}_{gi}$ and $\boldsymbol{\theta}_{gj}$ are as follows:*

$$\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gj}\partial \boldsymbol{\theta}'_{gi}} = -\frac{1}{4g_{jt}g_{it}}\frac{\partial g_{jt}}{\partial \boldsymbol{\theta}_{gj}}\frac{\partial g_{it}}{\partial \boldsymbol{\theta}'_{gi}}\boldsymbol{e}'_j\boldsymbol{P}_t^{-1}\boldsymbol{e}_i\boldsymbol{e}'_i\boldsymbol{z}_t\boldsymbol{z}'_t\boldsymbol{e}_j$$

for $i, j = 1, ..., N$, $i \neq j$, and

$$\begin{aligned}
\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi}\partial \boldsymbol{\theta}'_{gi}} &= -\frac{1}{2g_{it}}\{\frac{1}{g_{it}}\frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}}\frac{\partial g_{it}}{\partial \boldsymbol{\theta}'_{gi}} - \frac{\partial^2 g_{it}}{\partial \boldsymbol{\theta}_{gi}\partial \boldsymbol{\theta}'_{gi}}\}(\boldsymbol{e}'_i\boldsymbol{P}_t^{-1}\boldsymbol{z}_t\boldsymbol{z}'_t\boldsymbol{e}_i - 1) \\
&\quad -\frac{1}{4g_{it}^2}\frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}}\frac{\partial g_{it}}{\partial \boldsymbol{\theta}'_{gi}}\boldsymbol{e}'_i\boldsymbol{P}_t^{-1}(\boldsymbol{I}_N + \boldsymbol{e}_i\boldsymbol{e}'_i)\boldsymbol{z}_t\boldsymbol{z}'_t\boldsymbol{e}_i
\end{aligned} \qquad (16)$$

for $i = j$, $i = 1, ..., N$. In (16),

$$\frac{\partial^2 g_{it}}{\partial \boldsymbol{\theta}_{gi}\partial \boldsymbol{\theta}'_{gi}} = \begin{bmatrix} 0 & \boldsymbol{0}'_{r_i} & \boldsymbol{0}'_{2r_i} \\ \boldsymbol{0}_{3r_i} & \boldsymbol{0}_{3r_i \times r_i} & \boldsymbol{D}_{22i} \end{bmatrix}$$

where the $3r_i \times 2r_i$ matrix $\boldsymbol{D}_{22i}$ equals

$$\boldsymbol{D}_{22i} = \begin{bmatrix} diag(g_{i\delta\gamma 1t}, ..., g_{i\delta\gamma r_i t}) & diag(g_{i\delta c 1t}, ..., g_{i\delta c r_i t}) \\ diag(g_{i\gamma\gamma 1t}, ..., g_{i\gamma\gamma r_i t}) & diag(g_{i\gamma c 1t}, ..., g_{i\gamma c r_i t}) \\ diag(g_{ic\gamma 1t}, ..., g_{i\delta\gamma r_i t}) & diag(g_{icc 1t}, ..., g_{icc r_i t}) \end{bmatrix} \qquad (17)$$

with

$$\begin{aligned}
g_{i\delta\gamma jt} &= \delta_{ij}G_{ijt}(1 - G_{ijt})(t/T - c_{ij}) \\
g_{i\delta c jt} &= -\gamma_{ij}\delta_{ij}G_{ijt}(1 - G_{ijt}) \\
g_{i\gamma\gamma jt} &= \delta_{ij}G_{ijt}(1 - G_{ijt})(1 - 2G_{ijt})(t/T - c_{ij})^2 \\
g_{icc jt} &= \gamma^2\delta_{ij}G_{ijt}(1 - G_{ijt})(1 - 2G_{ijt})
\end{aligned}$$

and

$$g_{i\gamma c jt} = -\gamma_j\delta_{ij}G_{ijt}(1 - G_{ijt})(1 - 2G_{ijt})(t/T - c_{ij})$$

$j = 1, ..., r_i$.

Proof: See Appendix 12.

Lemma 5 *Setting $\boldsymbol{\theta}_\rho = (\boldsymbol{\rho}'_1, \boldsymbol{\rho}'_2, \gamma, c)'$, the cross derivatives of (7) with $\boldsymbol{D}_t = \boldsymbol{I}_N$ with respect to $\boldsymbol{\theta}_{gi}$ and $\boldsymbol{\theta}_\rho$ are*

$$\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi}\partial \boldsymbol{\theta}_\rho} = -\frac{1}{2g_{it}}\frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}}(\boldsymbol{e}_i \otimes \boldsymbol{e}_i)'(\boldsymbol{z}_t\boldsymbol{z}'_t\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1})\frac{\partial vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\theta}_\rho}. \qquad (18)$$

Furthermore, denoting $\boldsymbol{\rho} = (\boldsymbol{\rho}_1', \boldsymbol{\rho}_2')'$, the second partial derivatives of (7) with $\boldsymbol{D}_t = \boldsymbol{I}_N$ with respect to $\boldsymbol{\rho} = (\boldsymbol{\rho}_1', \boldsymbol{\rho}_2')'$ are

$$
\begin{aligned}
\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t | \boldsymbol{\theta}_1)}{\partial \boldsymbol{\rho}' \partial \boldsymbol{\rho}} = & -\frac{1}{2}\Big\{ \frac{\partial vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\rho}'} \{ \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} \\
& - (\boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} + \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1}) \} \frac{\partial vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\rho}} \Big\}. \quad (19)
\end{aligned}
$$

Proof: See Appendix 12.

The corresponding derivatives with respect to γ and c, the parameters in the transition function of the correlation matrix, have a slightly different structure and are formulated in the next lemma.

Lemma 6 *The second derivatives of (7) with $\boldsymbol{D}_t = \boldsymbol{I}_N$ with respect to the parameters γ and c in the transition function of the correlation matrix $\boldsymbol{P}_t$ are as follows:*

$$
\begin{aligned}
\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t | \boldsymbol{\theta}_1)}{\partial x \partial y} = & -\frac{1}{2} \frac{\partial^2 vec(\boldsymbol{P}_t)'}{\partial x \partial y} \{ vec(\boldsymbol{P}_t^{-1}) - vec(\boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1}) \} \\
& -\frac{1}{2} \frac{\partial vec(\boldsymbol{P}_t)'}{\partial x} \{ \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} \\
& - (\boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} + \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1}) \} \frac{\partial vec(\boldsymbol{P}_t)}{\partial y} \quad (20)
\end{aligned}
$$

for $x, y = \gamma, c$. In (20),

$$
\begin{aligned}
\frac{\partial^2 vec(\boldsymbol{P}_t)}{\partial \gamma^2} & = \frac{\partial^2}{\partial \gamma^2} \{ (1 - G_t) vec(\boldsymbol{P}_{(1)}) + G_t vec(\boldsymbol{P}_{(2)}) \} \\
& = G_t (1 - G_t)(1 - 2G_t)(t/T - c)^2 vec(\boldsymbol{P}_{(2)} - \boldsymbol{P}_{(1)}) \quad (21)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 vec(\boldsymbol{P}_t)}{\partial c^2} & = \frac{\partial^2}{\partial c^2} \{ (1 - G_t) vec(\boldsymbol{P}_{(1)}) + G_t vec(\boldsymbol{P}_{(2)}) \} \\
& = \gamma^2 G_t (1 - G_t)(1 - 2G_t) vec(\boldsymbol{P}_{(2)} - \boldsymbol{P}_{(1)}) \quad (22)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial^2 vec(\boldsymbol{P}_t)}{\partial \gamma \partial c} & = \frac{\partial^2}{\partial \gamma \partial c} \{ (1 - G_t) vec(\boldsymbol{P}_{(1)}) + G_t vec(\boldsymbol{P}_{(2)}) \} \\
& = -\gamma G_t (1 - G_t)(1 - 2G_t)(t/T - c) vec(\boldsymbol{P}_{(2)} - \boldsymbol{P}_{(1)}). \quad (23)
\end{aligned}
$$

Proof: See Appendix 12.

Ultimately, we present the second partial derivatives of (7) with $\boldsymbol{D}_t = \boldsymbol{I}_N$ with respect to $\boldsymbol{\rho}$ and $\boldsymbol{\gamma} = (\gamma, c)'$.

Lemma 7 *The second partial derivatives of (7) with $\boldsymbol{D}_t = \boldsymbol{I}_N$ with respect to $\boldsymbol{\rho}$ and $\boldsymbol{\gamma} = (\gamma, c)'$ are as follows:*

$$
\begin{aligned}
\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t | \boldsymbol{\theta}_1)}{\partial \boldsymbol{\rho}' \partial \boldsymbol{\gamma}} = & -\frac{1}{2} \frac{\partial^2 vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\rho}'} \{ vec(\boldsymbol{P}_t^{-1}) - vec(\boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1}) \} \\
& -\frac{1}{2} \frac{\partial vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\rho}'} \{ \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} \\
& - (\boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} + \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1}) \} \frac{\partial vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\gamma}}
\end{aligned}
$$

10

with

$$\frac{\partial^2 vec(\boldsymbol{P}_t)}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\rho}'} = \left[\frac{\partial^2 vec(\boldsymbol{P}_t)}{\partial \gamma \partial \boldsymbol{\rho}'}, \frac{\partial^2 vec(\boldsymbol{P}_t)}{\partial c \partial \boldsymbol{\rho}'} \right]. \tag{24}$$

The two $N(N-1) \times N^2$ blocks of the matrix (24) are

$$\frac{\partial^2 vec(\boldsymbol{P}_t)}{\partial x \partial \boldsymbol{\rho}'} = G_t(1 - G_t)a(x)\frac{\partial vec(\boldsymbol{P}_{(2)} - \boldsymbol{P}_{(1)})}{\partial \boldsymbol{\rho}'}.$$

with $x = \gamma, c$, where $a(x) = (t/T - c)$ for $x = \gamma$ and $a(x) = -\gamma$ for $x = c$.

Proof: Omitted.

6 Asymptotic normality

To prove asymptotic normality of the consistent ML estimator $\widehat{\boldsymbol{\theta}}_1$ of $\boldsymbol{\theta}_1$, we make the following additional assumption:

AN1. $\int \sup_{\boldsymbol{\theta}_1 \in \mathcal{N}_1} ||\frac{\partial}{\partial \boldsymbol{\theta}_1} f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)||d\boldsymbol{\zeta} < \infty$ and $\int \sup_{\boldsymbol{\theta}_1 \in \mathcal{N}_1} ||\frac{\partial^2}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'} f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)||d\boldsymbol{\zeta} < \infty$, where $|| \cdot ||$ is the Euclidean norm.

AN1 is a standard regularity condition. We state the following result:

Theorem 2 *Consider the time-varying smooth transition conditional correlation GARCH model (1), (14), (3), (5) and (6) where $\boldsymbol{D}_t = \boldsymbol{I}_N$. Let $\widehat{\boldsymbol{\theta}}_1$ be the maximum likelihood estimator of $\boldsymbol{\theta}_1^0$ defined in Theorem 1. Suppose that the conditions AG1-AG6 and AN1 hold with the modification that AG6 is strengthened to $\boldsymbol{\zeta}_t \sim iid\mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_N)$. Then*

$$\sqrt{T}(\widehat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^0) \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{B}(\boldsymbol{\theta}_1^0)^{-1})$$

as $T \to \infty$.

Proof: See Appendix 12.

The maximum likelihood estimator $\widehat{\boldsymbol{\theta}}_1$ is consistent and asymptotically normal but not efficient because it is assumed that $\boldsymbol{D}_t = \boldsymbol{I}_N$ in (7). The fact that $\widehat{\boldsymbol{\theta}}_1$ is consistent means, however, that the estimator provides a starting-point for further iterations during which the parameters in $\boldsymbol{D}_t$ are estimated as well.

7 Average Hessian and Expected Hessian

If the errors $\boldsymbol{\zeta}_t$ are not multivariate normal, the asymptotic covariance matrix of $\widehat{\boldsymbol{\theta}}_1$ has to be modified to fit this situation. For this purpose, we consider the average Hessian matrix of the log-likelihood and its expectation. Section 5.2 already contains the second partial derivatives of (8) with $\boldsymbol{D}_t = \boldsymbol{I}_N$ that are needed in these considerations. We state the following result:

Theorem 3 *Suppose that the assumptions of Theorem 2 hold. Denote $\boldsymbol{\theta}_1 = (\boldsymbol{\theta}'_g, \boldsymbol{\theta}'_\rho)' \in \Theta_1$ and let*

$$\boldsymbol{A}_T(\widehat{\boldsymbol{\theta}}_1) = \left[\begin{array}{cc} \boldsymbol{A}_{gT}(\widehat{\boldsymbol{\theta}}_g) & \boldsymbol{A}_{gPT}(\widehat{\boldsymbol{\theta}}_g, \widehat{\boldsymbol{\theta}}_\rho) \\ \boldsymbol{A}_{PgT}(\widehat{\boldsymbol{\theta}}_\rho, \widehat{\boldsymbol{\theta}}_g) & \boldsymbol{A}_{PT}(\widehat{\boldsymbol{\theta}}_\rho) \end{array} \right]$$

be the average Hessian evaluated at $\boldsymbol{\theta}_1 = \widehat{\boldsymbol{\theta}}_1$, where

$$\boldsymbol{A}_{gT}(\widehat{\boldsymbol{\theta}}_1) = \frac{1}{T} \sum_{t=1}^{T} \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t | \widehat{\boldsymbol{\theta}}_1)}{\partial \boldsymbol{\theta}_g \partial \boldsymbol{\theta}'_g}$$

$$\boldsymbol{A}_{gPT}(\widehat{\boldsymbol{\theta}}_1) = \frac{1}{T} \sum_{t=1}^{T} \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t | \widehat{\boldsymbol{\theta}}_1)}{\partial \boldsymbol{\theta}_g \partial \boldsymbol{\theta}'_\rho}$$

and

$$\boldsymbol{A}_{PT}(\widehat{\boldsymbol{\theta}}_1) = \frac{1}{T} \sum_{t=1}^{T} \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t | \widehat{\boldsymbol{\theta}}_1)}{\partial \boldsymbol{\theta}_\rho \partial \boldsymbol{\theta}'_\rho}.$$

Then, as $T \to \infty$, $\widehat{\boldsymbol{\theta}}_1 \xrightarrow{p} \boldsymbol{\theta}_1$ and, consequently,

$$\boldsymbol{A}_T(\widehat{\boldsymbol{\theta}}_1) \xrightarrow{p} \boldsymbol{A}(\boldsymbol{\theta}_1^0) = \left[\begin{array}{cc} \boldsymbol{A}_g(\boldsymbol{\theta}_1^0) & \boldsymbol{A}_{gP}(\boldsymbol{\theta}_1^0) \\ \boldsymbol{A}_{Pg}(\boldsymbol{\theta}_1^0) & \boldsymbol{A}_P(\boldsymbol{\theta}_1^0) \end{array} \right]$$

where $\boldsymbol{\theta}_1^0 = (\boldsymbol{\theta}_g^{0\prime}, \boldsymbol{\theta}_\rho^{0\prime})'$ is the true parameter. More specifically, the (j, i) element of $\boldsymbol{A}_g(\boldsymbol{\theta}_1^0)$ equals

$$[\boldsymbol{A}_g(\boldsymbol{\theta}_1^0)]_{ji} = -\frac{1}{4} \int_0^1 \frac{1}{g_{jr}^0 g_{ir}^0} \frac{\partial g_{jr}^0}{\partial \boldsymbol{\theta}_{gj}} \frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}'_{gi}} \boldsymbol{e}'_j (\boldsymbol{P}_r^0)^{-1} \boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{P}_r^0 \boldsymbol{e}_j \, dr$$

for $i \neq j$, and

$$[\boldsymbol{A}_g(\boldsymbol{\theta}_1^0)]_{ii} = -\frac{1}{4} \int_0^1 \frac{1}{(g_{it}^0)^2} \frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}_{gi}} \frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}'_{gi}} (1 + \boldsymbol{e}'_i (\boldsymbol{P}_r^0)^{-1} \boldsymbol{e}_i) \, dr$$

for $i = j$. Furthermore, the ith block of $\boldsymbol{A}_{gP}(\boldsymbol{\theta}_1^0)$ has the form

$$[\boldsymbol{A}_{gP}(\boldsymbol{\theta}_1^0)]_i = -\frac{1}{2} \int_0^1 \frac{1}{g_{ir}^0} \frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}_{gi}} (\boldsymbol{e}_i \otimes \boldsymbol{e}_i)' \{\boldsymbol{I}_N \otimes (\boldsymbol{P}_r^0)^{-1}\} \frac{\partial vec(\boldsymbol{P}_r^0)}{\partial \boldsymbol{\theta}'_\rho} \, dr$$

and, finally,

$$\boldsymbol{A}_P(\boldsymbol{\theta}_\rho^0) = -\frac{1}{2} \int_0^1 \frac{\partial vec(\boldsymbol{P}_r^0)}{\partial \boldsymbol{\theta}'_\rho} \{(\boldsymbol{P}_r^0)^{-1} \otimes (\boldsymbol{P}_r^0)^{-1}\} \frac{\partial vec(\boldsymbol{P}_r^0)}{\partial \boldsymbol{\theta}_\rho} \, dr.$$

Proof: See Appendix 12.

8 Maximum likelihood estimation of the complete model

As discussed in Amado and Teräsvirta (2013), ML estimation of MTV-GARCH models is numerically very difficult if the log-likelihood is maximised for all parameters at once. The situation does not become any easier if the model is a TVC model. It is therefore advantageous to split the estimation problem into two components and iterate between them. This is called maximisation by parts by Song et al. (2005) who give examples and consider the asymptotic theory behind this scheme. They are able to show that under regularity conditions, each iteration leads to consistent and asymptotically normal ML estimators, and the efficiency of the estimators increases with the increasing number of iterations. Amado and Teräsvirta (2013) applied maximisation by parts to the problem of maximising the log-likelihood of the MTV-GARCH model, and we shall do the same for the MTV-STCC-GARCH model. As in Section 3, the first parameter vector equals $\boldsymbol{\theta}_1 = (\boldsymbol{\theta}_g', \boldsymbol{\theta}_\rho')'$ and contains the parameters of the deterministic part, whereas $\boldsymbol{\theta}_h$ contains the GARCH parameters.

Supposing that our GARCH equations are of GJR-GARCH type, we make the following assumptions that hold for each of them in (1), see Amado and Teräsvirta (2013):

AH1. In (4), $\alpha_{i0} \geq 0$, $\alpha_{i0} + \kappa_{i1} > 0$, $\beta_{i1} \geq 0$, and $\alpha_{i1} + \kappa_{i1}/2 + \beta_{i1} < 1$ for $i = 1, ..., N$.

AH2. The parameter subspaces $\{\alpha_{i0} \times \kappa_i \times \alpha_i \times \beta_i\}$, $i = 1, ..., N$, are compact, the whole space Θ_h is compact, and the true parameter value $\boldsymbol{\theta}_h^0$ is an interior point of Θ_h.

AH3. $\mathsf{E}\phi_{it}^4 < \infty$.

AH1 is the necessary and sufficient weak stationarity condition for the ith first-order GJR-GARCH equation. Conditions AH2 and AH3 are standard regularity conditions required for proving asymptotic normality of ML estimators of $\boldsymbol{\theta}_{hi}$, $i = 1, ..., N$. We state the following result:

Theorem 4 *Consider the GARCH model (1) with (2) and $\boldsymbol{\varepsilon}_t = \boldsymbol{S}_t \boldsymbol{D}_t \boldsymbol{z}_t$ with $\boldsymbol{D}_t^0$. Suppose that Assumptions AG6 and AH1-3 hold. Furthermore, assume that $\boldsymbol{\phi}_t = \boldsymbol{S}_t^{-1} \boldsymbol{\varepsilon}_t$ holds. Then the maximum likelihood estimators of $\boldsymbol{\theta}_{hi}$, $i = 1, ..., N$, that is, the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}_h$ of $\boldsymbol{\theta}_h$, are consistent and*

$$\sqrt{T}(\widehat{\boldsymbol{\theta}}_h - \boldsymbol{\theta}_h) \xrightarrow{d} N(\mathbf{0}_N, \boldsymbol{V}_h) \tag{25}$$

as $T \to \infty$, where $\boldsymbol{V}_h$ is the asymptotic covariance matrix of $\widehat{\boldsymbol{\theta}}_h$ given $\boldsymbol{\theta}_1^0$.

Proof: This result is proved in Amado and Teräsvirta (2013) for a single GJR-GARCH equation. When assumptions AG6 and AH1-AH3 hold for all equations simultaneously and the ML estimator $\widehat{\boldsymbol{\theta}}_h$ is conditioned on $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^0$ that includes $\boldsymbol{\theta}_\rho = \boldsymbol{\theta}_\rho^0$ the asymptotic result (25) is valid. ■

In Theorem 4 it is assumed that $\boldsymbol{\theta}_1^0$ is known, which is not the case in practice, and we have to consider the situation in which $\boldsymbol{\theta}_1^0$ is estimated. The log-likelihood

$$L_T(\boldsymbol{\theta}) = (1/T)\sum_{t=1}^{T} \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}, \mathcal{F}_{t-1})$$

can be decomposed as follows:

$$L_T(\boldsymbol{\theta}) = L_T(\boldsymbol{\theta}_1) + L_T(\boldsymbol{\theta}_1, \boldsymbol{\theta}_h). \tag{26}$$

Given the available data, (26) is maximised numerically. Assuming the results in Theorems 2 and 4 are valid and using the result in Song et al. (2005, Theorem 3), we are able to conclude that after the kth iteration, the ML estimator $\widehat{\boldsymbol{\theta}}^{(k)} = ((\widehat{\boldsymbol{\theta}}_1^{(k)})', (\widehat{\boldsymbol{\theta}}_h^{(k)})')'$ is consistent and asymptotically normal, that is,

$$\sqrt{T}(\widehat{\boldsymbol{\theta}}^{(k)} - \boldsymbol{\theta}^0) \xrightarrow{d} N(\mathbf{0}_N, \boldsymbol{V}^{(k)})$$

where $\boldsymbol{V}^{(k)}$ is the asymptotic covariance matrix after k iterations, see Song et al. (2005). When $k \to \infty$, $\boldsymbol{V}^{(k)} \to \boldsymbol{V}$, where $\boldsymbol{V}$ is the asymptotic covariance matrix of $\sqrt{T}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)$, $\widehat{\boldsymbol{\theta}}$ being the final ML estimator of $\boldsymbol{\theta}^0 = ((\boldsymbol{\theta}_1^0)', (\boldsymbol{\theta}_h^0)')'$.

It should be noted that while the proof of Theorem 3 in Song et al. (2005) requires validity of the so-called information dominance condition, this condition is not needed here because consistency of $\widehat{\boldsymbol{\theta}}_1$ holds due to Theorem 1.

9   Numerical considerations

The slope parameter γ_{ij} in (3) may be numerically difficult to estimate when it is large. Reasons for this difficulty are discussed for example in Teräsvirta, Granger and Tjøstheim (2010, p. 403). In practice it may be useful to apply the transformation $\gamma_{ij} = \exp\{\eta_{ij}\}$, in which case η_{ij} need not be restricted. The motivation for this transformation is that estimating η_{ij} instead of γ_{ij} is numerically convenient in cases where γ_{ij} is large, see Goodwin, Holt and Prestemon (2011) or Silvennoinen and Teräsvirta (2016) for discussion. There is another advantage: the identifying restriction $\gamma > 0$ becomes redundant because $\exp\{\eta_{ij}\} > 0$ for any finite η_{ij}.

As already discussed, estimation of parameters is not a simple task, and the log-likelihood function has to be maximised such that every iteration is split into four stages. To study this, rewrite (26) as follows:

$$
\begin{aligned}
L_T(\boldsymbol{\theta}) ={}& -\sum_{i=1}^{N}\sum_{t=1}^{T} \ln g_i(\boldsymbol{\theta}_{gi}; t/T) - \sum_{i=1}^{N}\sum_{t=1}^{T} \ln h_{it}(\boldsymbol{\theta}_{hi}, \boldsymbol{\theta}_{gi}) \\
& -(1/2)\sum_{t=1}^{T} \ln |\boldsymbol{P}_t(\boldsymbol{\theta}_\rho)| \\
& -(1/2)\sum_{t=1}^{T} \boldsymbol{\varepsilon}_t'\{\boldsymbol{S}_t(\boldsymbol{\theta}_g)\boldsymbol{D}_t(\boldsymbol{\theta}_h, \boldsymbol{\theta}_g)\boldsymbol{P}_t(\boldsymbol{\theta}_\rho)\boldsymbol{D}_t(\boldsymbol{\theta}_h, \boldsymbol{\theta}_g)\boldsymbol{S}_t(\boldsymbol{\theta}_g)\}^{-1}\boldsymbol{\varepsilon}_t.
\end{aligned}
$$

Let

$$g_i(\boldsymbol{\theta}_{gi}; t/T) = \delta_{i0} + \sum_{j=1}^{r_i} \delta_{ij} G_{ij}(\gamma_{ij}, \boldsymbol{c}_{ij}; t/T) \tag{27}$$

(note the free intercept δ_{i0}). Estimation proceeds as follows.

STEP 1. Assume $\ln h_{it}(\boldsymbol{\theta}_{hi}, \boldsymbol{\theta}_{gi}) = 0$, $i = 1, ..., N$, and estimate parameters $\boldsymbol{\theta}_g = (\boldsymbol{\theta}_{g1}, ..., \boldsymbol{\theta}_{gN})'$, $i = 1, ..., N$, equation by equation, assuming $\boldsymbol{P}_t(\boldsymbol{\theta}_\rho) = \boldsymbol{I}_N$. Denote the estimate $\boldsymbol{S}_t(\widehat{\boldsymbol{\theta}}_g^{(1,1)})$. This means that the deterministic components $g_i(\boldsymbol{\theta}_{gi}; t/T)$ have been estimated once, including the intercept δ_{0i} in (27).

STEP 2. Estimate $\boldsymbol{P}_t(\boldsymbol{\theta}_\rho)$ given $\boldsymbol{\theta}_g = \widehat{\boldsymbol{\theta}}_g^{(1,1)}$. This requires a separate iteration because $\boldsymbol{P}_t(\boldsymbol{\theta}_\rho)$ is nonlinear in parameters, see (5) and (6). Denote the estimate $\boldsymbol{P}_t(\widehat{\boldsymbol{\theta}}_\rho^{(1,1)})$.

STEP 3. Re-estimate $\boldsymbol{S}_t(\boldsymbol{\theta}_g)$ assuming $\boldsymbol{P}_t(\boldsymbol{\theta}_\rho) = \boldsymbol{P}_t(\widehat{\boldsymbol{\theta}}_\rho^{(1,1)})$. This yields $\boldsymbol{S}_t(\widehat{\boldsymbol{\theta}}_g^{(1,2)})$. Then re-estimate $\boldsymbol{P}_t(\boldsymbol{\theta}_\rho)$ given $\boldsymbol{\theta}_g = \widehat{\boldsymbol{\theta}}_g^{(1,2)}$. Iterate until convergence. Let the result after R_1 iterations be $\boldsymbol{S}_t(\boldsymbol{\theta}_g) = \boldsymbol{S}_t(\widehat{\boldsymbol{\theta}}_g^{(1,R_1)})$ and $\boldsymbol{P}_t(\boldsymbol{\theta}_\rho) = \boldsymbol{P}_t(\widehat{\boldsymbol{\theta}}_\rho^{(1,R_1)})$. The resulting estimates are maximum likelihood ones under the assumption $\boldsymbol{D}_t(\boldsymbol{\theta}_h, \boldsymbol{\theta}_g) = \boldsymbol{I}_N$.

STEP 4. Estimate $\boldsymbol{\theta}_h$ from $\boldsymbol{D}_t(\boldsymbol{\theta}_h, \widehat{\boldsymbol{\theta}}_g^{(1,R_1)})$ using $\boldsymbol{P}_t(\boldsymbol{\theta}_\rho) = \boldsymbol{P}_t(\widehat{\boldsymbol{\theta}}_\rho^{(1,R_1)})$. This is a standard multivariate conditional correlation GARCH estimation step as in Bollerslev (1990), because $\boldsymbol{S}_t(\widehat{\boldsymbol{\theta}}_g^{(1,R_1)})$ is fixed and does not affect the maximum, and $\boldsymbol{P}_t(\widehat{\boldsymbol{\theta}}_\rho^{(1,R_1)})$ is known. In total, steps 1–4 form the first iteration of the maximisation algorithm. Denote the estimate $\widehat{\boldsymbol{\theta}}_h^{(1)}$.

STEP 5. Estimate $\boldsymbol{\theta}_g$ from $\boldsymbol{S}_t(\boldsymbol{\theta}_g)$ keeping $\boldsymbol{D}_t(\widehat{\boldsymbol{\theta}}_h^{(1)}, \widehat{\boldsymbol{\theta}}_g^{(1,R_1)})$ and $\boldsymbol{P}_t(\widehat{\boldsymbol{\theta}}_\rho^{(1,R_1)})$ fixed. This step is analogous to the first part of Step 3. The difference is that $\boldsymbol{D}_t(\widehat{\boldsymbol{\theta}}_h^{(1)}, \widehat{\boldsymbol{\theta}}_g^{(1,R_1)}) \neq \boldsymbol{I}_N$. Denote the estimator $\boldsymbol{S}_t(\widehat{\boldsymbol{\theta}}_g^{(2,1)})$.

STEP 6. Estimate $\boldsymbol{P}_t(\boldsymbol{\theta}_\rho)$ given $\boldsymbol{\theta}_g = \widehat{\boldsymbol{\theta}}_g^{(2,1)}$ and $\boldsymbol{\theta}_h = \widehat{\boldsymbol{\theta}}_h^{(1)}$. Denote the estimator $\boldsymbol{P}_t(\widehat{\boldsymbol{\theta}}_\rho^{(2,1)})$. Iterate until convergence, R_2 iterations. The result: $\boldsymbol{S}_t(\boldsymbol{\theta}_g) = \boldsymbol{S}_t(\widehat{\boldsymbol{\theta}}_g^{(2,R_2)})$ and $\boldsymbol{P}_t(\boldsymbol{\theta}_\rho) = \boldsymbol{P}_t(\widehat{\boldsymbol{\theta}}_\rho^{(2,R_2)})$.

STEP 7. Estimate $\boldsymbol{\theta}_h$ from $\boldsymbol{D}_t(\boldsymbol{\theta}_h, \widehat{\boldsymbol{\theta}}_g^{(2,R_2)})$ using $\boldsymbol{P}_t(\boldsymbol{\theta}_\rho) = \boldsymbol{P}_t(\widehat{\boldsymbol{\theta}}_\rho^{(2,R_2)})$ ($\boldsymbol{S}_t(\widehat{\boldsymbol{\theta}}_g^{(2,R_2)})$ is fixed). The result: $\boldsymbol{\theta}_h = \widehat{\boldsymbol{\theta}}_h^{(2)}$. This completes the second full iteration.

STEP 8. Repeat steps 5–7 and iterate until convergence.

For identification reasons, δ_{0i}, $i = 1, ..., N$, is frozen to $\delta_{0i} = \widehat{\delta}_{0i}^{(1,R_1)}$. This frees the intercepts in $\boldsymbol{\theta}_{hi}$. Any positive constant would do for δ_{0i}, but for numerical reasons the intercepts are fixed to the values they obtain after the first iteration when $\boldsymbol{\theta}_h$ is not yet estimated a single time.

10 Empirical example

As an illustration we consider daily log-returns of the S&P 500 index and the 30-year US Treasury bill from 3 January 2000 to 6 July 2015, 4046 observations. The series appear in Figures 1 and 2.

It is seen that the amplitude of volatility clusters varies strongly over time in both series. After a turbulent period until about the year 2004, relative calm prevails until late 2007 and again from about 2011-2012 onwards. The stock returns are more volatile than bond returns. There are a few extremely large absolute returns in Figure 1.

For a later comparison, we fit a first-order GJR-GARCH model to index returns ε_t^S and a standard first-order GARCH model to bond returns ε_t^B. This results in the following two equations:

$$\widehat{h}_t^S = \underset{(0.0027)}{0.0189} + \underset{(0.0150)}{0.164}\, I(\varepsilon_{t-1}^S < 0)(\varepsilon_{t-1}^S)^2 + \underset{(0.0081)}{0.901}\, \widehat{h}_{t-1}^S \qquad (28)$$

and

$$\widehat{h}_t^B = \underset{(2\times 10^{-4})}{5 \times 10^{-4}} + \underset{(0.0037)}{0.0290}(\varepsilon_{t-1}^B)^2 + \underset{(0.0040)}{0.968}\, \widehat{h}_{t-1}^B \qquad (29)$$

Persistence is high in both cases: $\widehat{\kappa}_1/2 + \widehat{\beta}_1 = 0.983$ in (28) and $\widehat{\alpha}_1 + \widehat{\beta}_1 = 0.997$ in (29), as the GARCH models try to accommodate the nonstationarity present in the series. We set $\alpha_1 = 0$ in (28) because it was originally estimated extremely close to the boundary value zero. The estimated conditional standard deviations $(\widehat{h}_t^S)^{1/2}$ and $(\widehat{h}_t^B)^{1/2}$ appear in Figures 3 and 4. It is seen that they do contain level shifts (equation (28)) and large bumps (equation (29)) that are indications of nonstationarity.

Next we fit a TVC model to these two series. The sequence of specification tests described in a companion paper Hall, Silvennoinen and Teräsvirta (2017) suggests three transitions for the TV-GARCH component of the index return and two for the bond return equation. The estimated equations can be found in Appendix 13. The deterministic components are graphed in Figures 5 and 6. A comparison of Figure 5 with Figure 3 shows that, as expected, the shape of $\widehat{g}_t^S$ follows level shifts in Figure 3. The first transition is very rapid, and the remaining two with opposite signs (see equation (68)) are responsible for the slightly asymmetric peak around 2009. Similarly, the shape of the deterministic component in Figure 6 agrees with movements in Figure 4.

The rescaling effect of the deterministic component becomes obvious from Figures 7 and 8 depicting rescaled returns $\varepsilon_t^S/(\widehat{g}_t^S)^{1/2}$ (index returns) and $\varepsilon_t^B/(\widehat{g}_t^B)^{1/2}$ (T-bill returns). The amplitude changes in Figures 1 and 2 have disappeared, and the series look more (weakly) stationary. Estimates of kurtosis have decreased from 13.7 to 8.2 (index returns) and from 7.3 to 5.2 (T-bill returns). However, the largest rescaled absolute S&P500 returns still stick out, although their relative magnitude is smaller than in Figure 1.

Another way of illustrating the effect of rescaling on the results is to look at the rescaled conditional standard deviations from equations (67) and (70). They can be found in Figures 9 and 10. Standard deviations for the index returns now look stationary, except two sharp peaks, and the persistence measure $\widehat{\kappa}_1/2 + \widehat{\beta}_1 = 0.954$,
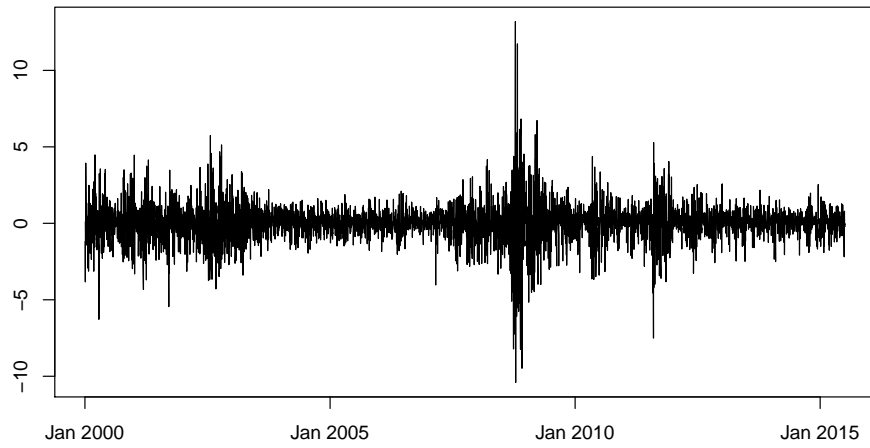
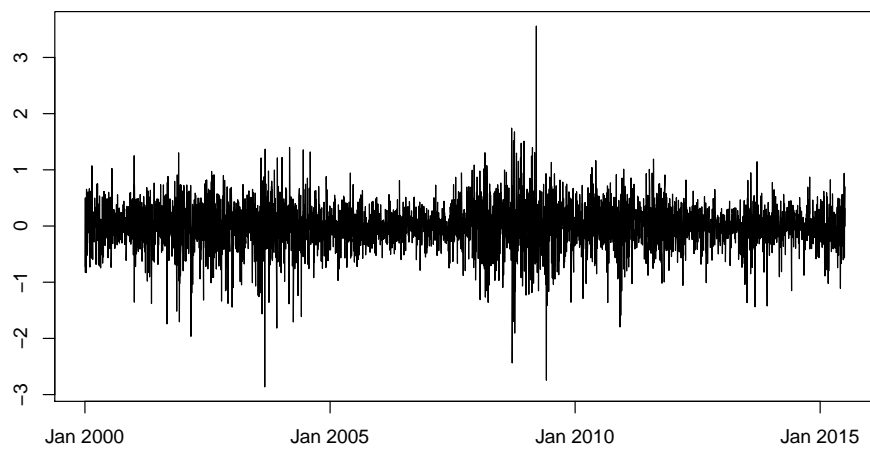Figure 1: S&P 500 index daily log-returns, 3 January 2000 - 6 July 2015.

Figure 2: 30-year Treasury bill daily log-returns, 3 January 2000 - 6 July 2015.
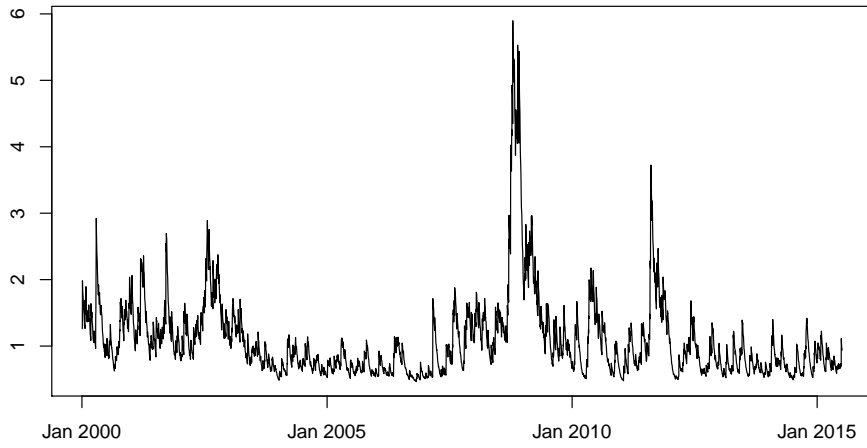
Figure 3: Estimated conditional standard deviations from the equation (28).

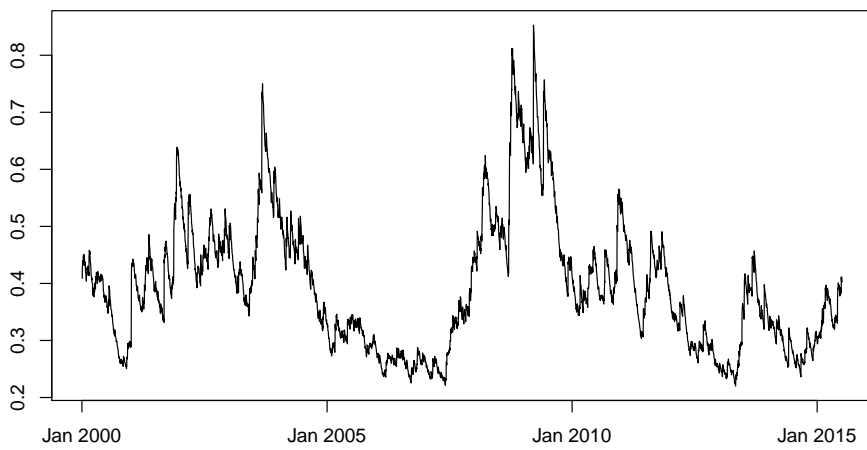Figure 4: Estimated conditional standard deviations from the equation (29).
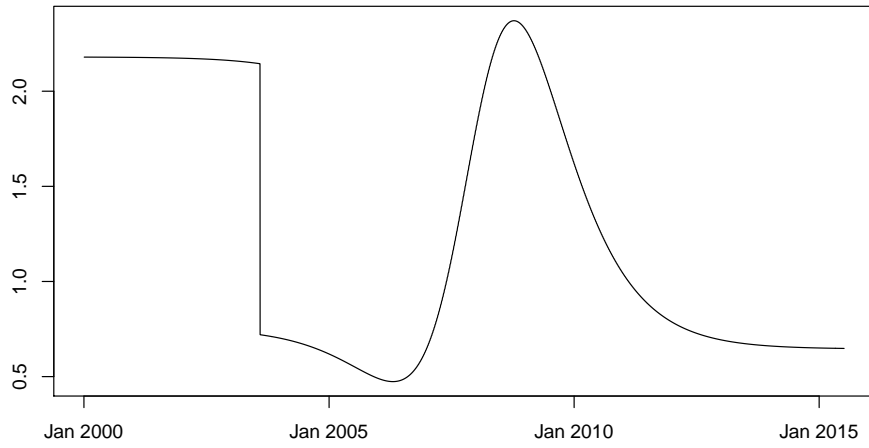
Figure 5: Estimated deterministic component $\widehat{g}_t^S$ from equation (68).
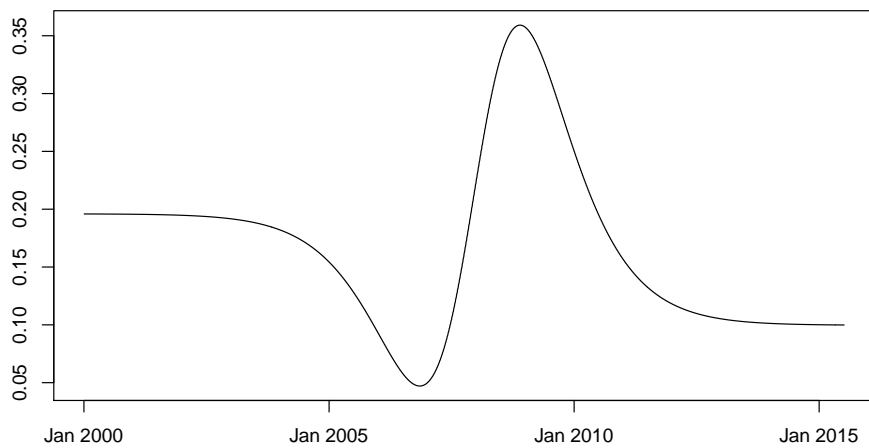
Figure 6: Estimated deterministic component $\widehat{g}_t^B$ from equation (71).

19

Figure 7: Rescaled S&P500 index returns, 3 January 2000 - 6 July 2015.

Figure 8: Rescaled 30-year Treasury bill returns, 3 January 2000 - 6 July 2015.

20

Figure 9: Rescaled conditional standard deviations for S&P500 index returns from equation (67), 3 January 2000 - 6 July 2015.

Figure 10: Rescaled conditional standard deviations for 30-year Treasury bill returns from equation (70), 3 January 2000 - 6 July 2015.

a decrease from 0.983. The corresponding ones for the T-bill returns seem somewhat more persistent. Even there, the persistence measure has decreased from 0.997 to 0.986.

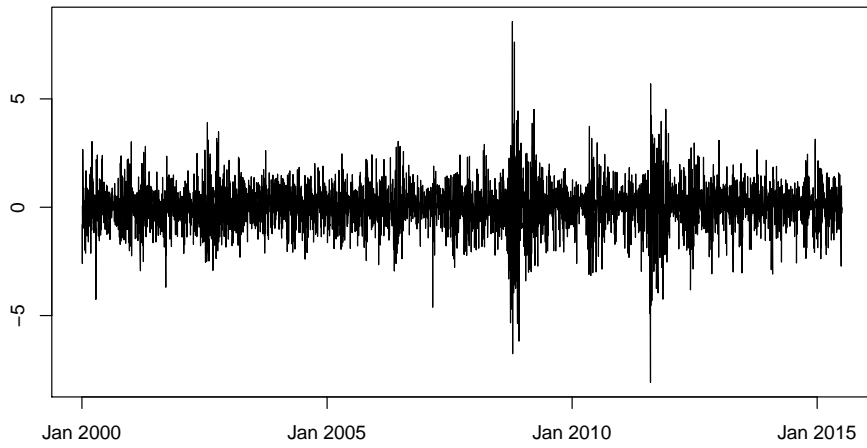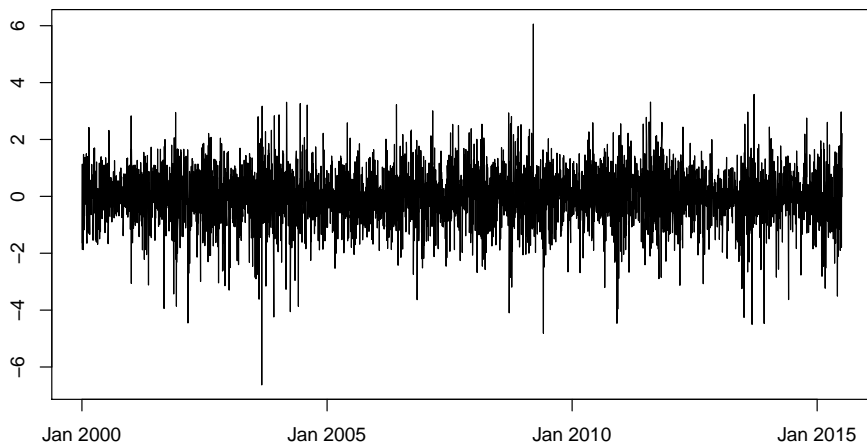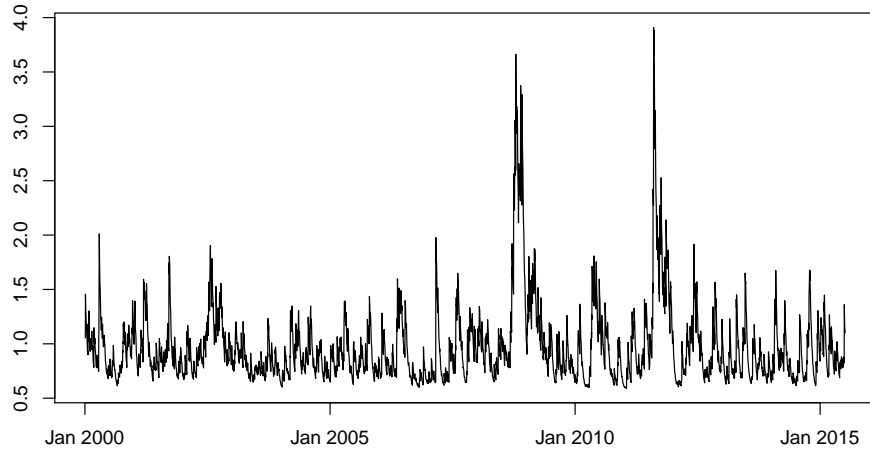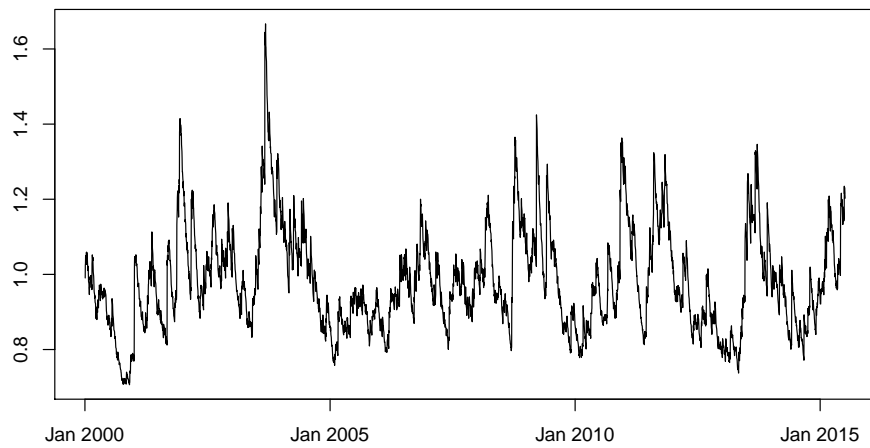Yet another way of looking at the contribution of g_t is to compare autocorrelations of ε_t^2 with those of $\varepsilon_t^2/\widehat{g}_t$. For the S&P500 returns, they appear in Figure 11. The original autocorrelations are positive until about lag 200, whereas the rescaled ones approach zero more quickly. Their decay rate is closer to exponential, which is the case for autocorrelations from the weakly stationary GARCH(1,1) model.

The squared autocorrelations for bond returns in Figure 12 look different.They are low but positive and persistent. Rescaling reduces them further. Even if we followed a standard procedure and fitted a first-order GARCH model to these returns, that may not seem an optimal choice. A look at the lower panel gives the impression that there is not much nonlinear dependence left to be modelled after rescaling. The question then is: what does the persistence value 0.972 for the GARCH component of the model really tell us? It seems that completing this figure with information from the autocorrelations of $\varepsilon_t^2/\widehat{g}_t$ is quite useful.

Finally, Figure 13 contains the time-varying correlations. They are clearly negative until 2003, rise slightly above zero thereafter and rather quickly drop to their previous level at the outset of the financial crisis in 2007. This shape also becomes visible, albeit smoother than in the figure, when a window of raw correlations between $\widehat{z}_t^S = \varepsilon_t^S/(\widehat{h}_t^S\widehat{g}_t^S)^{1/2}$ and $\widehat{z}_t^B = \varepsilon_t^B/(\widehat{h}_t^B\widehat{g}_t^B)^{1/2}$ is moved over the observation period. The shift is very systematic and as such difficult to describe by a VC- or DCC-GARCH model. The reason is that these models contain the assumption that the correlations fluctuate around a constant level during the observation period. It would seem possible to incorporate a more short-run stochastic component to the correlations, see Silvennoinen and Teräsvirta (2009a), but such an extension is left for the future.

11   Conclusions

In this paper we examine a special case of the MTV-STCC-GARCH with deterministically time-varying correlations or the TVC model in which the transition function is rescaled time. An advantage of TVC is that time-varying correlations can vary smoothly but also systematically. For example, the correlations can change from one level to another and remain there. In VC- or DCC-GARCH models the conditional correlations by definition fluctuate (or cluster) around constant values. We show that the ML estimators of parameters of the TVC model are consistent and asymptotically normal. This allows us, among other things, to construct tests to test constancy (CCC-GARCH) against TVC and, in particular, determine the number of transitions in TVC-type correlation matrices by sequential testing. Furthermore, it is possible to derive misspecification tests for this model. These modelling issues are considered in a companion paper Hall et al. (2017).

Engle and Kelly (2012) studied a special case of the DCC-GARCH, the dynamic equicorrelation model or DECO, in which, as the name suggests, all correlations are identical. This can be a useful model in situations where the number of series to be modelled and thus the number of parameters are both large. Deriving its
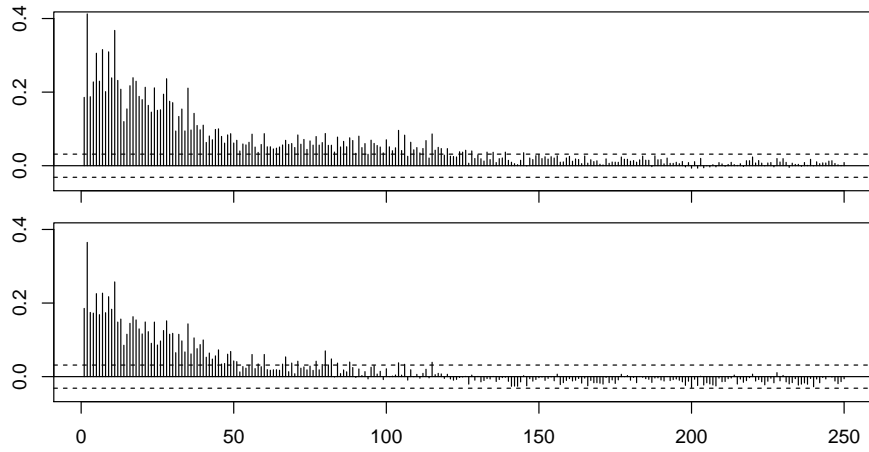
Figure 11: First 250 autocorrelations of daily squared (top panel) and rescaled squared returns of the S&P500 index, 3 January 2000 - 6 July 2015.
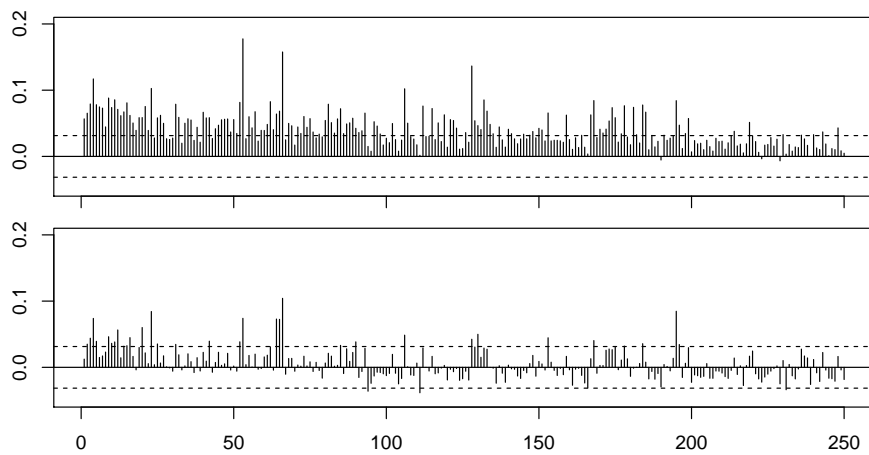

Figure 12: First 250 autocorrelations of daily squared (top panel) and rescaled squared returns of the 30-year Treasury bill, 3 January 2000 - 6 July 2015.

23
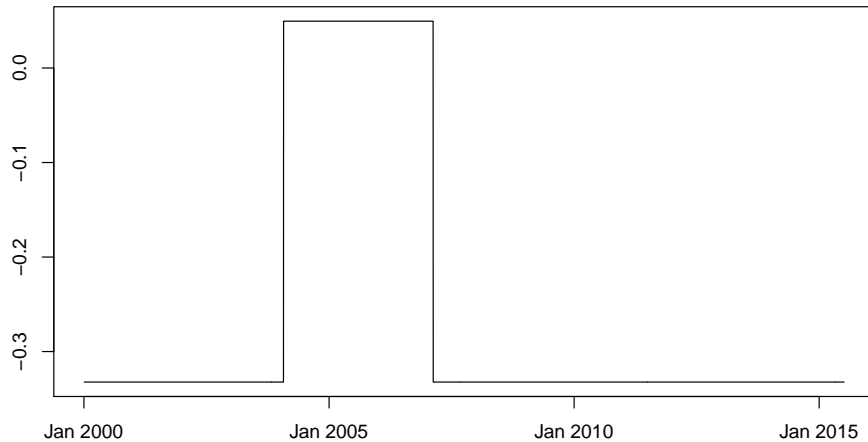
Figure 13: Time-varying correlations between daily S&P 500 index and 30-year Treasury bill returns from equation (72), 3 January 2000 - 6 July 2015.

TVC counterpart, the Time-Varying Equicorrelation (TVEC) model which in turn is a special case of the TVC model, is a straightforward exercise. Since asymptotic properties of the ML estimators are known in the general TVC case they are also known in the restricted TVEC case. This will allow us to build TVEC models using statistical tests, and we could also test TVEC against TVC. This work as well as extensions to the TVC model such as the double TVC model, see Silvennoinen and Teräsvirta (2009a), will be left for the future.

Finally, it should be mentioned once more that in this work focus has been on estimation of parameters and asymptotic properties of maximum likelihood estimators. Other aspects of modelling will be studied in detail in Hall et al. (2017).

References

Amado, C. and Teräsvirta, T.: 2008, Modelling conditional and unconditional heteroskedasticity with smoothly time-varying structure, *SSE/EFI Working Paper Series in Economics and Finance 691*, Stockholm School of Economics.

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

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

Amado, C. and Teräsvirta, T.: 2014b, Modelling changes in the unconditional variance of long stock return series, *Journal of Empirical Finance* **25**, 15–35.

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

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

Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O.: 1996, Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **74**, 3–30.

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

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

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

Bollerslev, T.: 1990, Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model, *Review of Economics and Statistics* **72**, 498–505.

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

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

Conrad, C., Karanasos, M. and Zeng, N.: 2011, Multivariate fractionally integrated APARCH modeling of stock market volatility: A multi-country study, *Journal of Empirical Finance* **18**, 147–159.

Dark, J.: 2015, A multivariate conditional correlation model with long memory dependence and asymmetries, *Working paper*, University of Melbourne.

Eklund, B. and Teräsvirta, T.: 2007, Testing constancy of the error covariance matrix in vector models, *Journal of Econometrics* **140**, 753–780.

Engle, R. F.: 2002, Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business and Economic Statistics* **20**, 339–350.

Engle, R. F. and Bollerslev, T.: 1986, Modelling the persistence of conditional variances, *Econometric Reviews* **5**, 1–50.

Engle, R. F. and Kelly, B.: 2012, Dynamic equicorrelation, *Journal of Business and Economic Statistics* **30**, 212–228.

Engle, R. F. and Kroner, K. F.: 1995, Multivariate simultaneous generalized ARCH, *Econometric Theory* **11**, 122–150.

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

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

Gallant, A. R.: 1981, On the bias in the flexible functional form and an essentially unbiased form. The Fourier Flexible Form, *Journal of Econometrics* **15**, 211–245.

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

Goodwin, B. K., Holt, M. T. and Prestemon, J. P.: 2011, North American oriented strand board markets, arbitrage activity, and market price dynamics: A smooth transition approach, *American Journal of Agricultural Economics* **93**, 993–1014.

Hafner, C. M. and Linton, O.: 2010, Efficient estimation of a multivariate multiplicative volatility model, *Journal of Econometrics* **159**, 55–73.

Hafner, C., van Dijk, D. and Franses, P. H.: 2006, Semiparametric modelling of correlation dynamics, *in* T. Fomby and C. Hill (eds), *Advances in Econometrics*, Vol. 20, Part A, Emerald, pp. 59–103.

Hall, A. D., Silvennoinen, A. and Teräsvirta, T.: 2017, Building multiplicative time-varying smooth transition conditional correlation GARCH models, *work in progress*, School of Economics and Finance, Queensland University of Technology.

Hillebrand, E., Medeiros, M. C. and Xu, J.: 2013, Asymptotic theory for regressions with smoothly changing parameters, *Journal of Time Series Econometrics* **5**, 133–162.

Jeantheau, T.: 1998, Strong consistency of estimators for multivariate ARCH models, *Econometric Theory* **14**, 70–86.

Jin, X. and Maheu, J. M.: 2016, Modeling covariance breakdowns in multivariate GARCH, *Journal of Econometrics* **194**, 1–23.

Li, W. K. and Mak, T. K.: 1994, On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity, *Journal of Time Series Analysis* **15**, 627–636.

Ling, S. and McAleer, M.: 2003, Asymptotic theory for a vector ARMA–GARCH model, *Econometric Theory* **19**, 280–310.

Long, X., Su, L. and Ullah, A.: 2011, Estimation and forecasting of dynamic conditional covariance: a semiparametric multivariate model, *Journal of Business and Economic Statistics* **29**, 109–125.

Lundbergh, S. and Teräsvirta, T.: 2002, Evaluating GARCH models, *Journal of Econometrics* **110**, 417–435.

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

Mazur, B. and Pipień, M.: 2012, On the empirical importance of periodicity in the volatility of financial returns - time varying GARCH as a second order APC(2) process, *Central European Journal of Economic Modelling and Econometrics* **4**, 95–116.

Nakatani, T. and Teräsvirta, T.: 2009, Testing for volatility interactions in the Constant Conditional Correlation GARCH model, *Econometrics Journal* **12**, 147–163.

Newey, W. K. and McFadden, D. L.: 1994, Large sample estimation and hypothesis testing, *in* R. F. Engle and D. L. McFadden (eds), *Handbook of Econometrics*, Vol. 4, Elsevier Science, Amsterdam, pp. 2111–2245.

Pedersen, R. S.: 2017, Inference and testing on the boundary in extended constant conditional correlation GARCH models, *Journal of Econometrics* **196**, 23–36.

Péguin-Feissolle, A. and Sanhaji, B.: 2016, Tests of the constancy of the conditional correlations of unknown functional form in multivariate GARCH models, *Annales d'économie et de statistique* pp. 77–101.

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

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

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

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

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

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

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

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

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

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

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

12   Proofs

Proof of Lemma 1: We begin by introducing the following lemma.

Lemma 8 *The first partial derivatives*

$$\frac{\partial \ln |\boldsymbol{P}_t|}{\partial vec(\boldsymbol{P}_t)} = vec(\boldsymbol{P}_t^{-1})$$

$$\frac{\partial (\boldsymbol{z}_t' \boldsymbol{P}_t^{-1} \boldsymbol{z}_t)}{\partial vec(\boldsymbol{P}_t^{-1})} = vec(\boldsymbol{z}_t \boldsymbol{z}_t')$$

and

$$\frac{\partial vec(\boldsymbol{P}_t^{-1})'}{\partial vec(\boldsymbol{P}_t)} = -(\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1}). \tag{30}$$

Proof: See Lütkepohl (1996).

Now we can prove Lemma 1. Consider first the ith element of (10):

$$\boldsymbol{s}_t(\boldsymbol{\theta}_{gi}) = \frac{\partial \ln f(\boldsymbol{\zeta}_t | \boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi}} = -\frac{\partial \ln g_{it}}{2 \partial \boldsymbol{\theta}_{gi}} - \frac{\partial \boldsymbol{z}_t' \boldsymbol{P}_t^{-1} \boldsymbol{z}_t}{2 \partial \boldsymbol{\theta}_{gi}}. \tag{31}$$

The second term on the r.h.s. becomes

$$
\begin{aligned}
\frac{\partial \boldsymbol{z}_t' \boldsymbol{P}_t^{-1} \boldsymbol{z}_t}{\partial \boldsymbol{\theta}_{gi}} &= \frac{\partial \boldsymbol{z}_t'}{\partial \boldsymbol{\theta}_{gi}} \frac{\partial \boldsymbol{z}_t' \boldsymbol{P}_t^{-1} \boldsymbol{z}_t}{\partial \boldsymbol{z}_t} \\
&= \frac{2 \partial \boldsymbol{z}_t'}{\partial \boldsymbol{\theta}_{gi}} \boldsymbol{P}_t^{-1} \boldsymbol{z}_t
\end{aligned} \tag{32}
$$

where, letting $\boldsymbol{e}_i = (\boldsymbol{0}'_{i-1}, 1, \boldsymbol{0}'_{N-i})'$, the $(3r_i + 1) \times N$ matrix $\partial \boldsymbol{z}'_t / \partial \boldsymbol{\theta}_{gi}$ equals

$$
\begin{aligned}
\frac{\partial \boldsymbol{z}'_t}{\partial \boldsymbol{\theta}_{gi}} &= \frac{\partial}{\partial \boldsymbol{\theta}_{gi}} (\varepsilon_{1t} g_{1t}^{-1/2}, ..., \varepsilon_{Nt} g_{Nt}^{-1/2}) \\
&= (\boldsymbol{0}_{3r_1+1}, ..., \boldsymbol{0}_{3r_{i-1}+1}, -\frac{\varepsilon_{it}}{2g_t^{3/2}} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}}, \boldsymbol{0}_{3r_{i+1}+1}, ..., \boldsymbol{0}_{3r_N+1}) \\
&= -\frac{1}{2g_{it}} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}} (0, ..., 0, \varepsilon_{it} g_{it}^{-1/2}, 0, ..., 0) = -z_{it} \frac{1}{2g_{it}} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}} \boldsymbol{e}'_i. \quad (33)
\end{aligned}
$$

Denoting $z_{it} = \boldsymbol{z}'_t \boldsymbol{e}_i$ and inserting (33) into (32) yields

$$
\frac{\partial \boldsymbol{z}'_t \boldsymbol{P}_t^{-1} \boldsymbol{z}_t}{\partial \boldsymbol{\theta}_{gi}} = -\frac{1}{g_{it}} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}} \boldsymbol{z}'_t \boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} \boldsymbol{z}_t.
$$

Plugging this into (31) yields the following $(3r_i + 1)$-vector:

$$
\boldsymbol{s}_t(\boldsymbol{\theta}_{gi}) = \frac{1}{2g_{it}} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}} (\boldsymbol{z}'_t \boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} \boldsymbol{z}_t - 1) \quad (34)
$$

$i = 1, ..., N$. In order to obtain (11), write

$$
\frac{\partial \ln |\boldsymbol{P}_t|}{\partial \boldsymbol{\theta}'_\rho} = \frac{\partial \ln |\boldsymbol{P}_t|}{\partial \mathrm{vec}(\boldsymbol{P}_t)'} \frac{\partial \mathrm{vec}(\boldsymbol{P}_t)}{\partial \boldsymbol{\theta}'_\rho} \quad (35)
$$

and

$$
\frac{\partial (\boldsymbol{z}'_t \boldsymbol{P}_t^{-1} \boldsymbol{z}_t)}{\partial \boldsymbol{\theta}'_\rho} = \frac{\partial (\boldsymbol{z}'_t \boldsymbol{P}_t^{-1} \boldsymbol{z}_t)}{\partial \mathrm{vec}(\boldsymbol{P}_t^{-1})'} \frac{\partial \mathrm{vec}(\boldsymbol{P}_t^{-1})}{\partial \mathrm{vec}(\boldsymbol{P}_t)'} \frac{\partial \mathrm{vec}(\boldsymbol{P}_t)}{\partial \boldsymbol{\theta}'_\rho}. \quad (36)
$$

Applying Lemma 8 to (35) and (36), one obtains

$$
\frac{\partial \ln |\boldsymbol{P}_t|}{\partial \boldsymbol{\theta}_\rho} = \frac{\partial \mathrm{vec}(\boldsymbol{P}_t)'}{\partial \boldsymbol{\theta}_\rho} \mathrm{vec}(\boldsymbol{P}_t^{-1})
$$

and

$$
\frac{\partial (\boldsymbol{z}'_t \boldsymbol{P}_t^{-1} \boldsymbol{z}_t)}{\partial \boldsymbol{\theta}_\rho} = -\frac{\partial \mathrm{vec}(\boldsymbol{P}_t)'}{\partial \boldsymbol{\theta}_\rho} (\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1}) \mathrm{vec}(\boldsymbol{z}_t \boldsymbol{z}'_t)
$$

so

$$
\begin{aligned}
\boldsymbol{s}_t(\boldsymbol{\theta}_\rho) &= \frac{\partial \ln f(\boldsymbol{\zeta}_t | \boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_\rho} = -(1/2) \frac{\partial \mathrm{vec}(\boldsymbol{P}_t)'}{\partial \boldsymbol{\theta}_\rho} \{\mathrm{vec}(\boldsymbol{P}_t^{-1}) \\
&\quad - (\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1}) \mathrm{vec}(\boldsymbol{z}_t \boldsymbol{z}'_t)\}.
\end{aligned}
$$

The expressions for $\partial \mathrm{vec}(\boldsymbol{P}_t)/\partial \boldsymbol{\rho}_i$, $i = 1, 2$, $\partial \mathrm{vec}(\boldsymbol{P}_t)/\partial \gamma$ and $\partial \mathrm{vec}(\boldsymbol{P}_t)/\partial c$ follow immediately from the definition of $\boldsymbol{P}_t$ in (5). The elements of $\partial \mathrm{vec}(\boldsymbol{P}_t)/\partial \boldsymbol{\rho}_i$ are either zeroes or equal G_t or $1 - G_t$. ∎

Proof of Theorem 1: In order to prove consistency, we verify the conditions of the following general result:

Theorem (Newey and McFadden, 1994, Theorem 2.5): Suppose that $\boldsymbol{\zeta}_t$, $t = 1, 2, ...$ are iid with p.d.f $f(\boldsymbol{\zeta}_t | \boldsymbol{\theta}_1^0)$ and that

(i) if $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^0$ then $f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1) \neq f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1^0)$;

(ii) $\boldsymbol{\theta}_1^0 \in \Theta_1$ is an interior point of Θ_1, which is compact;

(iii) $\ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)$ is continuous at each $\boldsymbol{\theta}_1 \in \Theta_1$ with probability 1;

(iv) $\mathsf{E}\sup_{\boldsymbol{\theta}_1 \in \Theta_1} |\ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)| < \infty$.

Proof: (i) This is an uniqueness condition. When $|\delta_{ij}| = 0$ or $\gamma_{ij} = 0$ for at least one $j = 1, ..., r_i$ and $i = 1, ..., N$, the model is not identified and (i) does not hold. Assumptions AG3 and AG4 exclude this case. The assumption $c_{i1} < ... < c_{ir_i}$, $j = 1, ..., r_i$, $i = 1, ..., N$, in AG4 determines the order of transition functions and is thus also required for identification (excludes exchangeability of transitions). The requirement in AG3 that for any pair (i, m), $i \neq m$, there exists at least one time-point t such that $g_{it} \neq g_{mt}$ excludes duplicate transitions. The correlation matrix is not identified if $\boldsymbol{P}_{(1)} = \boldsymbol{P}_{(2)}$ or $\gamma > 0$. AG5 excludes these possibilities. These conditions are required for uniqueness assumed in AG2. Note, however, that when $h_{it} \equiv 1$, $\delta_{i0} > 0$ is a free parameter, $i = 1, ..., N$.

(ii) This is satisfied due to Assumption AG1.

(iii) This follows from Lemma 1.

In order to show (iv), notice that $\boldsymbol{z}_t = (\boldsymbol{S}_t^0)^{-1}\boldsymbol{\varepsilon}_t$, where $\{\boldsymbol{z}_t\}$ is a sequence of independent random variables with $\mathsf{E}\boldsymbol{z}_t\boldsymbol{z}_t' = \boldsymbol{P}_t^0$, and

$$\boldsymbol{S}_t^0 = \text{diag}\{g^{1/2}(\boldsymbol{\theta}_{g1}^0; t/T), ..., g^{1/2}(\boldsymbol{\theta}_{gN}^0; t/T))\}. \tag{37}$$

In (37), $\boldsymbol{\theta}_{gi}^0$ is the true parameter vector for the ith diagonal element of $\boldsymbol{S}_t^0$. Likewise, $\boldsymbol{P}_t^0$ is the true correlation matrix and $(\boldsymbol{P}_t^0)^{1/2} = \boldsymbol{C}_t^0(\boldsymbol{\Lambda}_t^0)^{1/2}(\boldsymbol{C}_t^0)'$ is positive definite as $\boldsymbol{P}_t^0 = \boldsymbol{C}_t^0\boldsymbol{\Lambda}_t^0(\boldsymbol{C}_t^0)'$ is positive definite. Then

$$
\begin{aligned}
f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1) &= (1/2)\sum_{i=1}^{N} \ln g_{it}(\boldsymbol{\theta}_{gi}; t/T) - (1/2)\ln|\boldsymbol{P}_t| \\
&\quad -(1/2)\boldsymbol{z}_t'\boldsymbol{S}_t^0\boldsymbol{S}_t^{-1}\boldsymbol{P}_t^{-1}\boldsymbol{S}_t^{-1}\boldsymbol{S}_t^0\boldsymbol{z}_t.
\end{aligned}
$$

The absolute log density for observation t equals

$$
\begin{aligned}
|f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)| &= |(1/2)\sum_{i=1}^{N} \ln g(\boldsymbol{\theta}_{gi}; t/T) - (1/2)\ln|\boldsymbol{P}_t^{-1}| \\
&\quad -(1/2)\boldsymbol{z}_t'\boldsymbol{S}_t^0\boldsymbol{S}_t^{-1}\boldsymbol{P}_t^{-1}\boldsymbol{S}_t^{-1}\boldsymbol{S}_t^0\boldsymbol{z}_t|.
\end{aligned} \tag{38}
$$

Applying the triangle inequality to (38) leads to

$$
\begin{aligned}
|f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)| &\leq (1/2)\sum_{i=1}^{N} |\ln g(\boldsymbol{\theta}_{gi}; t/T)| + (1/2)\sum_{i=1}^{N} |\ln \lambda_{it}^{-1}(\boldsymbol{\theta}_\rho)| \\
&\quad +(1/2)\boldsymbol{z}_t'\boldsymbol{S}_t^0\boldsymbol{S}_t^{-1}\boldsymbol{P}_t^{-1}\boldsymbol{S}_t^{-1}\boldsymbol{S}_t^0\boldsymbol{z}_t \\
&\leq C_1 + (1/2)\sup_{\boldsymbol{\theta}_1 \in \Theta_1} \text{tr}(\boldsymbol{S}_t^{-1}\boldsymbol{P}_t^{-1}\boldsymbol{S}_t^{-1}\boldsymbol{S}_t^0\boldsymbol{z}_t\boldsymbol{z}_t'\boldsymbol{S}_t^0)
\end{aligned}
$$

where $\lambda_{it}(\boldsymbol{\theta}_\rho) > 0$ is the ith eigenvalue of $\boldsymbol{P}_t$, $i = 1, ..., N$, and C_1 is a generic positive constant. Then

$$
\begin{aligned}
&\mathsf{E}\sup_{\boldsymbol{\theta}_1 \in \Theta_1} \text{tr}(\boldsymbol{S}_t^{-1}\boldsymbol{P}_t^{-1}\boldsymbol{S}_t^{-1}\boldsymbol{S}_t^0\boldsymbol{z}_t\boldsymbol{z}_t'\boldsymbol{S}_t^0) \\
&= \sup_{\boldsymbol{\theta}_1 \in \Theta_1} \text{tr}(\boldsymbol{S}_t^{-1}\boldsymbol{P}_t^{-1}\boldsymbol{S}_t^{-1}\boldsymbol{S}_t^0\boldsymbol{P}_t^0\boldsymbol{S}_t^0) \leq C_2 < \infty
\end{aligned} \tag{39}
$$

where C_2 is a generic positive constant, because the matrices in (39) have finite elements for all $\boldsymbol{\theta}_1 \in \Theta_1$. Thus, for all $\boldsymbol{\theta}_1 \in \Theta_1$,

$$\mathsf{E}|f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)| \leq C_1 + C_2 < \infty$$

which verifies (iv) and concludes the proof. $\blacksquare$

Proof of Lemma 2: First define

$$\widehat{\boldsymbol{B}}_T(\boldsymbol{\theta}_1^0) = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{s}_t(\boldsymbol{\theta}_1^0)\boldsymbol{s}_t'(\boldsymbol{\theta}_1^0).$$

which is positive definite for $T > \sum_{i=1}^{N} r_i + N(N+2) + 2$, the dimension of $\boldsymbol{\theta}_1^0$. Using the law of large numbers and assuming that AG5 and AG6 hold,

$$\widehat{\boldsymbol{B}}_T(\boldsymbol{\theta}_1^0) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathsf{E}\boldsymbol{s}_t(\boldsymbol{\theta}_1^0)\boldsymbol{s}_t'(\boldsymbol{\theta}_1^0) + \boldsymbol{o}_p(1).$$

Due to the fact that the transition variable is (rescaled) time, we apply triangular array asymptotics, see for example Hillebrand, Medeiros and Xu (2013), although this is not emphasised in the proof. Let $[Tr] = t$, where $0 < r \leq 1$. It follows that

$$
\begin{aligned}
\widehat{\boldsymbol{B}}_T(\boldsymbol{\theta}_1^0) &= \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{s}_t(\boldsymbol{\theta}_1^0)\boldsymbol{s}_t'(\boldsymbol{\theta}_1^0) = \sum_{t=1}^{T} \int_{t/T}^{(t+1)/T} \boldsymbol{s}_{[Tr]}(\boldsymbol{\theta}_1^0)\boldsymbol{s}_{[Tr]}'(\boldsymbol{\theta}_1^0)\mathrm{d}r \\
&= \int_{1/T}^{(T+1)/T} \boldsymbol{s}_{[Tr]}(\boldsymbol{\theta}_1^0)\boldsymbol{s}_{[Tr]}'(\boldsymbol{\theta}_1^0)\mathrm{d}r \xrightarrow{p} \int_{0}^{1} \mathsf{E}\boldsymbol{s}_r(\boldsymbol{\theta}_1^0)\boldsymbol{s}_r'(\boldsymbol{\theta}_1^0)\mathrm{d}r \\
&= \boldsymbol{B}(\boldsymbol{\theta}_1^0)
\end{aligned}
$$

as $T \to \infty$. $\boldsymbol{B}(\boldsymbol{\theta}_1^0)$ is positive definite because $\widehat{\boldsymbol{B}}_T(\boldsymbol{\theta}_1^0)$ is positive definite for $T > \sum_{i=1}^{N} r_i + N(N+2) + 2$. $\blacksquare$

Proof of Lemma 3: From Lemma 1 it follows that

$$\boldsymbol{s}_t(\boldsymbol{\theta}_{gi}^0) = \frac{1}{2g_{it}^0} \frac{\partial g_{it}^0}{\partial \boldsymbol{\theta}_{gi}} (\boldsymbol{z}_t'(\boldsymbol{P}_t^0)^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\boldsymbol{z}_t - 1).$$

where $\boldsymbol{z}_t = (\boldsymbol{P}_t^0)^{1/2}\boldsymbol{\zeta}_t$. Write

$$\boldsymbol{e}_i'\boldsymbol{z}_t\boldsymbol{z}_t'(\boldsymbol{P}_t^0)^{-1}\boldsymbol{e}_i = (\boldsymbol{e}_i \otimes \boldsymbol{e}_i)'(\boldsymbol{I}_N \otimes (\boldsymbol{P}_t^0)^{-1})\mathrm{vec}(\boldsymbol{z}_t\boldsymbol{z}_t').$$

Let $i \neq j$ and consider

$$
\begin{aligned}
&(\boldsymbol{e}_i'\boldsymbol{z}_t\boldsymbol{z}_t'(\boldsymbol{P}_t^0)^{-1}\boldsymbol{e}_i - 1)(\boldsymbol{e}_j'\boldsymbol{z}_t\boldsymbol{z}_t'(\boldsymbol{P}_t^0)^{-1}\boldsymbol{e}_j - 1) \\
=\ &(\boldsymbol{e}_i \otimes \boldsymbol{e}_i)'(\boldsymbol{I}_N \otimes (\boldsymbol{P}_t^0)^{-1})\mathrm{vec}(\boldsymbol{z}_t\boldsymbol{z}_t')\mathrm{vec}(\boldsymbol{z}_t\boldsymbol{z}_t')'(\boldsymbol{I}_N \otimes (\boldsymbol{P}_t^0)^{-1})(\boldsymbol{e}_j \otimes \boldsymbol{e}_j) \\
&-\boldsymbol{e}_i'\boldsymbol{z}_t\boldsymbol{z}_t'(\boldsymbol{P}_t^0)^{-1}\boldsymbol{e}_i - \boldsymbol{e}_j'\boldsymbol{z}_t\boldsymbol{z}_t'(\boldsymbol{P}_t^0)^{-1}\boldsymbol{e}_j + 1.
\end{aligned}
$$

Since

$$
\begin{aligned}
&\mathsf{E}\mathrm{vec}(\boldsymbol{z}_t\boldsymbol{z}_t')\mathrm{vec}(\boldsymbol{z}_t\boldsymbol{z}_t')' \\
=\ &(\boldsymbol{P}_t^0 \otimes \boldsymbol{P}_t^0) + (\boldsymbol{I}_N \otimes \boldsymbol{P}_t^0)\boldsymbol{K}(\boldsymbol{I}_N \otimes \boldsymbol{P}_t^0) + \mathrm{vec}(\boldsymbol{P}_t^0)\mathrm{vec}(\boldsymbol{P}_t^0)' \quad\quad (40)
\end{aligned}
$$

31

see Anderson (2003, p. 64), and

$$\mathsf{E}e_i' z_t z_t' (P_t^0)^{-1} e_i = 1$$

one obtains

$$
\begin{aligned}
& \mathsf{E}(e_i' z_t z_t' (P_t^0)^{-1} e_i - 1)(e_j' z_t z_t' (P_t^0)^{-1} e_j - 1) \\
= \ & e_i' P_t^0 e_j e_i' (P_t^0)^{-1} e_j.
\end{aligned}
$$

and

$$\mathsf{E}s_t(\boldsymbol{\theta}_{gi}^0) s_t'(\boldsymbol{\theta}_{gj}^0) = \frac{1}{4 g_{it}^0 g_{jt}^0} \frac{\partial g_{it}^0}{\partial \boldsymbol{\theta}_{gi}} \frac{\partial g_{it}^0}{\partial \boldsymbol{\theta}_{gj}'} e_i' P_t^0 e_j e_i' (P_t^0)^{-1} e_j. \tag{41}$$

The expectation (41) exists due to AG6. Similarly, for $i = j$,

$$\mathsf{E}s_t(\boldsymbol{\theta}_{gi}^0) s_t'(\boldsymbol{\theta}_{gi}^0) = \frac{1}{4 (g_{it}^0)^2} \frac{\partial g_{it}^0}{\partial \boldsymbol{\theta}_{gi}} \frac{\partial g_{it}^0}{\partial \boldsymbol{\theta}_{gi}'} (1 + e_i' (P_t^0)^{-1} e_i).$$

Next, consider

$$
\begin{aligned}
\mathsf{E}s_t(\boldsymbol{\theta}_\rho^0) s_t'(\boldsymbol{\theta}_\rho^0) \ = \ & \mathsf{E}\frac{1}{4} \frac{\partial \mathrm{vec}(P_t^0)'}{\partial \boldsymbol{\theta}_\rho} [\mathrm{vec}((P_t^0)^{-1}) - \{(P_t^0)^{-1} \otimes (P_t^0)^{-1}\}(z_t \otimes z_t)] \\
& \times [\mathrm{vec}((P_t^0)^{-1}) - \{(P_t^0)^{-1} \otimes (P_t^0)^{-1}\}(z_t \otimes z_t)]' \\
& \times \frac{\partial \mathrm{vec}(P_t^0)}{\partial \boldsymbol{\theta}_\rho'}.
\end{aligned}
\tag{42}
$$

Since

$$\mathsf{E}(z_t \otimes z_t) = \mathsf{E}\mathrm{vec}(z_t z_t') = \mathrm{vec}(P_t^0)$$

one obtains

$$
\begin{aligned}
& \{(P_t^0)^{-1} \otimes (P_t^0)^{-1}\} \mathsf{E}\mathrm{vec}(z_t z_t') \mathrm{vec}((P_t^0)^{-1})' \\
= \ & \mathrm{vec}((P_t^0)^{-1}) \mathrm{vec}((P_t^0)^{-1})'.
\end{aligned}
\tag{43}
$$

From (40) it follows that

$$
\begin{aligned}
& \mathsf{E}\{(P_t^0)^{-1} \otimes (P_t^0)^{-1}\}(z_t \otimes z_t')(z_t \otimes z_t')'\{(P_t^0)^{-1} \otimes (P_t^0)^{-1}\} \\
= \ & (P_t^0)^{-1} \otimes (P_t^0)^{-1} + ((P_t^0)^{-1} \otimes I_N) K((P_t^0)^{-1} \otimes I_N) \\
& + \mathrm{vec}((P_t^0)^{-1}) \mathrm{vec}((P_t^0)^{-1})'.
\end{aligned}
\tag{44}
$$

Applying (43) and (44) to (42) yields

$$
\begin{aligned}
& \mathsf{E}s_t(\boldsymbol{\theta}_\rho^0) s_t'(\boldsymbol{\theta}_\rho^0) \\
= \ & \frac{1}{4} \frac{\partial \mathrm{vec}(P_t^0)'}{\partial \boldsymbol{\theta}_\rho} \{(P_t^0)^{-1} \otimes (P_t^0)^{-1} + ((P_t^0)^{-1} \otimes I_N) K((P_t^0)^{-1} \otimes I_N)\} \\
& \times \frac{\partial \mathrm{vec}(P_t^0)}{\partial \boldsymbol{\theta}_\rho'}
\end{aligned}
$$

Finally, consider

$$
\begin{aligned}
&\mathsf{E}\boldsymbol{s}_t(\boldsymbol{\theta}_\rho^0)\boldsymbol{s}_t'(\boldsymbol{\theta}_{gi}^0)\\
&=\ \frac{1}{4g_{it}^0}\frac{\partial\mathrm{vec}(\boldsymbol{P}_t^0)'}{\partial\boldsymbol{\theta}_\rho}\mathsf{E}\{\mathrm{vec}(\boldsymbol{P}_t^0)^{-1}-((\boldsymbol{P}_t^0)^{-1}\otimes(\boldsymbol{P}_t^0)^{-1})(\boldsymbol{z}_t\otimes\boldsymbol{z}_t)\}\\
&\qquad\times\{1-\boldsymbol{e}_i'\boldsymbol{z}_t\boldsymbol{z}_t'(\boldsymbol{P}_t^0)^{-1}\boldsymbol{e}_i\}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}'}\\
&=\ \frac{1}{4g_{it}^0}\frac{\partial\mathrm{vec}(\boldsymbol{P}_t^0)'}{\partial\boldsymbol{\theta}_\rho}\{\mathrm{vec}(\boldsymbol{P}_t^0)^{-1}(1-\boldsymbol{e}_i'\mathsf{E}\boldsymbol{z}_t\boldsymbol{z}_t'(\boldsymbol{P}_t^0)^{-1}\boldsymbol{e}_i)\\
&\qquad-((\boldsymbol{P}_t^0)^{-1}\otimes(\boldsymbol{P}_t^0)^{-1})\left(\mathsf{E}(\boldsymbol{z}_t\otimes\boldsymbol{z}_t)-\mathsf{E}(\boldsymbol{z}_t\otimes\boldsymbol{z}_t)\boldsymbol{e}_i'\boldsymbol{z}_t\boldsymbol{z}_t'(\boldsymbol{P}_t^0)^{-1}\boldsymbol{e}_i\right)\}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}'}\\
&=\ \frac{1}{4g_{it}^0}\frac{\partial\mathrm{vec}(\boldsymbol{P}_t^0)'}{\partial\boldsymbol{\theta}_\rho}\{((\boldsymbol{P}_t^0)^{-1}\otimes(\boldsymbol{P}_t^0)^{-1})\mathsf{E}(\boldsymbol{z}_t\boldsymbol{z}_t'\otimes\boldsymbol{z}_t\boldsymbol{z}_t')(\boldsymbol{e}_i\otimes(\boldsymbol{P}_t^0)^{-1}\boldsymbol{e}_i)\\
&\qquad-((\boldsymbol{P}_t^0)^{-1}\otimes(\boldsymbol{P}_t^0)^{-1})\mathsf{E}(\boldsymbol{z}_t\otimes\boldsymbol{z}_t)\}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}'}. \qquad(45)
\end{aligned}
$$

Application of (40) and (44) to (45) and some matrix manipulation give

$$
\begin{aligned}
&\mathsf{E}\boldsymbol{s}_t(\boldsymbol{\theta}_\rho)\boldsymbol{s}_t'(\boldsymbol{\theta}_{gi}^0)\\
&=\ \frac{1}{4g_{it}^0}\frac{\partial\mathrm{vec}(\boldsymbol{P}_t^0)'}{\partial\boldsymbol{\theta}_\rho}\{(\boldsymbol{P}_t^0)^{-1}\otimes\boldsymbol{I}_N+\boldsymbol{I}_N\otimes(\boldsymbol{P}_t^0)^{-1}\}(\boldsymbol{e}_i\otimes\boldsymbol{e}_i)\frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}'}
\end{aligned}
$$

which is the desired result. ∎

Proof of Lemma 4: As before, let

$$
\boldsymbol{z}_t=\boldsymbol{S}_t^{-1}\boldsymbol{\varepsilon}_t=(\frac{\varepsilon_{1t}}{g_{1t}^{1/2}},...,\frac{\varepsilon_{Nt}}{g_{Nt}^{1/2}})'.
$$

where $\boldsymbol{\varepsilon}_t=\boldsymbol{S}_t^0(\boldsymbol{P}_t^0)^{1/2}\boldsymbol{\zeta}_t$. The matrix of second derivatives of $\ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)$ with respect to the parameters in $\boldsymbol{S}_t$ are as follows:

$$
\begin{aligned}
\frac{\partial^2\ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}_{gj}\partial\boldsymbol{\theta}_{gi}'}\ =\ &\frac{1}{2}\frac{\partial}{\partial\boldsymbol{\theta}_{gj}}\frac{\partial\ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}_{gi}'}=\frac{1}{2}(\frac{\partial}{\partial\boldsymbol{\theta}_{gj}}\frac{1}{g_{it}})\frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}'}(\boldsymbol{e}_i'\boldsymbol{P}_t^{-1}\boldsymbol{z}_t\boldsymbol{z}_t'\boldsymbol{e}_i-1)\\
&+\frac{1}{2g_{it}}\frac{\partial^2 g_{it}}{\partial\boldsymbol{\theta}_{gj}\partial\boldsymbol{\theta}_{gi}'}(\boldsymbol{e}_i'\boldsymbol{P}_t^{-1}\boldsymbol{z}_t\boldsymbol{z}_t'\boldsymbol{e}_i-1)\\
&+\frac{1}{2g_{it}}\frac{\partial}{\partial\boldsymbol{\theta}_{gj}}(\boldsymbol{z}_t'\boldsymbol{e}_i\boldsymbol{e}_i'\boldsymbol{P}_t^{-1}\boldsymbol{z}_t-1)\frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}'} \qquad(46)
\end{aligned}
$$

when $i\neq j$. The first term on the r.h.s of (46) vanishes because $\partial g_{it}^{-1}/\partial\boldsymbol{\theta}_{gj}=\boldsymbol{0}_{3r_j+1}$

and the second one because $\partial^2 g_{it}^{-1}/(\partial\boldsymbol{\theta}_{gj}\partial\boldsymbol{\theta}'_{gi}) = \mathbf{0}_{3r_j+1\times 3r_i+1}$. Consider

$$
\begin{aligned}
\frac{\partial \boldsymbol{z}'_t \boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} \boldsymbol{z}_t}{\partial\boldsymbol{\theta}_{gj}} &= \frac{\partial \boldsymbol{z}_t}{\partial\boldsymbol{\theta}_{gj}} \frac{\partial \boldsymbol{z}'_t \boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} \boldsymbol{z}_t}{\partial \boldsymbol{z}_t} \\
&= \frac{\partial \boldsymbol{z}_t}{\partial\boldsymbol{\theta}_{gj}} (\boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} + \boldsymbol{P}_t^{-1} \boldsymbol{e}_i \boldsymbol{e}'_i) \boldsymbol{z}_t \\
&= -\frac{z_{jt}}{2g_{jt}} \frac{\partial g_{jt}}{\partial\boldsymbol{\theta}_{gj}} \boldsymbol{e}'_j (\boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} + \boldsymbol{P}_t^{-1} \boldsymbol{e}_i \boldsymbol{e}'_i) \boldsymbol{z}_t \\
&= -\frac{1}{2g_{jt}} \frac{\partial g_{jt}}{\partial\boldsymbol{\theta}_{gj}} \boldsymbol{e}'_j \boldsymbol{P}_t^{-1} \boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{z}_t \boldsymbol{z}'_t \boldsymbol{e}_j.
\end{aligned}
$$

Then

$$
\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}_{gj}\partial\boldsymbol{\theta}'_{gi}} = -\frac{1}{4g_{jt}g_{it}} \frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gj}} \frac{\partial g_{it}}{\partial\boldsymbol{\theta}'_{gi}} \boldsymbol{e}'_j \boldsymbol{P}_t^{-1} \boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{z}_t \boldsymbol{z}'_t \boldsymbol{e}_j.
$$

Next assume $i = j$. Then the first term on the r.h.s. of (46) becomes

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

Now consider

$$
\begin{aligned}
\frac{\partial \boldsymbol{z}'_t \boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} \boldsymbol{z}_t}{\partial\boldsymbol{\theta}_{gi}} &= \frac{\partial \boldsymbol{z}_t}{\partial\boldsymbol{\theta}_{gi}} \frac{\partial \boldsymbol{z}'_t \boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} \boldsymbol{z}_t}{\partial \boldsymbol{z}_t} \\
&= -\frac{z_{it}}{2g_{it}} \frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}} \boldsymbol{e}'_i (\boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} + \boldsymbol{P}_t^{-1} \boldsymbol{e}_i \boldsymbol{e}'_i) \boldsymbol{z}_t \\
&= -\frac{1}{2g_{it}} \frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}} \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} (\boldsymbol{I}_N + \boldsymbol{e}_i \boldsymbol{e}'_i) \boldsymbol{z}_t \boldsymbol{z}'_t \boldsymbol{e}_i.
\end{aligned}
$$

Finally,

$$
\begin{aligned}
\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}_{gi}\partial\boldsymbol{\theta}'_{gi}} &= -\left\{\frac{1}{2g_{it}^2} \frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}} \frac{\partial g_{it}}{\partial\boldsymbol{\theta}'_{gi}} - \frac{1}{2g_{it}} \frac{\partial^2 g_{it}}{\partial\boldsymbol{\theta}_{gi}\partial\boldsymbol{\theta}'_{gi}}\right\}(\boldsymbol{e}'_i \boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}'_t \boldsymbol{e}_i - 1) \\
&\quad -\frac{1}{4g_{it}^2} \frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}} \frac{\partial g_{it}}{\partial\boldsymbol{\theta}'_{gi}} \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} (\boldsymbol{I}_N + \boldsymbol{e}_i \boldsymbol{e}'_i) \boldsymbol{z}_t \boldsymbol{z}'_t \boldsymbol{e}_i.
\end{aligned}
$$

This is the desired result. The expressions for the elements of (17) can be found in Amado and Teräsvirta (2013, Lemma A.2). $\blacksquare$

In proving Lemma 5, we use the following result:

Lemma 9 *(Eklund and Teräsvirta, 2007):*

$$
\frac{\partial}{\partial\boldsymbol{\rho}}\left\{\frac{\partial vec(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}'} vec(\boldsymbol{P}_t^{-1})\right\} = -\frac{\partial vec(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}'}(\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1})\frac{\partial vec(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}} \tag{47}
$$

where $\boldsymbol{\rho} = (\boldsymbol{\rho}'_1, \boldsymbol{\rho}'_2)'$ (so the matrix (47) is $N(N-1) \times N(N-1)$), and

$$
\begin{aligned}
&\frac{\partial}{\partial\boldsymbol{\rho}}\left\{\frac{\partial vec(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}'} vec(\boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}'_t \boldsymbol{P}_t^{-1})\right\} \\
&= -\frac{\partial vec(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}'}(\boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}'_t \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} + \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}'_t \boldsymbol{P}_t^{-1})\frac{\partial vec(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}}. \tag{48}
\end{aligned}
$$

Proof of Lemma 5: Consider first $\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)/(\partial\boldsymbol{\theta}'_{gi}\partial\boldsymbol{\rho})$ and write

$$\frac{\partial}{\partial\boldsymbol{\rho}}\frac{\partial \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}'_{gi}} = \frac{1}{2g_{it}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}'_{gi}}\frac{\partial}{\partial\boldsymbol{\rho}}(z_{it}\boldsymbol{e}'_i\boldsymbol{P}_t^{-1}\boldsymbol{z}_t - 1).$$

where

$$z_{it}\boldsymbol{e}'_i\boldsymbol{P}_t^{-1}\boldsymbol{z}_t = z_{it}\mathrm{vec}(\boldsymbol{e}'_i\boldsymbol{P}_t^{-1}\boldsymbol{z}_t) = z_{it}(\boldsymbol{z}'_t \otimes \boldsymbol{e}'_i)\mathrm{vec}(\boldsymbol{P}_t^{-1}).$$

Differentiation gives

$$z_{it}\frac{\partial}{\partial\boldsymbol{\rho}}(\boldsymbol{z}'_t \otimes \boldsymbol{e}'_i)\mathrm{vec}(\boldsymbol{P}_t^{-1})$$

$$= z_{it}(\boldsymbol{z}'_t \otimes \boldsymbol{e}'_i)\frac{\mathrm{vec}(\boldsymbol{P}_t^{-1})}{\partial\boldsymbol{\rho}} = z_{it}(\boldsymbol{z}'_t \otimes \boldsymbol{e}'_i)\frac{\partial\mathrm{vec}(\boldsymbol{P}_t^{-1})}{\partial\mathrm{vec}(\boldsymbol{P}_t)}\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}}. \qquad (49)$$

Applying (30) to (49) yields

$$z_{it}\frac{\partial}{\partial\boldsymbol{\rho}}(\boldsymbol{z}'_t \otimes \boldsymbol{e}'_i)\mathrm{vec}(\boldsymbol{P}_t^{-1}) = -z_{it}(\boldsymbol{z}'_t \otimes \boldsymbol{e}'_i)(\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1})\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}}$$

$$= -(\boldsymbol{e}'_i\boldsymbol{z}_t\boldsymbol{z}'_t \otimes \boldsymbol{e}'_i)(\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1})\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}}$$

so

$$\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}_{gi}\partial\boldsymbol{\rho}} = -\frac{1}{2g_{it}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}'_{gi}}(\boldsymbol{e}'_i\boldsymbol{z}_t\boldsymbol{z}'_t \otimes \boldsymbol{e}'_i)(\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1})\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}}$$

$$= -\frac{1}{2g_{it}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}'_{gi}}(\boldsymbol{e}_i \otimes \boldsymbol{e}_i)'(\boldsymbol{z}_t\boldsymbol{z}'_t\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1})\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}}.$$

This takes care of the first part of the lemma. To prove the second part, note that the relevant block of the score equals

$$\frac{\partial \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial\boldsymbol{\rho}} = \boldsymbol{s}_t(\boldsymbol{\rho}) = -(1/2)\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}'}\{\mathrm{vec}(\boldsymbol{P}_t^{-1}) - \mathrm{vec}(\boldsymbol{P}_t^{-1}\boldsymbol{z}_t\boldsymbol{z}'_t\boldsymbol{P}_t^{-1})\}$$

so its derivative may be written as

$$\frac{\boldsymbol{s}_t(\boldsymbol{\rho})}{\partial\boldsymbol{\rho}} = -(1/2)\frac{\partial}{\partial\boldsymbol{\rho}}[\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}'}\{\mathrm{vec}(\boldsymbol{P}_t^{-1}) - \mathrm{vec}(\boldsymbol{P}_t^{-1}\boldsymbol{z}_t\boldsymbol{z}'_t\boldsymbol{P}_t^{-1})\}]. \qquad (50)$$

The elements of $\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}'}$ in (50) are either zeroes or ones multiplied by G_t or $1 - G_t$. Then, with slight abuse of notation, $\partial^2\mathrm{vec}(\boldsymbol{P}_t)/(\partial\boldsymbol{\rho}'\partial\boldsymbol{\rho}) = \boldsymbol{0}$. Applying Lemma 9 to (50) gives

$$\frac{\partial\boldsymbol{s}_t(\boldsymbol{\rho})}{\partial\boldsymbol{\rho}} = (1/2)\{\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}'}(\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1})\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}}$$

$$-\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}'}(\boldsymbol{P}_t^{-1}\boldsymbol{z}_t\boldsymbol{z}'_t\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} + \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1}\boldsymbol{z}_t\boldsymbol{z}'_t\boldsymbol{P}_t^{-1})\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\rho}}\}$$

which is the desired result. ∎

Proof of Lemma 6: The second derivative of (7) with respect to γ can be written as follows:

$$
\begin{aligned}
\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \gamma^2} &= \frac{\partial}{\partial \gamma} \frac{\partial \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \gamma} \\
&= -\frac{1}{2} \frac{\partial}{\partial \gamma} \left(\frac{\partial \mathrm{vec}(\boldsymbol{P}_t)'}{\partial \gamma} \{ \mathrm{vec}(\boldsymbol{P}_t^{-1}) - \mathrm{vec}(\boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1}) \} \right. \\
&= -\frac{1}{2} \frac{\partial^2 \mathrm{vec}(\boldsymbol{P}_t)'}{\partial \gamma^2} \{ \mathrm{vec}(\boldsymbol{P}_t^{-1}) - \mathrm{vec}(\boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1}) \} \\
&\quad -\frac{1}{2} \left(\frac{\partial \mathrm{vec}(\boldsymbol{P}_t)'}{\partial \gamma} \frac{\partial}{\partial \gamma} \{ \mathrm{vec}(\boldsymbol{P}_t^{-1}) - \mathrm{vec}(\boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1}) \}. \quad (51)
\end{aligned}
$$

Applying Lemma 9 to (51) yields

$$
\begin{aligned}
\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \gamma^2} &= -\frac{1}{2} \frac{\partial^2 \mathrm{vec}(\boldsymbol{P}_t)'}{\partial \gamma^2} \{ \mathrm{vec}(\boldsymbol{P}_t^{-1}) - \mathrm{vec}(\boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1}) \} \\
&\quad -\frac{1}{2} \frac{\partial \mathrm{vec}(\boldsymbol{P}_t)'}{\partial \gamma} \{ \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} \\
&\quad -(\boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} + \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1}) \} \frac{\partial \mathrm{vec}(\boldsymbol{P}_t)}{\partial \gamma}.
\end{aligned}
$$

Expressions for the derivatives $\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta})/\partial c^2$ and $\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta})/(\partial \gamma \partial c)$ are obtained in a similar fashion. The expressions for (21), (22) and (23) in (20) can be found by straightforward calculation. ∎

Proof of Theorem 2: To prove this result, we verify the conditions in the following theorem.

Theorem (Newey and McFadden, Theorem 3.3). Suppose that $\boldsymbol{\zeta}_t$, $t = 1, 2, ...$ are iid with p.d.f $f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1^0)$, the assumptions of the consistency theorem are satisfied, and that

(i) $\boldsymbol{\theta}_1^0$ is an interior point of Θ_1.

(ii) $f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)$ is twice continuously differentiable in a neighbourhood $\mathcal{N}_1$ of $\boldsymbol{\theta}_1^0$.

(iii) $\int \sup_{\boldsymbol{\theta}_1 \in \mathcal{N}_1} ||\frac{\partial}{\partial \boldsymbol{\theta}_1} f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)|| \mathrm{d}\boldsymbol{\zeta} < \infty$ and $\int \sup_{\boldsymbol{\theta}_1 \in \mathcal{N}_1} ||\frac{\partial^2}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'} f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)|| \mathrm{d}\boldsymbol{\zeta} < \infty$, where $|| \cdot ||$ is the Euclidean norm.

(iv) Matrix $\boldsymbol{B}(\boldsymbol{\theta}_1^0) = \mathsf{E}\frac{\partial}{\partial \boldsymbol{\theta}_1} \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1) \frac{\partial}{\partial \boldsymbol{\theta}_1'} \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)$ exists and is nonsingular.

(v) $\mathsf{E} \sup_{\boldsymbol{\theta}_1 \in \mathcal{N}_1} ||\frac{\partial^2}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'} \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)|| < \infty$.

Then $\sqrt{T}(\widehat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^0) \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{B}(\boldsymbol{\theta}_1^0)^{-1})$ as $T \to \infty$.

Condition (i) is satisfied by AG2. Condition (ii) holds due to Lemmas 1, 4, 5 and 6. Condition (iii) is satisfied due to AN1. Validity of (iv) follows from Lemma 2. To validate Condition (v), define

$$
||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'}|| = || \begin{bmatrix} \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_g \partial \boldsymbol{\theta}_g'} & \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_g \partial \boldsymbol{\theta}_\rho'} \\ \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_\rho \partial \boldsymbol{\theta}_g'} & \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_\rho \partial \boldsymbol{\theta}_\rho'} \end{bmatrix} ||
$$

where

$$
\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_g \partial \boldsymbol{\theta}_g'} = \mathrm{diag}(\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{g1} \partial \boldsymbol{\theta}_{g1}'}, ..., \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gN} \partial \boldsymbol{\theta}_{gN}'})
$$

and

$$\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_g \partial \boldsymbol{\theta}'_\rho} = (\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{g1} \partial \boldsymbol{\theta}'_\rho}, ..., \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gN} \partial \boldsymbol{\theta}'_\rho}).$$

Then, by definition,

$$\begin{aligned}
||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}'_1}|| &= \{\sum_{i=1}^{N} ||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_{gi}}||^2 + 2 \sum_{i=1}^{N} ||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_\rho}||^2 \\
&\quad + ||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_\rho \partial \boldsymbol{\theta}'_\rho}||^2\}^{1/2}
\end{aligned}$$

It follows, by triangular inequality, that

$$\begin{aligned}
||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}'_1}|| &\leq \sum_{i=1}^{N} ||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_{gi}}|| + 2 \sum_{i=1}^{N} ||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_\rho}|| \\
&\quad + ||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_\rho \partial \boldsymbol{\theta}'_\rho}||.
\end{aligned} \tag{52}$$

The terms in (52) will be studied one by one, beginning with terms in the first sum on the r.h.s. of (52). We state the following result:

Lemma 10 *The expectation of* $||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_{gi}}||$ *is finite for all t and $i = 1, ..., N$.*

Proof: Slight manipulation of the result in Lemma 4 gives

$$\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_{gi}} = -\frac{1}{g_{it}^2} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}'_{gi}} \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}'_t \boldsymbol{e}_i.$$

where g_{it} is positive and finite and the elements of

$$\frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}} = (1, G_{1t}, \frac{\partial G_{1t}}{\partial \gamma_1}, \frac{\partial G_{1t}}{\partial c_1}, ..., G_{r_1 t}, \frac{\partial G_{r_1 t}}{\partial \gamma_{r_1}}, \frac{\partial G_{r_1 t}}{\partial c_{r_1}})$$

are finite for all i, t, and $\boldsymbol{\theta}_{gi} \in \mathcal{N}_1$. Then

$$||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_{gi}}|| \leq C_1 |\boldsymbol{e}'_i \boldsymbol{P}_t^{-1} \boldsymbol{z}_t \boldsymbol{z}'_t \boldsymbol{e}_i|$$

where C_1 is a generic positive constant. $\boldsymbol{P}_t \to \boldsymbol{P}_t^0$ implies

$$\boldsymbol{e}'_i \boldsymbol{P}_t^{-1} \mathsf{E} \boldsymbol{z}_t \boldsymbol{z}'_t \boldsymbol{e}_i = \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} \boldsymbol{P}_t^0 \boldsymbol{e}_i \to 1$$

Then, for $\boldsymbol{\theta}_{gi} \in \mathcal{N}_1$, and for all i, t,

$$\mathsf{E}||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_{gi}}|| \leq C_1 |\boldsymbol{e}'_i \boldsymbol{P}_t^{-1} \boldsymbol{P}_t^0 \boldsymbol{e}_i| < \infty. \blacksquare$$

Lemma 11 *The expectation of* $||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_\rho}||$ *is finite for all t and $i = 1, ..., N$.*

Proof: Lemma 5 states that

$$\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_\rho} = -\frac{1}{2g_{it}} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}'_{gi}} (\boldsymbol{e}'_i \boldsymbol{z}_t \boldsymbol{z}'_t \otimes \boldsymbol{e}'_i)(\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1}) \frac{\partial \text{vec}(\boldsymbol{P}_t)}{\partial \boldsymbol{\theta}_\rho}. \tag{53}$$

Consider

$$\begin{aligned}
||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_\rho}|| &= \frac{1}{2|g_{it}|}||\frac{\partial g_{it}}{\partial \boldsymbol{\theta}'_{gi}}(\boldsymbol{e}'_i \boldsymbol{z}_t \boldsymbol{z}'_t \otimes \boldsymbol{e}'_i)(\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1}) \frac{\partial \text{vec}(\boldsymbol{P}_t)}{\partial \boldsymbol{\theta}_\rho}|| \\
&\leq \frac{1}{2|g_{it}|}||\frac{\partial g_{it}}{\partial \boldsymbol{\theta}'_{gi}}|| \times ||\boldsymbol{e}'_i \boldsymbol{z}_t \boldsymbol{z}'_t \otimes \boldsymbol{e}'_i|| \\
&\quad \times ||\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1}|| \times ||\frac{\partial \text{vec}(\boldsymbol{P}_t)}{\partial \boldsymbol{\theta}_\rho}|| \\
&\leq C_3 ||\boldsymbol{e}'_i \boldsymbol{z}_t \boldsymbol{z}'_t \otimes \boldsymbol{e}'_i||.
\end{aligned}$$

where C_3 is a generic positive constant, because also the elements of $\boldsymbol{P}_t^{-1}$ and $\frac{\partial \text{vec}(\boldsymbol{P}_t)}{\partial \boldsymbol{\theta}_\rho}$ are also finite for $\boldsymbol{\theta}_\rho \in \mathcal{N}_1$, and all t. Furthermore,

$$\mathsf{E}||\boldsymbol{e}'_i \boldsymbol{z}_t \boldsymbol{z}'_t \otimes \boldsymbol{e}'_i|| = ||\boldsymbol{e}'_i \boldsymbol{P}_t^0 \otimes \boldsymbol{e}'_i|| \leq C_4 < \infty$$

for $\boldsymbol{\theta}_\rho \in \mathcal{N}_1$, and all t, so

$$\mathsf{E}||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_\rho}|| = C_3 C_4 < \infty. \blacksquare$$

Lemma 12 *The expectation of* $||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_\rho \partial \boldsymbol{\theta}'_\rho}||$ *is finite for all* t.

Proof: The following block division is applied in proving the result:

$$\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_\rho \partial \boldsymbol{\theta}'_\rho} = \begin{bmatrix} \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \rho \partial \rho'} & \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \rho \partial \gamma} & \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \rho \partial c} \\ & \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \gamma^2} & \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \gamma \partial c} \\ & & \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial c^2} \end{bmatrix}. \tag{54}$$

The square of the Euclidean norm of (54) using these blocks equals

$$\begin{aligned}
||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_\rho \partial \boldsymbol{\theta}'_\rho}||^2 &= ||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \rho \partial \rho'}||^2 + 2||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \rho \partial \gamma}||^2 \\
&\quad + ||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \gamma^2}||^2 + ||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial c^2}||^2 \\
&\quad + 2||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \gamma \partial c}||^2
\end{aligned} \tag{55}$$

which implies

$$\begin{aligned}
||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_\rho \partial \boldsymbol{\theta}'_\rho}|| &\leq ||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \rho \partial \rho'}|| + \sqrt{2}||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \rho \partial \gamma}|| \\
&\quad + ||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \gamma^2}|| + ||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial c^2}|| \\
&\quad + \sqrt{2}||\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \gamma \partial c}||.
\end{aligned} \tag{56}$$

Consider the first term on the r.h.s. of (56). By Lemma 5 one can write

$$
\begin{aligned}
\left\|\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\rho} \partial \boldsymbol{\rho}'}\right\| &= \|\frac{1}{2}\{\frac{\partial \text{vec}(\boldsymbol{P}_t)}{\partial \boldsymbol{\rho}'}(\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1})\frac{\partial \text{vec}(\boldsymbol{P}_t)}{\partial \boldsymbol{\rho}}\| \\
&\quad + 2\|\frac{\partial \text{vec}(\boldsymbol{P}_t)}{\partial \boldsymbol{\rho}'}(\boldsymbol{P}_t^{-1}\boldsymbol{z}_t\boldsymbol{z}_t'\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1})\frac{\partial \text{vec}(\boldsymbol{P}_t)}{\partial \boldsymbol{\rho}}\| \\
&= C_5 + C_6\|\boldsymbol{P}_t^{-1}\boldsymbol{z}_t\boldsymbol{z}_t'\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1}\|
\end{aligned}
$$

where C_5 and C_6 are generic positive constants. Then

$$
\begin{aligned}
\mathsf{E}\|\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\rho} \partial \boldsymbol{\rho}'}\| &= C_5 + C_6\mathsf{E}\|\boldsymbol{P}_t^{-1}\boldsymbol{z}_t\boldsymbol{z}_t'\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1}\| \\
&\leq C_5 + C_6\|\boldsymbol{P}_t^{-1}\boldsymbol{P}_t^0\boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1}\|.
\end{aligned}
$$

As $\boldsymbol{\theta}_\rho \to \boldsymbol{\theta}_\rho^0$, $\|\boldsymbol{P}_t^{-1}\boldsymbol{P}_t^0\boldsymbol{P}_t^{-1}\otimes\boldsymbol{P}_t^{-1}\| \to \|\boldsymbol{P}_t^{-1}\otimes\boldsymbol{P}_t^{-1}\| < \infty$. It follows that $\|\boldsymbol{P}_t^{-1}\boldsymbol{P}_t^0\boldsymbol{P}_t^{-1}\otimes\boldsymbol{P}_t^{-1}\| \leq C_7 < \infty$ for $\boldsymbol{\theta}_\rho \in \mathcal{N}_1$, and

$$
\mathsf{E}\|\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\rho} \partial \boldsymbol{\rho}'}\| \leq C_5 + C_6 C_7 < \infty
$$

for $\boldsymbol{\theta}_\rho \in \mathcal{N}_1$, and all t. Similar arguments apply to the other blocks as the elements of $\frac{\partial \text{vec}(\boldsymbol{P}_t)}{\partial \gamma}$ and $\frac{\partial \text{vec}(\boldsymbol{P}_t)}{\partial c}$ are finite for all t. Thus,

$$
\mathsf{E}\|\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_\rho \partial \boldsymbol{\theta}_\rho'}\| < \infty
$$

for $\boldsymbol{\theta}_\rho \in \mathcal{N}_1$, and all t. ∎

Condition (v) can now be verified by applying Lemmas 10–12 to (52). This completes the proof of Theorem 2. ∎

Proof of Theorem 3: First we introduce some notation that agrees with the one used in the theorem. Let

$$
\boldsymbol{A}_T(\boldsymbol{\theta}_1) = \left[\begin{array}{cc} \boldsymbol{A}_{gT}(\boldsymbol{\theta}_1) & \boldsymbol{A}_{gPT}(\boldsymbol{\theta}_1) \\ \boldsymbol{A}_{Pg}(\boldsymbol{\theta}_1) & \boldsymbol{A}_{PT}(\boldsymbol{\theta}_1) \end{array}\right] \tag{57}
$$

where $\boldsymbol{\theta}_1 = (\boldsymbol{\theta}_g', \boldsymbol{\theta}_\rho')'$, be the average Hessian evaluated at $\boldsymbol{\theta}_1 \in \Theta_1$, and

$$
\boldsymbol{A}_{gT}(\boldsymbol{\theta}_1) = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_g \partial \boldsymbol{\theta}_g'} \tag{58}
$$

$$
\boldsymbol{A}_{gPT}(\boldsymbol{\theta}_1) = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_g \partial \boldsymbol{\theta}_\rho'} \tag{59}
$$

and

$$
\boldsymbol{A}_{PT}(\boldsymbol{\theta}_1) = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_\rho \partial \boldsymbol{\theta}_\rho'}.
$$

The bulk of the proof consists of proving the following result:

Lemma 13 *Suppose that the assumptions of Theorem 2 hold and define a neigbour-hood $\mathcal{N}_\delta \subset \Theta_1$ of $\boldsymbol{\theta}_1^0$ such that for all $\boldsymbol{\theta}_1 \in \mathcal{N}_\delta$, $||\boldsymbol{\theta}_1 - \boldsymbol{\theta}_1^0|| < \delta$, $\delta > 0$. Let (57) now be the average Hessian evaluated at $\boldsymbol{\theta}_1 \in \mathcal{N}_\delta$, whereas*

$$\boldsymbol{A}_T(\boldsymbol{\theta}_1^0) = \frac{1}{T}\sum_{t=1}^T \mathsf{E}\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1^0)}{\partial\boldsymbol{\theta}_1\partial\boldsymbol{\theta}_1'}$$

is the expected Hessian. Then, for $\boldsymbol{\theta}_1 \to \boldsymbol{\theta}_1^0$, the (j,i) block of (58), $j \neq i$, becomes

$$\begin{aligned}\boldsymbol{A}_{g_jg_iT}(\boldsymbol{\theta}_1) &\to -\frac{1}{4}\int_0^1 \frac{1}{g_{jr}^0 g_{ir}^0}\frac{\partial g_{jr}^0}{\partial\boldsymbol{\theta}_{gj}}\frac{\partial g_{ir}^0}{\partial\boldsymbol{\theta}_{gi}'}\boldsymbol{e}_j'(\boldsymbol{P}_r^0)^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\boldsymbol{P}_r^0\boldsymbol{e}_j\, dr \\ &= \boldsymbol{A}_{g_jg_i}(\boldsymbol{\theta}_1^0)\end{aligned}$$

and for $i = j$,

$$\begin{aligned}\boldsymbol{A}_{g_ig_iT}(\boldsymbol{\theta}_1) &\to -\frac{1}{4}\int_0^1 \frac{1}{(g_{ir}^0)^2}\frac{\partial g_{ir}^0}{\partial\boldsymbol{\theta}_{gi}}\frac{\partial g_{ir}^0}{\partial\boldsymbol{\theta}_{gi}'}(1 + \boldsymbol{e}_i'(\boldsymbol{P}_r^0)^{-1}\boldsymbol{e}_i)\, dr \\ &= \boldsymbol{A}_{g_ig_i}(\boldsymbol{\theta}_1^0).\end{aligned}$$

The ith block of (59) has the form

$$\begin{aligned}\boldsymbol{A}_{gPT}(\boldsymbol{\theta}_1) &\to -\frac{1}{2}\int_0^1 \frac{1}{g_{ir}^0}\frac{\partial g_{ir}^0}{\partial\boldsymbol{\theta}_{gi}}(\boldsymbol{e}_i \otimes \boldsymbol{e}_i)'\{\boldsymbol{I}_N \otimes (\boldsymbol{P}_r^0)^{-1}\}\frac{\partial vec(\boldsymbol{P}_r^0)}{\partial\boldsymbol{\theta}_\rho'}\, dr \\ &= \boldsymbol{A}_{gP}(\boldsymbol{\theta}_1^0)\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{A}_{PT}(\boldsymbol{\theta}_1) &\to -\frac{1}{2}\int_0^1 \frac{\partial vec(\boldsymbol{P}_r^0)}{\partial\boldsymbol{\theta}_\rho'}\{(\boldsymbol{P}_r^0)^{-1} \otimes (\boldsymbol{P}_r^0)^{-1}\}\frac{\partial vec(\boldsymbol{P}_r^0)}{\partial\boldsymbol{\theta}_\rho}\, dr \\ &= \boldsymbol{A}_P(\boldsymbol{\theta}_1^0).\end{aligned}$$

as $T \to \infty$. Summing up,

$$\boldsymbol{A}_T(\boldsymbol{\theta}_1) \xrightarrow{p} \boldsymbol{A}(\boldsymbol{\theta}_1^0) = \begin{bmatrix} \boldsymbol{A}_g(\boldsymbol{\theta}_1^0) & \boldsymbol{A}_{gP}(\boldsymbol{\theta}_1^0) \\ \boldsymbol{A}_{Pg}(\boldsymbol{\theta}_1^0) & \boldsymbol{A}_P(\boldsymbol{\theta}_1^0) \end{bmatrix}$$

as $\boldsymbol{\theta}_1 \to \boldsymbol{\theta}_1^0$ and $T \to \infty$.

Proof: First consider

$$\boldsymbol{A}_{gT}(\boldsymbol{\theta}_1) = \begin{bmatrix} \boldsymbol{A}_{g_1g_1T}(\boldsymbol{\theta}_1) & \dots & \boldsymbol{A}_{g_1g_NT}(\boldsymbol{\theta}_1) \\ & \dots & \\ \boldsymbol{A}_{g_Ng_1T}(\boldsymbol{\theta}_1) & & \boldsymbol{A}_{g_Ng_NT}(\boldsymbol{\theta}_1) \end{bmatrix}$$

from (58), where, for $i \neq j$, the (j,i)-block equals

$$\begin{aligned}\boldsymbol{A}_{g_jg_iT}(\boldsymbol{\theta}_1) &= \frac{1}{T}\sum_{t=1}^T \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}_{gj}\partial\boldsymbol{\theta}_{gi}'} \\ &= -\frac{1}{T}\sum_{t=1}^T \frac{1}{g_{jt}g_{it}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gj}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}'}\boldsymbol{e}_j'\boldsymbol{P}_t^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\boldsymbol{z}_t\boldsymbol{z}_t'\boldsymbol{e}_j.\end{aligned}$$

40

As $\{\boldsymbol{\zeta}_t\}$ is a sequence of iid$(\mathbf{0}_N, \boldsymbol{I}_N)$ vectors and $\boldsymbol{z}_t = (\boldsymbol{S}_t)^{-1}\boldsymbol{S}_t^0(\boldsymbol{P}_t^0)^{1/2}\boldsymbol{\zeta}_t$, where $\boldsymbol{S}_t^0 = \mathrm{diag}(g_{1t}^0, ..., g_{Nt}^0)$ and $g_{it}^0 = g_i(\boldsymbol{\theta}_{gi}^0; t/T)$, $i = 1, ..., N$, the corresponding block of the expected Hessian evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_1$ equals

$$
\begin{aligned}
\boldsymbol{A}_{g_j g_i T}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) &= \frac{1}{T}\sum_{t=1}^{T}\mathsf{E}\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}_{gj}\partial\boldsymbol{\theta}'_{gi}} \\
&= -\frac{1}{4T}\sum_{t=1}^{T}\frac{1}{g_{jt}g_{it}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gj}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}'_{gi}}\boldsymbol{e}'_j\boldsymbol{P}_t^{-1}\boldsymbol{e}_i\boldsymbol{e}'_i\mathsf{E}\boldsymbol{z}_t\boldsymbol{z}'_t\boldsymbol{e}_j \\
&= -\frac{1}{4T}\sum_{t=1}^{T}\frac{1}{g_{jt}g_{it}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gj}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}'_{gi}}\boldsymbol{e}'_j\boldsymbol{P}_t^{-1}\boldsymbol{e}_i\boldsymbol{e}'_i(\boldsymbol{M}_t^0)^2\boldsymbol{e}_j
\end{aligned}
$$

where

$$
(\boldsymbol{M}_t^0)^2 = (\boldsymbol{S}_t)^{-1}\boldsymbol{S}_t^0(\boldsymbol{P}_t^0)\boldsymbol{S}_t^0(\boldsymbol{S}_t)^{-1}.
$$

Then, for $\boldsymbol{\theta}_1 \in \mathcal{N}_\delta \subset \Theta_1$,

$$
\boldsymbol{A}_{g_j g_i T}(\boldsymbol{\theta}_1) = \boldsymbol{A}_{g_j g_i T}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) + \boldsymbol{o}_p(1).
$$

Since g_{jt} and g_{it} are deterministic, we again apply triangular array asymptotics to find the probability limit of $\boldsymbol{A}_{g_j g_i T}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0)$. Denoting $t = [Tr]$ and assuming $i \neq j$ leads to

$$
\begin{aligned}
& \boldsymbol{A}_{g_j g_i T}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) \\
&= -\frac{1}{4T}\sum_{t=1}^{T}\frac{1}{g_{j[Tr]}g_{i[Tr]}}\frac{\partial g_{j[Tr]}}{\partial\boldsymbol{\theta}_{gj}}\frac{\partial g_{i[Tr]}}{\partial\boldsymbol{\theta}'_{gi}}\boldsymbol{e}'_j\boldsymbol{P}_{[Tr]}^{-1}\boldsymbol{e}_i\boldsymbol{e}'_i(\boldsymbol{M}_{[Tr]}^0)^2\boldsymbol{e}_j\} \\
&= -\frac{1}{4}\sum_{t=1}^{T}\int_{t/T}^{(t+1)/T}\frac{1}{g_{j[Tr]}g_{i[Tr]}}\frac{\partial g_{j[Tr]}}{\partial\boldsymbol{\theta}_{gj}}\frac{\partial g_{i[Tr]}}{\partial\boldsymbol{\theta}'_{gi}}\boldsymbol{e}'_j\boldsymbol{P}_{[Tr]}^{-1}\boldsymbol{e}_i\boldsymbol{e}'_i(\boldsymbol{M}_{[Tr]}^0)^2\boldsymbol{e}_j\mathrm{d}r \\
&= -\frac{1}{4}\int_{1/T}^{(T+1)/T}\frac{1}{g_{j[Tr]}g_{i[Tr]}}\frac{\partial g_{j[Tr]}}{\partial\boldsymbol{\theta}_{gj}}\frac{\partial g_{i[Tr]}}{\partial\boldsymbol{\theta}'_{gi}}\boldsymbol{e}'_j\boldsymbol{P}_{[Tr]}^{-1}\boldsymbol{e}_i\boldsymbol{e}'_i(\boldsymbol{M}_{[Tr]}^0)^2\boldsymbol{e}_j\mathrm{d}r \\
&\to -\frac{1}{4}\int_0^1\frac{1}{g_{jr}g_{ir}}\frac{\partial g_{jr}}{\partial\boldsymbol{\theta}_{gj}}\frac{\partial g_{ir}}{\partial\boldsymbol{\theta}'_{gi}}\boldsymbol{e}'_j\boldsymbol{P}_r^{-1}\boldsymbol{e}_i\boldsymbol{e}'_i(\boldsymbol{M}_r^0)^2\boldsymbol{e}_j\mathrm{d}r = \boldsymbol{A}_{g_j g_i}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) \quad (60)
\end{aligned}
$$

as $t \to T$ and $T \to \infty$ with $t/T \to 1$. That $\boldsymbol{P}_{[Tr]}^{-1} \to \boldsymbol{P}_r^{-1}$ when $T \to \infty$ follows from $\boldsymbol{P}_{[Tr]} \to \boldsymbol{P}_r$ and continuity of the inverse transformation. Likewise,

$$
\begin{aligned}
\boldsymbol{A}_{g_i g_i T}(\boldsymbol{\theta}_1) &= \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}_{gi}\partial\boldsymbol{\theta}'_{gi}} \\
&= -\frac{1}{2T}\sum_{t=1}^{T}\frac{1}{g_{it}}[\{\frac{1}{g_{it}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}'_{gi}} - \frac{\partial^2 g_{it}}{\partial\boldsymbol{\theta}_{gi}\partial\boldsymbol{\theta}'_{gi}}\}(\boldsymbol{e}'_i\boldsymbol{P}_t^{-1}\boldsymbol{z}_t\boldsymbol{z}'_t\boldsymbol{e}_i - 1) \\
&\quad -\frac{1}{2g_{it}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}_{gi}}\frac{\partial g_{it}}{\partial\boldsymbol{\theta}'_{gi}}\boldsymbol{e}'_i\boldsymbol{P}_t^{-1}(\boldsymbol{I}_N + \boldsymbol{e}_i\boldsymbol{e}'_i)\boldsymbol{z}_t\boldsymbol{z}'_t\boldsymbol{e}_i]
\end{aligned}
$$

41

and

$$\boldsymbol{A}_{g_i g_i T}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) = \frac{1}{T} \sum_{t=1}^{T} \mathsf{E} \frac{\partial^2 \ln f(\boldsymbol{\zeta}_t | \boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_{gi}}$$

$$= -\frac{1}{2T} \sum_{t=1}^{T} \frac{1}{g_{it}} [\{\frac{1}{g_{it}} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}'_{gi}} - \frac{\partial^2 g_{it}}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_{gi}}\}(\boldsymbol{e}'_i \boldsymbol{P}_t^{-1} (\boldsymbol{M}_t^0)^2 \boldsymbol{e}_i - 1)$$

$$- \frac{1}{2g_{it}} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}'_{gi}} \boldsymbol{e}'_i \boldsymbol{P}_t^{-1} (\boldsymbol{I}_N + \boldsymbol{e}_i \boldsymbol{e}'_i)(\boldsymbol{M}_t^0)^2 \boldsymbol{e}_i].$$

Proceeding as above,

$$\boldsymbol{A}_{g_i g_i T}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0)$$
$$\xrightarrow{p} -\frac{1}{2} \int_0^1 [\frac{1}{g_{ir}} \{\frac{1}{g_{ir}} \frac{\partial g_{ir}}{\partial \boldsymbol{\theta}_{gi}} \frac{\partial g_{ir}}{\partial \boldsymbol{\theta}'_{gi}} - \frac{\partial^2 g_{ir}}{\partial \boldsymbol{\theta}_{gi} \partial \boldsymbol{\theta}'_{gi}}\}(\boldsymbol{e}'_i \boldsymbol{P}_r^{-1} (\boldsymbol{M}_r^0)^2 \boldsymbol{e}_i - 1)] \mathrm{d}r$$
$$- \frac{1}{4} \int_0^1 \frac{1}{g_{it}^2} \frac{\partial g_{ir}}{\partial \boldsymbol{\theta}_{gi}} \frac{\partial g_{ir}}{\partial \boldsymbol{\theta}'_{gi}} \boldsymbol{e}'_i \boldsymbol{P}_r^{-1} (\boldsymbol{I}_N + \boldsymbol{e}_i \boldsymbol{e}'_i)(\boldsymbol{M}_r^0)^2 \boldsymbol{e}_i \mathrm{d}r$$
$$= \boldsymbol{A}_{g_i g_i}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0). \tag{61}$$

Let $\boldsymbol{\theta}_1 \to \boldsymbol{\theta}_1^0$, which implies $\boldsymbol{S}_t \to \boldsymbol{S}_t^0$ and $\boldsymbol{P}_t \xrightarrow{p} \boldsymbol{P}_t^0$. Since the inverse transformation is continuous, the continuous mapping theorem applies, so $\boldsymbol{S}_t^{-1} \to (\boldsymbol{S}_t^0)^{-1}$, which leads to $\boldsymbol{S}_t^{-1} \boldsymbol{S}_t^0 \to \boldsymbol{I}_N$ or $(\boldsymbol{M}_t^0)^2 \to \boldsymbol{P}_t^0$. Furthermore, by the same theorem, $\boldsymbol{P}_t^{-1} \to (\boldsymbol{P}_t^0)^{-1}$. Applying this to (60) yields

$$\boldsymbol{A}_{g_j g_i}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) \to -\frac{1}{4} \int_0^1 \frac{1}{g_{jr}^0 g_{ir}^0} \frac{\partial g_{jr}^0}{\partial \boldsymbol{\theta}_{gj}} \frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}'_{gi}} \boldsymbol{e}'_j (\boldsymbol{P}_r^0)^{-1} \boldsymbol{e}_i \boldsymbol{e}'_i \boldsymbol{P}_r^0 \boldsymbol{e}_j \mathrm{d}r$$
$$= \boldsymbol{A}_{g_j g_i}(\boldsymbol{\theta}_1^0)$$

where $g_{\ell r}^0 = g(r, \boldsymbol{\theta}_{g\ell}^0)$, and $\partial g_{\ell r}^0 / \partial \boldsymbol{\theta}_{g\ell} = \partial g_{\ell r} / \partial \boldsymbol{\theta}_{g\ell}|_{\boldsymbol{\theta}_{g\ell} = \boldsymbol{\theta}_{g\ell}^0}$, $\ell = i, j$, for all $i, j = 1, ..., N$. Note that $\boldsymbol{e}'_j (\boldsymbol{P}_r^0)^{-1} \boldsymbol{e}_i$ is the (j, i) element of $(\boldsymbol{P}_r^0)^{-1}$ and $\boldsymbol{e}'_i \boldsymbol{P}_r^0 \boldsymbol{e}_j$ is the (i, j) element of $\boldsymbol{e}'_i \boldsymbol{P}_r^0 \boldsymbol{e}_j$.

Applying the same arguments to (61) yields

$$\boldsymbol{A}_{g_i g_i}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) \to -\frac{1}{4} \int_0^1 \frac{1}{(g_{it}^0)^2} \frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}_{gi}} \frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}'_{gi}} (1 + \boldsymbol{e}'_i (\boldsymbol{P}_r^0)^{-1} \boldsymbol{e}_i) \mathrm{d}r$$
$$= \boldsymbol{A}_{g_i g_i}(\boldsymbol{\theta}_1^0)$$

because $\boldsymbol{P}_r^0$ is a correlation matrix, so $\boldsymbol{e}'_i \boldsymbol{P}_r^0 \boldsymbol{e}_i = 1$, $i = 1, ..., N$.

More generally,

$$\boldsymbol{A}_{gT}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) \xrightarrow{p} [\boldsymbol{A}_{g_j g_i}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0)]$$
$$= \boldsymbol{A}_g(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0)$$

as $T \to \infty$, and

$$\boldsymbol{A}_g(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) \to [\boldsymbol{A}_{g_j g_i}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0)]$$
$$= \boldsymbol{A}_g(\boldsymbol{\theta}_1^0).$$

as $\boldsymbol{\theta}_1 \to \boldsymbol{\theta}_1^0$.

To obtain the desired result, apply the triangular inequality using previous results to the difference $||\boldsymbol{A}_{gT}(\boldsymbol{\theta}_1) - \boldsymbol{A}_g(\boldsymbol{\theta}_1^0)||$. This yields

$$
\begin{aligned}
&||\boldsymbol{A}_{gT}(\boldsymbol{\theta}_1) - \boldsymbol{A}_g(\boldsymbol{\theta}_1^0)|| \\
\leq\ &||\boldsymbol{A}_{gT}(\boldsymbol{\theta}_1) - \boldsymbol{A}_{gT}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0)|| + ||\boldsymbol{A}_{gT}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) - \boldsymbol{A}_g(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0)|| \\
&+ ||\boldsymbol{A}_g(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) - \boldsymbol{A}_g(\boldsymbol{\theta}_1^0)|| \\
\leq\ &\boldsymbol{o}_p(1) + ||\boldsymbol{A}_{gT}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) - \boldsymbol{A}_g(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0)|| \\
&+ \sup_{\boldsymbol{\theta}_1 \in \mathcal{N}_\delta} ||\boldsymbol{A}_g(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) - \boldsymbol{A}_g(\boldsymbol{\theta}_1^0)|| \\
&\xrightarrow{p} 0
\end{aligned}
\tag{62}
$$

as $\boldsymbol{\theta}_1 \to \boldsymbol{\theta}_1^0$ and $T \to \infty$, where

$$
[\boldsymbol{A}_g(\boldsymbol{\theta}_1^0)]_{ji} = -\frac{1}{4} \int_0^1 \frac{1}{g_{jr}^0 g_{ir}^0} \frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}_{gj}} \frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}_{gi}'} \boldsymbol{e}_j'(\boldsymbol{P}_r^0)^{-1} \boldsymbol{e}_i \boldsymbol{e}_i' \boldsymbol{P}_r^0 \boldsymbol{e}_j \mathrm{d}r.
$$

Similarly, taking expectations and letting $T \to \infty$,

$$
\begin{aligned}
&\boldsymbol{A}_{gPT}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) \\
=\ &-\frac{1}{2T} \sum_{t=1}^T \frac{1}{g_{it}} \frac{\partial g_{it}}{\partial \boldsymbol{\theta}_{gi}} (\boldsymbol{e}_i \otimes \boldsymbol{e}_i)'(\mathsf{E}\boldsymbol{z}_t \boldsymbol{z}_t' \boldsymbol{P}_t^{-1} \otimes \boldsymbol{P}_t^{-1}) \frac{\partial \mathrm{vec}(\boldsymbol{P}_t)}{\partial \boldsymbol{\theta}_\rho'} \\
&\xrightarrow{p} -\frac{1}{2} \int_0^1 \frac{1}{g_{ir}} \frac{\partial g_{ir}}{\partial \boldsymbol{\theta}_{gi}} (\boldsymbol{e}_i \otimes \boldsymbol{e}_i)'\{(\boldsymbol{M}_r^0)^2 \boldsymbol{P}_r^{-1} \otimes \boldsymbol{P}_r^{-1}\} \frac{\partial \mathrm{vec}(\boldsymbol{P}_r)}{\partial \boldsymbol{\theta}_\rho'} \\
=\ &\boldsymbol{A}_{gP}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0).
\end{aligned}
\tag{63}
$$

Letting $\boldsymbol{\theta}_1 \to \boldsymbol{\theta}_1^0$ in (63) gives

$$
\begin{aligned}
&\boldsymbol{A}_{gP}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) \to -\frac{1}{2} \int_0^1 \frac{1}{g_{ir}^0} \frac{\partial g_{ir}^0}{\partial \boldsymbol{\theta}_{gi}} (\boldsymbol{e}_i \otimes \boldsymbol{e}_i)'\{\boldsymbol{I}_N \otimes (\boldsymbol{P}_r^0)^{-1}\} \frac{\partial \mathrm{vec}(\boldsymbol{P}_r^0)}{\partial \boldsymbol{\theta}_\rho'} \mathrm{d}r \\
=\ &\boldsymbol{A}_{gP}(\boldsymbol{\theta}_1^0).
\end{aligned}
$$

Considerations similar to the ones in (62) lead to

$$
||\boldsymbol{A}_{gPT}(\boldsymbol{\theta}_1) - \boldsymbol{A}_{gP}(\boldsymbol{\theta}_1^0)|| \xrightarrow{p} 0
\tag{64}
$$

as $\boldsymbol{\theta}_1 \to \boldsymbol{\theta}_1^0$ and $T \to \infty$.

Denoting

$$
\frac{\partial \mathrm{vec}(\boldsymbol{P}_r^0)}{\partial \boldsymbol{\theta}_\rho} = \frac{\partial \mathrm{vec}(\boldsymbol{P}_r)}{\partial \boldsymbol{\theta}_\rho}\Big|_{\boldsymbol{\theta}_\rho = \boldsymbol{\theta}_\rho^0}
$$

one obtains

$$
\begin{aligned}
\boldsymbol{A}_{PT}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) &= \frac{1}{T}\sum_{t=1}^{T}\mathsf{E}\frac{\partial^2 \ln f(\boldsymbol{\zeta}_t|\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}'_\rho\partial\boldsymbol{\theta}_\rho} \\
&= -\frac{1}{2T}\sum_{t=1}^{T}\frac{\partial^2\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\theta}'_\rho\partial\boldsymbol{\theta}_\rho}\{\mathrm{vec}(\boldsymbol{P}_t^{-1}) - \mathrm{vec}(\boldsymbol{P}_t^{-1}(\boldsymbol{M}_r^0)^2\boldsymbol{P}_t^{-1})\} \\
&\quad -\frac{1}{2T}\sum_{t=1}^{T}\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\theta}'_\rho}\{\boldsymbol{P}_t^{-1}\otimes\boldsymbol{P}_t^{-1} \\
&\quad -(\boldsymbol{P}_t^{-1}(\boldsymbol{M}_r^0)^2\boldsymbol{P}_t^{-1}\otimes\boldsymbol{P}_t^{-1} + \boldsymbol{P}_t^{-1}\otimes\boldsymbol{P}_t^{-1}(\boldsymbol{M}_r^0)^2\boldsymbol{P}_t^{-1})\}\frac{\partial\mathrm{vec}(\boldsymbol{P}_t)}{\partial\boldsymbol{\theta}_\rho}.
\end{aligned}
$$

As $T \to \infty$,

$$
\begin{aligned}
\boldsymbol{A}_{PT}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) &\xrightarrow{p} -\frac{1}{2}\int_0^1\frac{\partial\mathrm{vec}(\boldsymbol{P}_r)}{\partial\boldsymbol{\theta}'_\rho}\{\boldsymbol{P}_r^{-1}\otimes\boldsymbol{P}_r^{-1} \\
&\quad -(\boldsymbol{P}_r^{-1}(\boldsymbol{M}_r^0)^2\boldsymbol{P}_r^{-1}\otimes\boldsymbol{P}_r^{-1} + \boldsymbol{P}_r^{-1}\otimes\boldsymbol{P}_r^{-1}(\boldsymbol{M}_r^0)^2\boldsymbol{P}_r^{-1})\}\frac{\partial\mathrm{vec}(\boldsymbol{P}_r)}{\partial\boldsymbol{\theta}_\rho} \\
&= \boldsymbol{A}_P(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0). \qquad (65)
\end{aligned}
$$

Finally, letting $\boldsymbol{\theta}_1 \to \boldsymbol{\theta}_1^0$ in (65),

$$
\begin{aligned}
\boldsymbol{A}_P(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^0) &\to -\frac{1}{2}\int_0^1\frac{\partial\mathrm{vec}(\boldsymbol{P}_r^0)}{\partial\boldsymbol{\theta}'_\rho}\{(\boldsymbol{P}_r^0)^{-1}\otimes(\boldsymbol{P}_r^0)^{-1}\}\frac{\partial\mathrm{vec}(\boldsymbol{P}_r^0)}{\partial\boldsymbol{\theta}_\rho}\mathrm{d}r \\
&= \boldsymbol{A}_P(\boldsymbol{\theta}_1^0).
\end{aligned}
$$

As in the two previous cases, using the triangular inequality leads to the conclusion that

$$
||\boldsymbol{A}_{PT}(\boldsymbol{\theta}_1) - \boldsymbol{A}_P(\boldsymbol{\theta}_1^0)|| \xrightarrow{p} 0 \qquad (66)
$$

as $\boldsymbol{\theta}_1 \to \boldsymbol{\theta}_1^0$ and $T \to \infty$. Finally, due to (62), (64) and (66),

$$
\begin{aligned}
||\boldsymbol{A}_T(\boldsymbol{\theta}_1) - \boldsymbol{A}(\boldsymbol{\theta}_1^0)|| &\le ||\boldsymbol{A}_{gT}(\boldsymbol{\theta}_1) - \boldsymbol{A}_g(\boldsymbol{\theta}_1^0)|| + 2||\boldsymbol{A}_{gPT}(\boldsymbol{\theta}_1) - \boldsymbol{A}_{gP}(\boldsymbol{\theta}_1^0)|| \xrightarrow{p} \\
&\quad +||\boldsymbol{A}_{gPT}(\boldsymbol{\theta}_1) - \boldsymbol{A}_{gP}(\boldsymbol{\theta}_1^0)|| \xrightarrow{p} 0
\end{aligned}
$$

as $\boldsymbol{\theta}_1 \to \boldsymbol{\theta}_1^0$ and $T \to \infty$. This completes the proof. $\blacksquare$

The theorem is now proved by replacing $\boldsymbol{\theta}_1 \subset \mathcal{N}_\delta$ above by the consistent estimator $\widehat{\boldsymbol{\theta}}_1$ and applying Lemma 13. $\blacksquare$

13 Estimated TVC equations

The estimated equations for the variance components and the estimated correlation equation are as follows (the figures in parentheses are standard deviation estimates):

S&P 500 returns:

$$
\widehat{h}_t^S = \underset{(0.0054)}{0.043} + \underset{(0.0156)}{0.164}\, I(\varepsilon_{t-1}^S < 0)\frac{(\varepsilon_{t-1}^S)^2}{\widehat{g}_{t-1}^S} + \underset{(0.0110)}{0.872}\widehat{h}_{t-1}^S \qquad (67)
$$

where

$$\widehat{g}_t^S = 2.18 - \underset{(0.145)}{1.42}\, G_{1t}^S + \underset{(1.42)}{8.71} G_{2t}^S - \underset{(1.42)}{8.83} G_{3t}^S \tag{68}$$

with

$$G_{1t}^S = (1 + \exp\{-\exp\underset{(-)}{\{12.6\}}(t/T - \underset{(0.0008)}{0.232})\})^{-1} \tag{69}$$

$$G_{2t}^S = (1 + \exp\{-\exp\{\underset{(0.387)}{3.27}\}(t/T - \underset{(0.0034)}{0.54})\})^{-1}$$

and

$$G_{3t}^S = (1 + \exp\{-\exp\{\underset{(0.057)}{2.79}\}(t/T - \underset{(0.011)}{0.63})\})^{-1}.$$

T-bill returns:

$$\widehat{h}_t^B = \underset{(0.0045)}{0.0142} + \underset{(0.0044)}{0.0277}\frac{(\varepsilon_{t-1}^B)^2}{\widehat{g}_{t-1}^B} + \underset{(0.0077)}{0.958}\,\widehat{h}_{t-1}^B. \tag{70}$$

where

$$\widehat{g}_t^B = 0.196 + \underset{(0.39)}{6.59}G_{1t}^B - \underset{(0.39)}{6.69}G_{2t}^B \tag{71}$$

with

$$G_{1t}^B = (1 + \exp\{-\exp\{\underset{(0.106)}{3.15}\}(t/T - \underset{(0.0008)}{0.518})\})^{-1}$$

and

$$G_{2t}^B = (1 + \exp\{-\exp\{\underset{(0.113)}{3.03}\}(t/T - \underset{(0.0008)}{0.521})\})^{-1}.$$

The correlation equation has the following form:

$$\widehat{\rho}_t = \underset{(0.037)}{0.0495}G_t - \underset{(0.015)}{0.332}(1 - G_t) \tag{72}$$

where

$$G_t = (1 + \exp\{-\exp\{13.0\}(t/T - \underset{(0.0001)}{0.26})(t/T - \underset{(0.0008)}{0.46})\})^{-1}. \tag{73}$$

Standard deviations of the large slope estimates $\widehat{\eta}$ in (69) and (73) are not given (they have been computed) because the corresponding transitions are very steep. The steep slopes, almost thresholds, make the Hessian ill-conditioned, and in computing this matrix these components have been removed. The purpose of not reporting the standard deviations is to remind the reader that the corresponding partial derivatives have been excluded from the Hessian. Also recall that for identification reasons, the intercept in equations (68) and (71) is fixed.

Research Papers
2017

CREATES
Center for Research in Econometric
Analysis of Time Series