Supplementary Material:

Non-Gaussian score-driven conditionally heteroskedastic models with a macroeconomic application

Szabolcs Blazsek [a], Álvaro Escribano [b,*], and Adrián Licht [a]

[a] School of Business, Universidad Francisco Marroquín, Guatemala City, Guatemala

[b] Department of Economics, Universidad Carlos III de Madrid, Getafe (Madrid), Spain

[*] Corresponding author.

E-mail addresses: sblazsek@ufm.edu (S. Blazsek), alvaroe@eco.uc3m.es (A. Escribano), adrian.licht@ufm.edu (A. Licht).

1. Maximum likelihood (ML) estimator

We estimate all models by using the ML method. The ML estimates of parameters are given by:

$$\hat{\Theta} = \arg\max_{\Theta} \mathrm{LL}(Y_1, \ldots, Y_T, \Theta) = \arg\max_{\Theta} \sum_{t=1}^{T} \ln f(Y_t|\mathcal{F}_{t-1}, \Theta) \tag{S.1}$$

where LL is the log-likelihood function and $\mathcal{F}_{t-1} = \sigma(Y_1, \ldots, Y_{t-1}, X_1, \Omega_1)$.

In the following, the gradient vector $G_t(\Theta)$ and the Hessian matrix $H_t(\Theta)$ of LL are defined. The $T \times S$ matrix of contributions to the gradient $G(Y_1, \ldots, Y_T, \Theta)$ is defined by its elements:

$$G_{t,i}(\Theta) = -\frac{\partial \ln f(Y_t|\mathcal{F}_{t-1}; \Theta)}{\partial \Theta_i} \tag{S.2}$$

for period $t = 1, \ldots, T$, and parameter $i = 1, \ldots, S$. The t-th row of $G(Y_1, \ldots, Y_T, \Theta)$ is denoted by using $G_t(\Theta)$, which is the score vector for the t-th observation. Under the ML assumptions of the next section, the maximization problem of Equation (S.1) is equivalent to:

$$\frac{1}{T}\sum_{t=1}^{T} G_t(\hat{\Theta})' = \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} G_{t,1}(\hat{\Theta}) \\ \vdots \\ G_{t,S}(\hat{\Theta}) \end{bmatrix} = \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} -\frac{\partial \ln f(Y_t|\mathcal{F}_{t-1};\hat{\Theta})}{\partial \Theta_1} \\ \vdots \\ -\frac{\partial \ln f(Y_t|\mathcal{F}_{t-1};\hat{\Theta})}{\partial \Theta_S} \end{bmatrix} = 0_{S\times 1} \tag{S.3}$$

According to the mean-value expansion about the true values of parameters Θ_0:

$$\frac{1}{T}\sum_{t=1}^{T} G_t(\hat{\Theta})' = \frac{1}{T}\sum_{t=1}^{T} G_t(\Theta_0)' + \frac{1}{T}\left[\sum_{t=1}^{T} H_t(\bar{\Theta})\right](\hat{\Theta} - \Theta_0) \tag{S.4}$$

where each row of the $S \times S$ Hessian matrix:

$$H_t(\Theta) = -\frac{\partial^2 \ln f(Y_t|\mathcal{F}_{t-1};\Theta)}{\partial\Theta\partial\Theta'} \tag{S.5}$$

is evaluated at S different mean values $\bar{\Theta}$ of Equation (S.4). Each $\bar{\Theta}$ is located between Θ_0 and $\hat{\Theta}$: $||\bar{\Theta} - \Theta_0|| \leq ||\hat{\Theta} - \Theta_0||$, where $||\cdot||$ is the Euclidean norm. From Equations (S.3) and (S.4):

$$\sqrt{T}(\hat{\Theta} - \Theta_0) = \left[-\frac{1}{T}\sum_{t=1}^{T} H_t(\bar{\Theta})\right]^{-1}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} G_t(\Theta_0)'\right] \tag{S.6}$$

The asymptotic covariance matrix of parameters $\hat{\Theta}$ is estimated by using the inverse information matrix: $\{(1/T)\sum_{t=1}^{T}[G_t(\hat{\Theta})'G_t(\hat{\Theta})]\}^{-1}$. We prove the consistency and asymptotic normality of the ML parameter estimates, and the consistency of inverse information matrix-based estimator of standard errors of parameters in the remainder of the Supplementary Material.

2. Assumptions

(A1) $\tilde{\Theta}$ is the parameter set, for which $\Theta \in \tilde{\Theta} \subset \mathbb{R}^S$, and $\tilde{\Theta}$ is compact.

(A2) Asymptotically, $f(Y_t|\mathcal{F}_{t-1};\Theta_0) = p_0(Y_t|\mathcal{F}_{t-1};\Theta_0)$ for Θ_0 from the parameter set $\tilde{\Theta} \subset \mathbb{R}^S$, where p_0 is the true conditional density, and Θ_0 represents the true values of Θ.

(A3) Asymptotically, $f(Y_t|\mathcal{F}_{t-1};\Theta_0)$ for Θ_0 is a dynamically complete density [Wooldridge (1994)].

(A4) $\lambda_{i,t}$ for $i = 1,\ldots,\mathcal{N}$ are uniformly bounded, i.e. $\exists\ \lambda_{\max} \in \mathbb{R}^+$ such that $|\lambda_{i,t}| \leq \lambda_{\max} < \infty$ for $i = 1,\ldots,\mathcal{N}$ and for all t and $\Theta \in \tilde{\Theta}$.

(A5) Y_t is strictly stationary for $T \to \infty$ and ergodic on $\mathbb{R}^{\mathcal{N}}$ for all $\Theta \in \tilde{\Theta}$.

(A6) $\ln f(\cdot|\mathcal{F}_{t-1};\Theta) : \mathbb{R}^{\mathcal{N}} \times \tilde{\Theta} \to \mathbb{R}$ is a real-valued function.

(A7) For each $\Theta \in \tilde{\Theta}$, $\ln f(\cdot|\mathcal{F}_{t-1};\Theta)$ is a Borel measurable function on $\mathbb{R}^{\mathcal{N}}$.

(A8) For each $Y_t \in \mathbb{R}^{\mathcal{N}}$, $\ln f(Y_t|\mathcal{F}_{t-1};\cdot)$ is a continuous function on $\tilde{\Theta}$.

(A9) $\exists$ function $b(\cdot) : \mathbb{R}^{\mathcal{N}} \to \mathbb{R}$ such that $|\ln f(Y_t|\mathcal{F}_{t-1};\Theta)| \leq b(Y_t)$ for all Θ, and $E[b(Y_t)] < \infty$.

(A10) $\int_{\mathbb{R}^{\mathcal{N}}} f(Y_t|\mathcal{F}_{t-1};\Theta)dY_t = 1$ for all Θ.

(A11) Θ_0 is a unique solution to:

$$\max_{\Theta \in \tilde{\Theta}} \text{plim}_{T\to\infty} T^{-1} \sum_{t=1}^{T} E[\ln f(y_t|\mathcal{F}_{t-1}, \Theta)] \tag{S.7}$$

(A12) Each element of $H_t(\Theta)$ is strictly stationary for $T \to \infty$ and ergodic.

(A13) For each element of $H_t(\Theta)$, $H_{i,j,t}(\Theta) : \mathbb{R}^{\mathcal{N}} \times \tilde{\Theta} \to \mathbb{R}$ is a real-valued function.

(A14) For each $\Theta \in \tilde{\Theta}$, each element of $H_t(\Theta)$ is a Borel measurable function on $\mathbb{R}^{\mathcal{N}}$.

(A15) For each $Y_t \in \mathbb{R}^{\mathcal{N}}$, each element of $H_t(\Theta)$ is a continuous function on $\tilde{\Theta}$.

(A16) $\exists$ function $b(\cdot) : \mathbb{R}^{\mathcal{N}} \to \mathbb{R}$ such that, for all elements of $H_t(\Theta)$, $|H_{i,j,t}(\Theta)| \leq b[H_{i,j,t}(\Theta)]$ for all Θ, and $E\{b[H_{i,j,t}(\Theta)]\} < \infty$.

(A17) $E[G_t(\Theta_0)G_t(\Theta_0)'] < \infty$ for $T \to \infty$.

(A18) $(1/\sqrt{T}) \sum_{t=1}^{T} E[G_t(\Theta_0)'] \to 0_{S\times 1}$ for $T \to \infty$.

(A19) $(1/\sqrt{T}) \sum_{t=1}^{T} G_t(\Theta_0)' \to_d N(0, B_0)$ for $T \to \infty$, where:

$$B_0 = \lim_{T\to\infty} \text{Var}\left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} G_t(\Theta_0)'\right] \tag{S.8}$$

(A20) Θ_0 is an interior point within $\tilde{\Theta} \subset \mathbb{R}^S$.

(A21) For each Y_t, $\ln f(Y_t|\mathcal{F}_{t-1}; \cdot)$ is twice continuously differentiable on all interior points of $\tilde{\Theta}$.

(A22) $\partial[\int_{\mathbb{R}^{\mathcal{N}}} f(Y_t|\mathcal{F}_{t-1}; \Theta)dY_t]/\partial\Theta = \int_{\mathbb{R}^{\mathcal{N}}} [\partial f(Y_t|\mathcal{F}_{t-1}; \Theta)/\partial\Theta]dY_t$.

(A23) $\partial[\int_{\mathbb{R}^{\mathcal{N}}} G_t(\Theta)' f(Y_t|\mathcal{F}_{t-1}; \Theta)dY_t]/\partial\Theta = \int_{\mathbb{R}^{\mathcal{N}}} [\partial G_t(\Theta)' f(Y_t|\mathcal{F}_{t-1}; \Theta)/\partial\Theta]dY_t$.

(A24) The following matrix is positive definite:

$$A_0 = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E[H_t(\Theta_0)] = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \text{Var}[G_t(\Theta_0)'] \tag{S.9}$$

3. Properties of the score functions

3.1. Boundedness of the score functions and their derivatives

The results of this section are true for all t and $\Theta \in \tilde{\Theta}$. In this section, we prove the existence of all unconditional moments of $u_{i,t}$ and $e_{i,t}$ for $i = 1, \ldots, \mathcal{N}$, and their derivatives.

(i) From Equation (20) of the paper, u_t as a function of the structural form errors is:

$$u_t = [(\nu - 2)\nu]^{1/2} D\Omega_t \times \frac{\tilde{\epsilon}_t}{\nu - 2 + \tilde{\epsilon}_t' \tilde{\epsilon}_t} \tag{S.10}$$

where $\tilde{\epsilon}_t \sim t[0, I_{\mathcal{N}} \times (\nu - 2)/\nu, \nu]$ with $\nu > 2$ is an i.i.d. multivariate t-distribution with zero mean and identity covariance matrix. Since Ω_t is a diagonal matrix, the i-th element of u_t is:

$$u_{i,t} = [(\nu - 2)\nu]^{1/2} D_{i,i} \exp(\lambda_{i,t}) \times \frac{\tilde{\epsilon}_{i,t}}{\nu - 2 + \sum_{j=1}^{\mathcal{N}} \tilde{\epsilon}_{j,t}^2} \tag{S.11}$$

First, $\exp(\lambda_{i,t}) < \infty$ due to assumption (A4). Second, for the last multiplier of Equation (S.11):

$$\frac{\tilde{\epsilon}_{i,t}}{\nu - 2 + \sum_{j=1}^{\mathcal{N}} \tilde{\epsilon}_{j,t}^2} \to_p 0 \qquad \text{if} \qquad |\tilde{\epsilon}_{i,t}| \to \infty \tag{S.12}$$

for $i = 1, \ldots, \mathcal{N}$. In addition, since Equation (S.12) is a continuous function of $\tilde{\epsilon}_{i,t}$, Equation (S.12) is a bounded function of $\tilde{\epsilon}_{i,t}$. Hence, all unconditional moments of u_t exist.

(ii) From Equation (16) of the paper, $e_{i,t}$ as a function of the reduced form errors is:

$$e_{i,t} = \frac{\partial \ln f_i(Y_{i,t}|\mathcal{F}_{t-1}, \Theta)}{\partial \lambda_{i,t}} = \frac{(\nu + 1) v_{i,t}^2}{\nu \exp(2\lambda_{i,t}) + v_{i,t}^2} - 1 \tag{S.13}$$

for $i = 1, \ldots, \mathcal{N}$. We also write $e_{i,t}$ as a function of the structural form errors:

$$e_{i,t} = \left\{ \frac{(\nu + 1)(\nu - 2)}{\nu} \times \frac{(D_{i,1}\tilde{\epsilon}_{1,t} + \ldots + D_{i,\mathcal{N}}\tilde{\epsilon}_{\mathcal{N},t})^2}{\nu + (D_{i,1}\tilde{\epsilon}_{1,t} + \ldots + D_{i,\mathcal{N}}\tilde{\epsilon}_{\mathcal{N},t})^2} \right\} - 1 \tag{S.14}$$

where $\tilde{\epsilon}_t \sim t[0, I_{\mathcal{N}} \times (\nu - 2)/\nu, \nu]$ with $\nu > 2$ is an i.i.d. multivariate t-distribution with zero mean and identity covariance matrix. The multiplier

$$\frac{(D_{i,1}\tilde{\epsilon}_{1,t} + \ldots + D_{i,\mathcal{N}}\tilde{\epsilon}_{\mathcal{N},t})^2}{\nu + (D_{i,1}\tilde{\epsilon}_{1,t} + \ldots + D_{i,\mathcal{N}}\tilde{\epsilon}_{\mathcal{N},t})^2} \in (0, 1) \tag{S.15}$$

Thus, $e_{i,t}$ is a bounded function of $\tilde{\epsilon}_{i,t}$, and all unconditional moments of $e_{i,t}$ exist.

(iii) For $i = 1, \ldots, \mathcal{N}$, $g(e_{i,t}) \equiv \alpha_i e_{i,t-1} + \alpha_i^* \text{sgn}(-\epsilon_{i,t-1})(e_{i,t-1}+1)$, due to the boundedness of the sgn function, $g(e_{i,t})$ is a bounded function of $\tilde{\epsilon}_{i,t}$, and all unconditional moments of $g(e_{i,t})$ exist.

(iv) By using similar arguments, under assumption (A4), we also have that $\partial u_t/\partial(CX_{t-1})$, $\partial u_t/\partial\lambda_{i,t}$, $\partial e_{i,t}/\partial(CX_{t-1})$, $\partial e_{i,t}/\partial\lambda_{i,t}$, $\partial g(e_{i,t})/\partial(CX_{t-1})$, and $\partial g(e_{i,t})/\partial\lambda_{i,t}$ are bounded functions of $\tilde{\epsilon}_{i,t}$ for $i = 1, \ldots, \mathcal{N}$. Hence, all moments and covariances of $u_{i,t}$, $e_{i,t}$, $g(e_{i,t})$, and their derivatives exist.

3.2. The score functions are white noise

The results in this section hold asymptotically at the true values of parameters Θ_0. We study the consequences of assumption (A3) on the score functions:

(i) u_t is a martingale difference sequence (MDS) due to the following arguments. First, due to (A3) $G_t(\Theta_0)'$ is a MDS:

$$E_{t-1}\left[\frac{\partial \ln f(Y_t|\mathcal{F}_{t-1},\Theta)}{\partial\Theta'}\right] = E_{t-1}\left[\frac{\partial \ln f(Y_t|\mathcal{F}_{t-1},\Theta)}{\partial[CX_{t-1}]'}\right] \times \frac{\partial[CX_{t-1}]}{\partial\Theta'} = 0_{1\times S} \tag{S.16}$$

where E_{t-1} indicates expectations that are conditional on $\mathcal{F}_{t-1}$. Since $\partial[CX_{t-1}]/\partial\Theta' \neq 0_{\mathcal{N}\times S}$,

$$E_{t-1}\left[\frac{\partial \ln f(Y_t|\mathcal{F}_{t-1},\Theta)}{\partial CX_{t-1}}\right] = E_{t-1}\left(\frac{\nu+\mathcal{N}}{\nu}\Sigma_t^{-1} \times u_t\right) = \frac{\nu+\mathcal{N}}{\nu}\Sigma_t^{-1}E_{t-1}(u_t) = 0_{\mathcal{N}\times 1} \tag{S.17}$$

Since $[(\nu+\mathcal{N})/\nu]\Sigma_t^{-1} \neq 0_{\mathcal{N}\times\mathcal{N}}$, we conclude that $E_{t-1}(u_t) = 0_{\mathcal{N}\times 1}$, i.e. u_t is a MDS.

(ii) $e_{i,t}$ for $i = 1, \ldots, \mathcal{N}$ are MDSs due to the following arguments. Due to (A3), $G_t(\Theta_0)'$ is a MDS:

$$E_{t-1}\left[\frac{\partial \ln f_i(Y_t|\mathcal{F}_{t-1},\Theta)}{\partial\Theta'}\right] = E_{t-1}\left[\frac{\partial \ln f_i(Y_t|\mathcal{F}_{t-1},\Theta)}{\partial\lambda_{i,t}}\right] \times \frac{\partial\lambda_{i,t}}{\partial\Theta'} = 0_{1\times S} \tag{S.18}$$

Since $(\partial\lambda_{i,t}/\partial\Theta') \neq 0_{1\times S}$,

$$E_{t-1}\left[\frac{\partial \ln f_i(Y_t|\mathcal{F}_{t-1},\Theta)}{\partial\lambda_{i,t}}\right] = E_{t-1}(e_{i,t}) = 0 \tag{S.19}$$

Thus, $e_{i,t}$ for $i = 1, \ldots, \mathcal{N}$ are MDSs.

(iii) For $i = 1, \ldots, \mathcal{N}$, $g(e_{i,t}) \equiv \alpha_i e_{i,t-1} + \alpha_i^* \text{sgn}(-\epsilon_{i,t-1})(e_{i,t-1}+1)$ is MDS.

(iv) Due to the law of iterated expectations, $E(u_t) = 0_{\mathcal{N}\times 1}$, $E(e_{i,t}) = 0$, and $E[g(e_{i,t})] = 0$.

(v) u_t is MDS and the second moments of u_t exist, hence u_t is white noise vector [White (2001)].

(vi) For $i = 1, \ldots, \mathcal{N}$, $e_{i,t}$ is MDS and $\mathrm{Var}(e_{i,t}) < \infty$, hence $e_{i,t}$ is white noise [White (2001)].

(vii) For $i = 1, \ldots, \mathcal{N}$, $g(e_{i,t}) \equiv \alpha_i e_{i,t-1} + \alpha_i^* \mathrm{sgn}(-\epsilon_{i,t-1})(e_{i,t-1} + 1)$ is MDS and $\mathrm{Var}[g(e_{i,t})] < \infty$, hence $g(e_{i,t})$ is white noise [White (2001)].

3.3. The score functions are stationary and ergodic

We show that u_t and $e_{i,t}$ for $i = 1, \ldots, \mathcal{N}$ are stationary and ergodic for all t and $\Theta \in \tilde{\Theta}$.

(i) Scaled score function u_t is a continuous function of $(\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_t)$. Under (A4), u_t is an $\mathcal{F}$-measurable function of $(\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_t)$ [White (2001)].

(ii) Scaled score function u_t is strictly stationary and ergodic, because u_t is an $\mathcal{F}$-measurable function of $(\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_t)$, and $\tilde{\epsilon}_t$ is strictly stationary and ergodic [White (2001, Theorem 3.35)].

(iii) Score function $e_{i,t}$ is i.i.d., because $e_{i,t}$ is a continuous function of $\tilde{\epsilon}_t$, and $\tilde{\epsilon}_t$ is i.i.d. [White (2001)].

(iv) Score function $e_{i,t}$ is an $\mathcal{F}$-measurable function of $\tilde{\epsilon}_t$, because $e_{i,t}$ is a continuous function of the $\mathcal{F}$-measurable $\tilde{\epsilon}_t$ error term [White (2001)].

(v) Score function $e_{i,t}$ is strictly stationary and ergodic, because $e_{i,t}$ is an $\mathcal{F}$-measurable function of $\tilde{\epsilon}_t$, and $\tilde{\epsilon}_t$ is strictly stationary and ergodic [White (2001, Theorem 3.35)].

(vi) For $i = 1, \ldots, \mathcal{N}$, $g(e_{i,t}) \equiv \alpha_i e_{i,t-1} + \alpha_i^* \mathrm{sgn}(-\epsilon_{i,t-1})(e_{i,t-1} + 1)$ is strictly stationary and ergodic, because $g(e_{i,t})$ is an $\mathcal{F}$-measurable transformation of $\tilde{\epsilon}_t$ [White (2001, Theorem 3.35)].

4. Propositions and proofs

The following Propositions 1(a) and 1(b) use arguments from proofs of the work of Harvey (2013).

Proposition 1(a): If the maximum modulus of the eigenvalues of $A < 1$, and BD^{-1} is non-zero, then X_t, asymptotically and at the true values of parameters Θ_0, is covariance stationary.

Proof: For filter $X_t = AX_{t-1} + Bu_t$, scaled score function u_t is white noise, asymptotically and at the true values of parameters Θ_0, with zero mean and a well-defined covariance matrix for $\nu > 2$. If the maximum modulus of the eigenvalues of A is less than one and BD^{-1} is non-zero, then X_t is covariance stationary, asymptotically and at the true values of parameters Θ_0. *QED*

Proposition 1(b): For $i = 1, \ldots, \mathcal{N}$, if $|\beta_i| < 1$, and α_i or α_i^* is non-zero, then $\lambda_{i,t}$, asymptotically and at the true values of parameters Θ_0, is covariance stationary.

Proof: For filter $\lambda_{i,t} = \omega_i + \beta_i \lambda_{i,t-1} + g(e_{i,t})$, the updating terms $g(e_{i,t})$ for $i = 1, \ldots, \mathcal{N}$ are white noise, asymptotically and at the true values of parameters Θ_0, with zero mean and finite variance. If $|\beta_i| < 1$ and α_i or α_i^* is non-zero, then $\lambda_{i,t}$ is covariance stationary, asymptotically and at the true values of parameters Θ_0. *QED*

The following Propositions 2(a) and 2(b) adopt conditions from the works of Elton (1990), Alsmeyer (2003), and Gerencsér et al. (2008).

Proposition 2(a): X_t converges almost surely (a.s.) to a unique strictly stationary and ergodic vector sequence for all $\Theta \in \tilde{\Theta}$, when the following conditions hold: (i) Define $\mathcal{X}_t = \partial X_t / \partial X_{t-1}'$ ($\mathcal{N} \times \mathcal{N}$). Suppose that $E(\ln^+ ||\mathcal{X}_1||_2) < \infty$, where $\ln^+(x) = 0$ if $0 \leq x \leq 1$ and $\ln^+(x) = \ln(x)$ if $x > 1$, and $||\mathcal{W}||_2$ is the spectral norm. (ii) $E(\ln^+ ||BD^{-1}u_1||_2 < \infty$. (iii) The Lyapunov exponent is:

$$\text{Inv}_\mu = \sup_{\Theta \in \tilde{\Theta}} \left\{ \inf_{n \geq 1} \frac{1}{n} E \left[\ln \left|\left| \prod_{t=1}^{n} \mathcal{X}_t \right|\right|_2 \right] \right\} < 0 \tag{S.20}$$

(iv) $BD^{-1}u_t$ is strictly stationary and ergodic. (v) $\mathcal{X}_t$ is strictly stationary and ergodic.

Proof: (i) and (ii) hold due to the results of Section 3.1 of this Supplementary Material. (iii) is a maintained assumption. (iv) is due to the properties of the scaled score function u_t. (v) is due to the following arguments: $\mathcal{X}_t$ is an $\mathcal{F}$-measurable function of $(\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_t)$. Variable $\mathcal{X}_t$ is strictly stationary and ergodic, because $\tilde{\epsilon}_t$ is strictly stationary and ergodic [White (2001, Theorem 3.35)]. Due to the results of Elton (1990), Alsmeyer (2003), and Gerencsér et al. (2008), X_t converges a.s. to a unique strictly stationary and ergodic vector sequence for all $\Theta \in \tilde{\Theta}$. *QED*

Proposition 2(b): $\lambda_{i,t}$ for $i = 1, \ldots, \mathcal{N}$ converge a.s. to unique strictly stationary and ergodic sequences for all $\Theta \in \tilde{\Theta}$, if: (i) Define $\Lambda_{i,t} = \partial \lambda_{i,t} / \partial \lambda_{i,t-1}$ for $i = 1, \ldots, \mathcal{N}$. Suppose that $E(\ln^+ |\Lambda_{i,1}|) < \infty$ for $i = 1, \ldots, \mathcal{N}$. (ii) $E(\ln^+ |g(e_{i,1})| < \infty$ for $i = 1, \ldots, \mathcal{N}$. (iii) For $i = 1, \ldots, \mathcal{N}$, the Lyapunov exponents are:

$$\text{Inv}_{\lambda,i} = \sup_{\Theta \in \tilde{\Theta}} \left\{ \inf_{n \geq 1} \frac{1}{n} E \left[\ln \left| \prod_{t=1}^{n} \Lambda_{i,t} \right| \right] \right\} < 0 \tag{S.21}$$

(iv) $g(e_{i,t})$ for $i = 1, \dots, \mathcal{N}$ is strictly stationary and ergodic. (v) $\Lambda_{i,t}$ for $i = 1, \dots, \mathcal{N}$ is strictly stationary and ergodic.

Proof: (i) and (ii) hold due to the results of Section 3.1 of this Supplementary Material. (iii) is a maintained assumption. (iv) is due to the properties of the score functions $e_{i,t}$ for $i = 1, \dots, \mathcal{N}$. (v) is due to the following arguments: $\Lambda_{i,t}$ for $i = 1, \dots, \mathcal{N}$ are $\mathcal{F}$-measurable functions of $(\tilde{\epsilon}_1, \dots, \tilde{\epsilon}_{t-1})$. Variables $\Lambda_{i,t}$ for $i = 1, \dots, \mathcal{N}$ are strictly stationary and ergodic, because $\tilde{\epsilon}_t$ is strictly stationary and ergodic [White (2001, Theorem 3.35)]. Due to the results of Elton (1990), Alsmeyer (2003), and Gerencsér et al. (2008), $\lambda_{i,t}$ for $i = 1, \dots, \mathcal{N}$ converge a.s. to unique strictly stationary and ergodic sequences for all $\Theta \in \tilde{\Theta}$. *QED*

Proposition 3: If assumptions (A1), (A5), (A6), (A7), (A8), and (A9) hold, then $\ln f(Y_t|\mathcal{F}_{t-1}; \Theta)$ for all $\Theta \in \tilde{\Theta}$ satisfies the uniform weak law of large numbers (UWLLN) [Wooldridge (1994)].

Proof: For (A5), we use Propositions 2(a-b) and White (2001, Theorem 3.35): In the heteroskedastic t-QVAR model, X_t and $\lambda_{i,t}$, for $i = 1, \dots, \mathcal{N}$, are transformed to Y_t, using an $\mathcal{F}$-measurable function. Therefore, Y_t is strictly stationary and ergodic for $T \to \infty$. (A1), (A6), (A7), (A8), and (A9) hold for the heteroskedastic score-driven t-QVAR model. Due to the result of Wooldridge (1994, Theorem 4.1), $\ln f(Y_t|\mathcal{F}_{t-1}; \Theta)$ for all $\Theta \in \tilde{\Theta}$ satisfies the UWLLN. *QED*

Proposition 4: If the following assumptions hold: (A1), (A2), (A7), (A8), (A10), (A11), and log-density $\ln f(Y_t|\mathcal{F}_{t-1}; \Theta)$ satisfies the UWLLN, then $\hat{\Theta}$ is weakly consistent, i.e. $\hat{\Theta} \to_p \Theta_0$.

Proof: (A2) and (A11) are maintained assumptions. (A1), (A7), (A8), and (A10) hold for the heteroskedastic score-driven t-QVAR model. The assumption '$\ln f(Y_t|\mathcal{F}_{t-1}; \Theta)$ satisfies the UWLLN' holds due to Proposition 3. Due to the results of Wooldridge (1994), $\hat{\Theta} \to_p \Theta_0$. *QED*

Proposition 5: If the following assumptions hold: (A1), (A12), (A13), (A14), (A15), and (A16), then $H_t(\Theta)$ for all $\Theta \in \tilde{\Theta}$ satisfies the UWLLN.

Proof: For (A12), we use the following result for the heteroskedastic score-driven t-QVAR model: $H_t(\Theta)$ converges a.s. to a unique strictly stationary and ergodic sequence for all $\Theta \in \tilde{\Theta}$ for $T \to \infty$, which is proven in Proposition 11. (A1), (A13), (A14), and (A15) hold the heteroskedastic score-driven t-QVAR model. (A16) is maintained. Due to the results of Wooldridge (1994, Theorem 4.1), $H_t(\Theta)$ for all $\Theta \in \tilde{\Theta}$ satisfies the UWLLN. *QED*

Proposition 6: If the following assumptions hold: (A17), (A18), and (A19), then $G_t(\Theta_0)$ satisfies the central limit theorem (CLT) with asymptotic variance B_0.

Proof: For (A17), we use the following result: $E[H_t(\Theta_0)] = \text{Var}[G_t(\Theta_0)')] = E[G_t(\Theta_0)'G_t(\Theta_0)] < \infty$, where the equalities hold due to (A22) [Wooldridge (1994, p. 2674)] and (A23) [Wooldridge (1994, p. 2675)], respectively. Inequality $E[G_t(\Theta_0)'G_t(\Theta_0)] < \infty$ is shown in Proposition 9, which implies $E[G_t(\Theta_0)G_t(\Theta_0)'] < \infty$, because the terms of the sum defined by $G_t(\Theta_0)G_t(\Theta_0)'$ are in the diagonal of $G_t(\Theta_0)'G_t(\Theta_0)$. For (A18), we use the following result: $E[G_t(\Theta_0)'] = 0_{S\times 1}$, which holds under (A22) [Wooldridge (1994, p. 2674)]. For (A19) we, use White (2001, Theorem 5.16): (i) $G_t(\Theta_0)'$ is a MDS, which holds under (A3) [Wooldridge (1994, p. 2677)]. Therefore, $G_t(\Theta_0)$ is an adapted mixingale [White (2001, Definition 5.15, p. 125)]. (ii) $G_t(\Theta)'$ converges a.s. to a unique strictly stationary and ergodic sequence for all $\Theta \in \tilde{\Theta}$ for $T \to \infty$, which is proven in Proposition 10. (i) and (ii) provide (A18). Due to the results of Wooldridge (1994, Definition 4.3), $G_t(\Theta_0)$ satisfies the CLT with asymptotic variance B_0. *QED*

Proposition 7: If the following assumptions hold: (A1), (A2), (A3), (A7), (A8), (A10), (A11), (A20), (A21), (A22), (A23), (A24), $\ln f(Y_t|\mathcal{F}_{t-1};\Theta)$ satisfies the UWLLN, $H_t(\Theta)$ satisfies the UWLLN, and $G_t(\Theta_0)$ satisfies the CLT with asymptotic variance:

$$B_0 = \lim_{T\to\infty} \text{Var}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} G_t(\Theta_0)'\right], \tag{S.22}$$

then

$$\sqrt{T}(\hat{\Theta} - \Theta_0) \to_d N_S\left(0_{S\times 1}, A_0^{-1}B_0A_0^{-1}\right) = N_S\left(0_{S\times 1}, A_0^{-1}\right) \quad \text{as} \quad T \to \infty \tag{S.23}$$

The equality in Equation (S.23) is due to (A3), which provides: (i) $G_t(\Theta_0)'$ is a MDS [Wooldridge (1994, p. 2677)], (ii) $G_t(\Theta_0)'$ is serially uncorrelated [Wooldridge (1994, pp. 2676-2677)], and (iii) $A_0 = B_0$ [Wooldridge (1994, p. 2676)]. The equality in Equation (S.33) is due to (iii).

Proof: (A2), (A3), (A11), (A20), and (A24) are maintained assumptions. (A1), (A7), (A8), (A10), (A21), (A22), and (A23) hold for the score-driven ABCD representations of this paper. The assumption '$\ln f(Y_t|\mathcal{F}_{t-1};\Theta)$ satisfies the UWLLN' holds due to Proposition 3. The assumption '$H_t(\Theta)$ satisfies the UWLLN' holds due to Proposition 5. The assumption '$G_t(\Theta_0)$ satisfies the CLT asymptotic variance B_0' holds due to Proposition 6. Due to the results of Wooldridge (1994, Theorem 5.2), Equation (S.23) holds. *QED*

The following Propositions 8 and 9 use arguments from proofs of the work of Harvey (2013).

Proposition 8: For the heteroskedastic score-driven t-QVAR model, asymptotically at the true values of parameters Θ_0, the expected value of the gradient is time-invariant if Parts 1 and 2 hold. Part 1: The maximum modulus of the eigenvalues of $E(\mathcal{X}_t) < 1$, where $\mathcal{X}_t \equiv \left[A + BD^{-1}\partial u_t/\partial X'_{t-1}\right]$. Part 2: $|E(\Lambda_{i,t})| < 1$ for $i = 1, \ldots, \mathcal{N}$, where $\Lambda_{i,t} \equiv \{\beta_i + [\alpha_i + \alpha_i^* \text{sgn}(-\epsilon_{i,t-1})]\partial e_{i,t-1}/\partial\lambda_{i,t-1}\}$.

Proof: Part 1. We focus on the derivatives of X_t, with respect to an element of A, which is denoted $A_{i,j}$, and which is in $G_t(\Theta_0)'$ and $H_t(\Theta_0)$. The partial derivative of X_t, with respect to $A_{i,j}$, is:

$$\frac{\partial X_t}{\partial A_{i,j}} = \underbrace{\left[A + BD^{-1}\frac{\partial u_t}{\partial X'_{t-1}}\right]}_{\mathcal{X}_t} \frac{\partial X_{t-1}}{\partial A_{i,j}} + W_{i,j}X_{t-1} \tag{S.24}$$

where element (i, j) of matrix $W_{i,j}$ ($\mathcal{N} \times \mathcal{N}$) takes the value one and the remaining elements are zeros. We have similar first-order dynamic equations if we consider the partial derivative of X_t, with respect to the elements of BD^{-1}. The expectation of the latter equation is:

$$E\left(\frac{\partial X_t}{\partial A_{i,j}}\right) = E\left(\mathcal{X}_t\right) E\left(\frac{\partial X_{t-1}}{\partial A_{i,j}}\right) + \text{Cov}\left(\mathcal{X}_t, \frac{\partial X_{t-1}}{\partial A_{i,j}}\right) + W_{i,j}E(X_{t-1}) \tag{S.25}$$

Under the asymptotic covariance stationarity of X_t, at the true values of parameters Θ_0, the expectations on the right side of Equation (S.25) are finite due to the results of Section 3.1 of this Supplementary Material. The covariances on the right side of Equation (S.25) are finite if the variances of the random variables within those covariances are finite, which hold under the same conditions. Due to these arguments, $E\left(\partial X_t/\partial A_{i,j}\right) < \infty$ if the maximum modulus of the eigenvalues of $E\left(\mathcal{X}_t\right)$ is less than one, asymptotically at Θ_0.

Part 2. We focus on the derivative of $\lambda_{i,t}$, with respect to α_i for $i = 1, \ldots, \mathcal{N}$, which is in $G_t(\Theta_0)'$

and $H_t(\Theta_0)$. The partial derivative of $\lambda_{i,t}$, with respect to α_i, is:

$$\frac{\partial \lambda_{i,t}}{\partial \alpha_i} = \left\{ \beta_i + [\alpha_i + \alpha_i^* \text{sgn}(-\epsilon_{i,t-1})] \frac{\partial e_{i,t-1}}{\partial \lambda_{i,t-1}} \right\} \frac{\partial \lambda_{i,t-1}}{\partial \alpha_i} + e_{i,t-1} = \Lambda_{i,t} \frac{\partial \lambda_{i,t-1}}{\partial \alpha_i} + e_{i,t-1} \tag{S.26}$$

We have similar first-order dynamic equations if we consider the partial derivative of $\lambda_{i,t}$, with respect to ω_i, β_i and α_i^*. The expectation of the latter equation, that is conditional on $\mathcal{F}_{t-2}$, is:

$$E\left(\frac{\partial \lambda_{i,t}}{\partial \alpha_i} | \mathcal{F}_{t-2}\right) = E\left(\Lambda_{i,t} | \mathcal{F}_{t-2}\right) \frac{\partial \lambda_{i,t-1}}{\partial \alpha_i} + E\left(e_{i,t-1} | \mathcal{F}_{t-2}\right) \tag{S.27}$$

where $\partial \lambda_{i,t-1} / \partial \alpha_i$ is outside the conditional expectation, because it is determined by $\mathcal{F}_{t-2}$. We consider the unconditional expectation of Equation (S.27), and we start with the term

$$E\left[E\left(\Lambda_{i,t} | \mathcal{F}_{t-2}\right) \frac{\partial \lambda_{i,t-1}}{\partial \alpha_i}\right] = E\left(\Lambda_{i,t}\right) E\left(\frac{\partial \lambda_{i,t-1}}{\partial \alpha_i}\right) + \text{Cov}\left[E\left(\Lambda_{i,t} | \mathcal{F}_{t-2}\right), \frac{\partial \lambda_{i,t-1}}{\partial \alpha_i}\right] \tag{S.28}$$

We show the boundedness of all terms in the latter equation. First, $E\left(\Lambda_{i,t}\right) < \infty$, due to the boundedness of the sgn function and the boundedness of $e_{i,t-1} / \partial \lambda_{i,t-1}$. Second,

$$E\left(\frac{\partial \lambda_{i,t-1}}{\partial \alpha_i}\right) = E\left(\frac{\partial \sum_{j=0}^{\infty} \beta_i^j g(u_{t-j-2})}{\partial \alpha_i}\right) < \infty \tag{S.29}$$

where the first equation is under the covariance stationarity of $\lambda_{i,t}$ (asymptotically and at the true values of parameters Θ_0), and finiteness is due to the results of Section 3.1 of this Supplementary Material. Third, the covariance term in Equation (S.28) is bounded if the variance of both random variables within the covariance is finite. With respect to $E\left(\Lambda_{i,t} | \mathcal{F}_{t-2}\right)$,

$$\text{Var}[E\left(\Lambda_{i,t} | \mathcal{F}_{t-2}\right)] = E[E^2\left(\Lambda_{i,t} | \mathcal{F}_{t-2}\right)] - E^2[E\left(\Lambda_{i,t} | \mathcal{F}_{t-2}\right)] \tag{S.30}$$

$$= E[E^2\left(\Lambda_{i,t} | \mathcal{F}_{t-2}\right)] - E^2(\Lambda_{i,t}) \leq E[E\left(\Lambda_{i,t}^2 | \mathcal{F}_{t-2}\right)] - E^2(\Lambda_{i,t}) = E\left(\Lambda_{i,t}^2\right) - E^2(\Lambda_{i,t}) < \infty$$

where the first inequality is due to Jensen's inequality, and the second inequality is due to the results of Section 3.1 of this Supplementary Material. With respect to $\partial \lambda_{i,t-1} / \partial \alpha_i$,

$$\text{Var}\left(\frac{\partial \lambda_{i,t-1}}{\partial \alpha_i}\right) = \text{Var}\left[\frac{\partial \sum_{j=0}^{\infty} \beta_i^j g(u_{t-j-2})}{\partial \alpha_i}\right] < \infty \tag{S.31}$$

which is due to the boundedness of the sgn function and the results of Section 3.1 of this Supplementary Material. Therefore, the unconditional expectation of Equation (S.27) is

$$E\left[E\left(\frac{\partial\lambda_{i,t}}{\partial\alpha_i}|\mathcal{F}_{t-2}\right)\right]=E\left[E\left(\Lambda_{i,t}|\mathcal{F}_{t-2}\right)\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}\right]+E\left[E\left(e_{i,t-1}|\mathcal{F}_{t-2}\right)\right] \tag{S.32}$$

which is equivalent to

$$E\left(\frac{\partial\lambda_{i,t}}{\partial\alpha_i}\right)=E\left(\Lambda_{i,t}\right)E\left(\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}\right)+\mathrm{Cov}\left[E\left(\Lambda_{i,t}|\mathcal{F}_{t-2}\right),\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}\right]+E\left(e_{i,t-1}\right) \tag{S.33}$$

Due to the previous arguments, $E\left(\partial\lambda_{i,t}/\partial\alpha_i\right)<\infty$ if $|E\left(\Lambda_{i,t}\right)|<1$, asymptotically at Θ_0.

Due to the proofs of Parts 1 and 2, the expected value of the gradient it time-invariant. *QED*

Proposition 9: For the heteroskedastic score-driven t-QVAR model, asymptotically at the true values of parameters Θ_0, the expected value of the Hessian matrix is time-invariant if the following Parts 1 and 2 hold. Part 1: The maximum modulus of the eigenvalues of $E(\mathcal{X}_t\otimes\mathcal{X}_t)$, where $\otimes$ denotes the Kronecker product, is less than one. Part 2: $|E(\Lambda_{i,t}^2)|<1$ for $i=1,\ldots,\mathcal{N}$. It is enough to consider these two parts because the information matrix with respect to the parameters of X_t and $\lambda_{i,t}$ for $i=1,\ldots,\mathcal{N}$ is block diagonal.

Proof: Part 1. We focus on the derivatives of X_t, with respect to elements of A, which are denoted $A_{i,j}$ and $A_{k,l}$, and which contribute to $H_t(\Theta_0)$:

$$\frac{\partial X_t}{\partial A_{i,j}}\frac{\partial X_t'}{\partial A_{k,l}}=\mathcal{X}_t\frac{\partial X_{t-1}}{\partial A_{i,j}}\frac{\partial X_{t-1}'}{\partial A_{k,l}}\mathcal{X}_t'+\mathcal{X}_t\frac{\partial X_{t-1}}{\partial A_{i,j}}X_{t-1}'W_{k,l}' \tag{S.34}$$

$$+W_{i,j}X_{t-1}\frac{\partial X_{t-1}'}{\partial A_{k,l}}\mathcal{X}_t'+W_{i,j}X_{t-1}X_{t-1}'W_{k,l}'$$

$$\mathrm{vec}\left(\frac{\partial X_t}{\partial A_{i,j}}\frac{\partial X_t'}{\partial A_{k,l}}\right)=(\mathcal{X}_t\otimes\mathcal{X}_t)\mathrm{vec}\left(\frac{\partial X_{t-1}}{\partial A_{i,j}}\frac{\partial X_{t-1}'}{\partial A_{k,l}}\right)+W_{k,l}\mathcal{X}_t\mathrm{vec}\left(\frac{\partial X_{t-1}}{\partial A_{i,j}}X_{t-1}'\right) \tag{S.35}$$

$$+\mathcal{X}_tW_{i,j}\mathrm{vec}\left(X_{t-1}\frac{\partial X_{t-1}'}{\partial A_{k,l}}\right)+W_{k,l}W_{i,j}\mathrm{vec}\left(X_{t-1}X_{t-1}'\right)$$

The expectation of the latter equation is:

$$E\left[\mathrm{vec}\left(\frac{\partial X_t}{\partial A_{i,j}}\frac{\partial X_t'}{\partial A_{k,l}}\right)\right]=E(\mathcal{X}_t\otimes\mathcal{X}_t)E\left[\mathrm{vec}\left(\frac{\partial X_{t-1}}{\partial A_{i,j}}\frac{\partial X_{t-1}'}{\partial A_{k,l}}\right)\right] \tag{S.36}$$

$$+W_{k,l}E(\mathcal{X}_t)E\left[\text{vec}\left(\frac{\partial X_{t-1}}{\partial A_{i,j}}X'_{t-1}\right)\right]$$

$$+E(\mathcal{X}_t)W_{i,j}E\left[\text{vec}\left(X_{t-1}\frac{\partial X'_{t-1}}{\partial A_{k,l}}\right)\right]+W_{k,l}W_{i,j}E[\text{vec}\left(X_{t-1}X'_{t-1}\right)]+\text{Cov}^*$$

For the unconditional expectation of the first three terms on the right side of Equation (S.36), covariances appear in the same way as explained for Equation (S.25). We summarize those covariance terms by using the notation Cov*. The expectations and the covariances on the right side of Equation (S.36) are finite due to the asymptotic covariance stationarity of X_t at Θ_0 and the results of Section 3.1 of this Supplementary Material. Hence, $E\left\{\text{vec}\left[(\partial X_t/\partial A_{i,j})(\partial X'_t/\partial A_{k,l})\right]\right\}$ is time-invariant if the maximum modulus of eigenvalues of $E(\mathcal{X}_t\otimes\mathcal{X}_t)$ is less than one.

Part 2. We focus on the following derivative for $i=1,\ldots,\mathcal{N}$, which contributes to $H_t(\Theta_0)$:

$$\left(\frac{\partial\lambda_{i,t}}{\partial\alpha_i}\right)^2=\Lambda_{i,t}^2\left(\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}\right)^2+2\Lambda_{i,t}\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}e_{i,t-1}+e_{i,t-1}^2 \tag{S.37}$$

We have similar first-order dynamic equations if we consider the partial derivatives of $\lambda_{i,t}$, with respect to other combinations of ω_i, β_i, α_i, and α_i^*. The expectation of the latter equation, that is conditional on $\mathcal{F}_{t-2}$, is:

$$E\left[\left(\frac{\partial\lambda_{i,t}}{\partial\alpha_i}\right)^2|\mathcal{F}_{t-2}\right]=E(\Lambda_{i,t}^2|\mathcal{F}_{t-2})\left(\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}\right)^2 \tag{S.38}$$

$$+2E(\Lambda_{i,t}e_{i,t-1}|\mathcal{F}_{t-2})\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}+E(e_{i,t-1}^2|\mathcal{F}_{t-2})$$

where $\partial\lambda_{i,t-1}/\partial\alpha_i$ and $(\partial\lambda_{i,t-1}/\partial\alpha_i)^2$ are outside the conditional expectation, because they are determined by $\mathcal{F}_{t-2}$. We consider the unconditional expectation of Equation (S.38):

$$E\left[\left(\frac{\partial\lambda_{i,t}}{\partial\alpha_i}\right)^2\right]=E(\Lambda_{i,t}^2)E\left[\left(\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}\right)^2\right]+\text{Cov}\left[E(\Lambda_{i,t}^2|\mathcal{F}_{t-2}),\left(\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}\right)^2\right] \tag{S.39}$$

$$+2E(\Lambda_{i,t}e_{i,t-1})E\left(\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}\right)+2\text{Cov}\left[E(\Lambda_{i,t}e_{i,t-1}|\mathcal{F}_{t-2}),\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}\right]+E(e_{i,t-1}^2)$$

The expectations on the right side of Equation (S.39) are finite due to the results of Section 3.1 of this Supplementary Material. The covariances on the right side of Equation (S.39) are finite if the variances of the random variables within those covariances are finite. In the following we show the finiteness of those variances:

$$\text{Var}\left[E(\Lambda_{i,t}^2|\mathcal{F}_{t-2})\right] = E\left[E^2(\Lambda_{i,t}^2|\mathcal{F}_{t-2})\right] - E^2\left[E(\Lambda_{i,t}^2|\mathcal{F}_{t-2})\right] \tag{S.40}$$

$$= E\left[E^2(\Lambda_{i,t}^2|\mathcal{F}_{t-2})\right] - E^2(\Lambda_{i,t}^2) \leq E\left[E(\Lambda_{i,t}^4|\mathcal{F}_{t-2})\right] - E^2(\Lambda_{i,t}^2)$$

$$= E(\Lambda_{i,t}^4) - E^2(\Lambda_{i,t}^2) < \infty$$

$$\text{Var}\left[\left(\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}\right)^2\right] < \infty \tag{S.41}$$

$$\text{Var}\left[E(\Lambda_{i,t}e_{i,t-1}|\mathcal{F}_{t-2})\right] = E\left[E^2(\Lambda_{i,t}e_{i,t-1}|\mathcal{F}_{t-2})\right] - E^2\left[E(\Lambda_{i,t}e_{i,t-1}|\mathcal{F}_{t-2})\right] \tag{S.42}$$

$$= E\left[E^2(\Lambda_{i,t}e_{i,t-1}|\mathcal{F}_{t-2})\right] - E^2(\Lambda_{i,t}e_{i,t-1}) \leq E\left[E(\Lambda_{i,t}^2e_{i,t-1}^2|\mathcal{F}_{t-2})\right] - E^2(\Lambda_{i,t}e_{i,t-1})$$

$$= E(\Lambda_{i,t}^2e_{i,t-1}^2) - E^2(\Lambda_{i,t}e_{i,t-1}) < \infty$$

$$\text{Var}\left(\frac{\partial\lambda_{i,t-1}}{\partial\alpha_i}\right) < \infty \tag{S.43}$$

where $\leq$ in Equations (S.40) and (S.42) is due to Jensen's inequality. Equations (S.40) to (S.43) are finite due to the results of Section 3.1 of this Supplementary Material. Hence, asymptotically at Θ_0, $E\left[(\partial\lambda_{i,t}/\partial\alpha_i)^2\right]$ in Equation (S.39) is time-invariant if $|E(\Lambda_{i,t}^2)| < 1$ for $i = 1, \ldots, \mathcal{N}$.

Due to the proofs of Parts 1 and 2, the expected value of the Hessian is time-invariant. *QED*

The following Propositions 10 and 11 adopt conditions from the works of Elton (1990), Alsmeyer (2003), and Gerencsér et al. (2008).

Proposition 10: Vector $G_t(\Theta)'$ converges a.s. to a unique strictly stationary and ergodic sequence for all $\Theta \in \tilde{\Theta}$, when the conditions of the following Parts 1 and 2 hold.

<u>Part 1.</u> We repeat the dynamic equation, which contributes to the gradient:

$$\frac{\partial X_t}{\partial A_{i,j}} = \mathcal{X}_t \frac{\partial X_{t-1}}{\partial A_{i,j}} + W_{i,j}X_{t-1} \tag{S.44}$$

The conditions of Part 1 are the following: (i) Define

$$\mathcal{X}_t^{(1)} = \frac{\partial(\partial X_t/\partial A_{i,j})}{\partial(\partial X_{t-1}/\partial A_{i,j})'} \tag{S.45}$$

Suppose that $E(\ln^+ ||\mathcal{X}_1^{(1)}||_2) < \infty$. (ii) $E(\ln^+ ||X_1||_2) < \infty$. (iii) The Lyapunov exponent is:

$$\sup_{\Theta\in\tilde{\Theta}} \left\{ \inf_{n\geq 1} \frac{1}{n} E\left[\ln \left|\left| \prod_{t=1}^{n} \mathcal{X}_t^{(1)} \right|\right|_2 \right] \right\} < 0 \tag{S.46}$$

(iv) X_t is strictly stationary and ergodic. (v) $\mathcal{X}_t^{(1)}$ is strictly stationary and ergodic.

Part 2. We repeat the dynamic equation, which contributes to the gradient:

$$\frac{\partial \lambda_{i,t}}{\partial \alpha_i} = \Lambda_{i,t} \frac{\partial \lambda_{i,t-1}}{\partial \alpha_i} + e_{i,t-1} \tag{S.47}$$

for $i = 1, \ldots, \mathcal{N}$. The conditions of Part 2 are the following: (i) Define the variable $\Lambda_{i,t}^{(1)} = \partial(\partial\lambda_{i,t}/\partial\alpha_i)/\partial(\partial\lambda_{i,t-1}/\partial\alpha_i)$. Suppose that $E(\ln^+ |\Lambda_{i,1}^{(1)}|) < \infty$ for $i = 1, \ldots, \mathcal{N}$. (ii) $E(\ln^+ |e_{i,1}|) < \infty$ for $i = 1, \ldots, \mathcal{N}$. (iii) For $i = 1, \ldots, \mathcal{N}$, the Lyapunov exponents are:

$$\sup_{\Theta\in\tilde{\Theta}} \left\{ \inf_{n\geq 1} E\left[\ln \left| \prod_{t=1}^{n} \Lambda_{i,t}^{(1)} \right| \right] \right\} < 0 \tag{S.48}$$

(iv) $e_{i,t}$ for $i = 1, \ldots, \mathcal{N}$ are strictly stationary and ergodic. (v) $\Lambda_{i,t}^{(1)}$ for $i = 1, \ldots, \mathcal{N}$ are strictly stationary and ergodic.

Proof: Proof of the conditions of Part 1: (i) and (ii) hold due to the results of Section 3.1 of this Supplementary Material. (iii) is a maintained assumption. (iv) is due to the following arguments: X_{t-1} is an $\mathcal{F}$-measurable function of $(\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_{t-1})$. X_{t-1} is strictly stationary and ergodic, because $\tilde{\epsilon}_t$ is strictly stationary and ergodic [White (2001, Theorem 3.35)]. (v) is due to the following arguments: $\mathcal{X}_t^{(1)}$ is an $\mathcal{F}$-measurable function of $(\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_t)$. $\mathcal{X}_t^{(1)}$ is strictly stationary and ergodic, because $\tilde{\epsilon}_t$ is strictly stationary and ergodic [White (2001, Theorem 3.35)].

Proof of the conditions of Part 2: (i) and (ii) hold due to the results of Section 3.1 of this Supplementary Material. (iii) is a maintained assumption. (iv) is due to the properties of the score functions $e_{i,t-1}$ for $i = 1, \ldots, \mathcal{N}$. (v) is due to the following arguments: $\Lambda_{i,t}^{(1)}$ for $i = 1, \ldots, \mathcal{N}$ are

$\mathcal{F}$-measurable functions of $(\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_{t-1})$. $\Lambda_{i,t}^{(1)}$ for $i = 1, \ldots, \mathcal{N}$ are strictly stationary and ergodic, because $\tilde{\epsilon}_t$ is strictly stationary and ergodic [White (2001, Theorem 3.35)].

Due to the results of Elton (1990), Alsmeyer (2003), and Gerencsér et al. (2008), vector $G_t(\Theta)'$ converges a.s. to a unique strictly stationary and ergodic sequence for all $\Theta \in \tilde{\Theta}$. *QED*

Proposition 11: Matrix $H_t(\Theta)$ converges a.s. to a unique strictly stationary and ergodic sequence for all $\Theta \in \tilde{\Theta}$, when the conditions of the following Parts 1 and 2 hold.

Part 1. We repeat the dynamic equation, which contributes to the Hessian:

$$\text{vec}\left(\frac{\partial X_t}{\partial A_{i,j}}\frac{\partial X_t'}{\partial A_{k,l}}\right) = (\mathcal{X}_t \otimes \mathcal{X}_t)\text{vec}\left(\frac{\partial X_{t-1}}{\partial A_{i,j}}\frac{\partial X_{t-1}'}{\partial A_{k,l}}\right) + W_{k,l}\mathcal{X}_t\text{vec}\left(\frac{\partial X_{t-1}}{\partial A_{i,j}}X_{t-1}'\right) \tag{S.49}$$

$$+\mathcal{X}_t W_{i,j}\text{vec}\left(X_{t-1}\frac{\partial X_{t-1}'}{\partial A_{k,l}}\right) + W_{k,l}W_{i,j}\text{vec}\left(X_{t-1}X_{t-1}'\right)$$

$$= (\mathcal{X}_t \otimes \mathcal{X}_t)\text{vec}\left(\frac{\partial X_{t-1}}{\partial A_{i,j}}\frac{\partial X_{t-1}'}{\partial A_{k,l}}\right) + X_{t-1}^*$$

where X_{t-1}^* is defined in the last equality of Equation (S.49).

The conditions of Part 1 are the following: (i) Define

$$\mathcal{X}_t^{(2)} = \frac{\partial[(\partial X_t/\partial A_{i,j}) \times (\partial X_t/\partial A_{i,j})]}{\partial[(\partial X_{t-1}/\partial A_{i,j}) \times (\partial X_{t-1}/\partial A_{i,j})]'} \tag{S.50}$$

Suppose that $E(\ln^+ ||\mathcal{X}_1^{(2)}||_2) < \infty$. (ii) $E(\ln^+ ||X_1^*||_2) < \infty$. (iii) The Lyapunov exponent is:

$$\sup_{\Theta\in\tilde{\Theta}}\left\{\inf_{n\geq 1}E\left[\ln\left|\left|\prod_{t=1}^{n}\mathcal{X}_t^{(2)}\right|\right|_2\right]\right\} < 0 \tag{S.51}$$

(iv) X_t^* is strictly stationary and ergodic. (v) $\mathcal{X}_t^{(2)}$ is strictly stationary and ergodic.

Part 2. We repeat the dynamic equation, which contributes to the gradient:

$$\left(\frac{\partial \lambda_{i,t}}{\partial \alpha_i}\right)^2 = \Lambda_{i,t}^2\left(\frac{\partial \lambda_{i,t-1}}{\partial \alpha_i}\right)^2 + 2\Lambda_{i,t}\frac{\partial \lambda_{i,t-1}}{\partial \alpha_i}e_{i,t-1} + e_{i,t-1}^2 \tag{S.52}$$

$$= \Lambda_{i,t}^2\left(\frac{\partial \lambda_{i,t-1}}{\partial \alpha_i}\right)^2 + \Lambda_{i,t-1}^*$$

for $i = 1, \ldots, \mathcal{N}$, where $\Lambda^*_{i,t-1}$ is defined in Equation (S.52).

The conditions of Part 2 are the following: (i) Define

$$\Lambda^{(2)}_{i,t} = \frac{\partial[(\partial\lambda_{i,t}/\partial\alpha_i) \times (\partial\lambda_{i,t}/\partial\alpha_i)]}{\partial[(\partial\lambda_{i,t-1}/\partial\alpha_i) \times (\partial\lambda_{i,t-1}/\partial\alpha_i)]} \tag{S.53}$$

Suppose that $E(\ln^+ |\Lambda^{(2)}_{i,1}|) < \infty$ for $i = 1, \ldots, \mathcal{N}$. (ii) $E(\ln^+ |\Lambda^*_{i,1}|) < \infty$ for $i = 1, \ldots, \mathcal{N}$. (iii) For $i = 1, \ldots, \mathcal{N}$, the Lyapunov exponents are:

$$\sup_{\Theta\in\tilde{\Theta}} \left\{ \inf_{n\geq 1} E\left[\ln \left| \prod_{t=1}^{n} \Lambda^{(2)}_{i,t} \right| \right] \right\} < 0 \tag{S.54}$$

(iv) $\Lambda^*_{i,t}$ for $i = 1, \ldots, \mathcal{N}$ are strictly stationary and ergodic. (v) $\Lambda^{(2)}_{i,t}$ for $i = 1, \ldots, \mathcal{N}$ are strictly stationary and ergodic.

Proof: Proof of the conditions of Part 1. (i) and (ii) hold due to the results of Section 3.1 of this Supplementary Material. (iii) is a maintained assumption. (iv) is due to the following arguments: X^*_{t-1} is an $\mathcal{F}$-measurable function of $(\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_{t-1})$. X_{t-1} is strictly stationary and ergodic, because $\tilde{\epsilon}_t$ is strictly stationary and ergodic [White (2001, Theorem 3.35)]. (v) is due to the following arguments: $\mathcal{X}^{(2)}_t$ is an $\mathcal{F}$-measurable function of $(\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_t)$. $\mathcal{X}^{(2)}_t$ is strictly stationary and ergodic, because $\tilde{\epsilon}_t$ is strictly stationary and ergodic [White (2001, Theorem 3.35)].

Proof of the conditions of Part 2. (i) and (ii) hold due to the results of Section 3.1 of this Supplementary Material. (iii) is a maintained assumption. (iv) is due to the following arguments: $\Lambda^*_{i,t-1}$ for $i = 1, \ldots, \mathcal{N}$ are $\mathcal{F}$-measurable functions of $(\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_{t-1})$. $\Lambda^*_{i,t-1}$ for $i = 1, \ldots, \mathcal{N}$ are strictly stationary and ergodic, because $\tilde{\epsilon}_t$ is strictly stationary and ergodic (White 2001, Theorem 3.35). (v) is due to the following arguments: $\Lambda^{(2)}_{i,t}$ for $i = 1, \ldots, \mathcal{N}$ are $\mathcal{F}$-measurable functions of $(\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_{t-1})$. $\Lambda^{(2)}_{i,t}$ for $i = 1, \ldots, \mathcal{N}$ are strictly stationary and ergodic, because $\tilde{\epsilon}_t$ is strictly stationary and ergodic [White (2001, Theorem 3.35)].

Due to the results of Elton (1990), Alsmeyer (2003), and Gerencsér et al. (2008), matrix $H_t(\Theta)$ converges a.s. to a unique strictly stationary and ergodic sequence for all $\Theta \in \tilde{\Theta}$. *QED*

References

Alsmeyer, Gerold (2003) On the Harris recurrence of iterated random Lipschitz functions and related convergence rate results. *Journal of Theoretical Probability* 16, 217–247.

Elton, John H. (1990) A multiplicative ergodic theorem for Lipschitz maps. *Stochastic Processes and their Applications* 34, 39–47.

Gerencsér, László., György Michaletzky, and Zsanett Orlovits (2008) Stability of block-triangular stationary random matrices. *Systems & Control Letters* 57, 620–625.

Harvey, Andrew C. (2013) *Dynamic Models for Volatility and Heavy Tails.* Econometric Society Monograph. Cambridge, UK: Cambridge University Press.

White, Halbert (2001) *Asymptotic Theory for Econometricians.* San Diego: Academic Press.

Wooldridge, Jeffrey M. (1994) Estimation and inference for dependent processes. In Robert F. Engle and Daniel L. McFadden (eds.), *Handbook of Macroeconomics*, Volume IV, Chapter 45, pp. 2639–2738, Elsevier.