Discussion Papers

Department of Economics

University of Copenhagen

No. 15-07

Nonstationary ARCH and GARCH with t-Distributed Innovations

Rasmus Søndergaard Pedersen & Anders Rahbek

Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K., Denmark

Tel.: +45 35 32 30 01 – Fax: +45 35 32 30 00

http://www.econ.ku.dk

ISSN: 1601-2461 (E)

NONSTATIONARY ARCH AND GARCH WITH t-DISTRIBUTED INNOVATIONS

By Rasmus Søndergaard Pedersen† and Anders Rahbek‡

April 10, 2015

Abstract

Consistency and asymptotic normality are established for the maximum likelihood estimators in the nonstationary ARCH and GARCH models with general t-distributed innovations. The results hold for joint estimation of (G)ARCH effects and the degrees of freedom parameter parametrizing the t-distribution. With T denoting sample size, $\sqrt{T}$-convergence is shown to hold with closed form expressions for the multivariate covariances.

Keywords: ARCH, GARCH, asymptotic normality, asymptotic theory, consistency, t-distribution, maximum likelihood, nonstationarity.

JEL Classification: C32.

1 Introduction

Asymptotic theory is presented for likelihood inference in nonstationary linear ARCH and GARCH models with (scaled) t_ν-distributed innovations, where ν denotes the degrees of freedom. It is shown that when the parameters lie in the region where no stationary solution exists, maximum likelihood (ML) estimation of the (G)ARCH parameters and the degrees of freedom parameter yields estimators which are $\sqrt{T}$-consistent and have a joint normal limiting distribution. Furthermore, the covariances are shown to have closed form expressions.

Existing theory for likelihood-based estimation of nonstationary GARCH deals with Gaussian (quasi-) maximum likelihood (QML) estimation as in Jensen and Rahbek

†Department of Economics, University of Copenhagen, Denmark. ‡Department of Economics, University of Copenhagen, and CREATES, Denmark. Funding from the Danish National Research Foundation is gratefully acknowledged (Sapere Aude, Advanced Grant no. 12-124980). Correspondence to: Anders Rahbek, Department of Economics, University of Copenhagen, Øster Farimagsgade 5, 1353 Copenhagen K, Denmark; email: Anders.Rahbek@econ.ku.dk.

(2004a,b). As, by definition, the t_ν-likelihood function differs from the Gaussian, the arguments employed here are different and the results new. In particular, by exploiting results for the Beta-distribution, we are able to explicitly characterize distributional properties of key ratios appearing in the derivatives of the t_ν-likelihood function.

Theory for t_ν-likelihood estimation for the *stationary* case is given in Berkes and Horváth (2004) and Straumann (2005, Ch.6). More generally for the stationary case, asymptotic properties of estimation from maximization of the Gaussian likelihood function, that is, QML estimation, have been widely studied in the literature, see e.g. Weiss (1986), Lee and Hansen (1994), Lumsdaine (1996), Berkes, Horváth, and Kokoszka (2003), Berkes and Horváth (2003), Hall and Yao (2003), Straumann (2005) and Kristensen and Rahbek (2005). In terms of *nonstationary* QML estimation of GARCH models, in addition to Jensen and Rahbek (2004a,b), Francq and Zakoïan (2012,2013) consider nonstationary (asymmetric) GARCH QML estimation and testing for stationarity, whereas Linton, Pan, and Wang (2010) consider QML estimation allowing for dependent innovations. Moreover, Aknouche (2014) has recently proposed least-squares-based estimation of nonstationary ARCH.

To simplify the presentation, as well as the structure of the proofs of the main results in Theorems 1 and 2, initially the theory for the simple ARCH model is given in Section 2, which is next extended in Section 2.1 to GARCH. The main difference between the two cases is that the derivations for the GARCH model are notationally more involved due to the extended recursive structure as reflected in the proofs which are located in the appendix.

NOTATION: With x a positive scalar, $\Gamma(x)$ denotes the gamma function, and $\psi(x)$ and $\psi'(x)$ the digamma and trigamma functions, respectively, i.e. $\psi(x) := \mathrm{d}\ln\Gamma(x)/\mathrm{d}x$ and $\psi'(x) := \mathrm{d}^2\ln\Gamma(x)/\mathrm{d}x^2$; see Davis (1964) for more details and properties of $\Gamma(x)$. All limits are taken as $T \to \infty$ unless stated otherwise, and $\xrightarrow{P}$, $\xrightarrow{w}$ denote convergence in probability and convergence in distribution, respectively.

2 NONSTATIONARY t_ν-ARCH

Consider initially the simple ARCH model of order one introduced in Engle (1982) given by

$$x_t = \sigma_t z_t, \quad \sigma_t^2 = \omega + \alpha x_{t-1}^2, \tag{1}$$

for $t = 1, 2, ..., T$ and with the initial value x_0 fixed in the statistical analysis. The parameters α and ω are positive, $\alpha, \omega > 0$, and $\{z_t\}_{t=1,2,...}$ is a sequence of i.i.d. innovations following a scaled t-distribution with $\nu > 2$ degrees of freedom, denoted t_ν. That is, $z_t = \sqrt{(\nu-2)/\nu}\tilde{z}_t$ where $\tilde{z}_t$ has a Student's t-distribution with ν degrees of freedom, and $E[z_t] = 0$ and $E[z_t^2] = 1$.

The t_ν-log-likelihood includes ν as a parameter and is given by

$$L_T(\theta) := \frac{1}{T}\sum_{t=1}^{T} l_t(\theta), \quad l_t(\theta) := -\frac{1}{2}\ln(\sigma_t^2(\theta)) + \ln(g_\nu(x_t/\sigma_t(\theta))), \tag{2}$$

where $\theta := (\alpha, \nu, \omega)'$, the conditional variance $\sigma_t^2(\theta) := \omega + \alpha x_{t-1}^2$, and the density $g_\nu(\cdot)$ is given by

$$g_\nu(x) := \frac{\gamma(\nu)}{\sqrt{(\nu-2)\pi}}\left(1 + \frac{x^2}{\nu-2}\right)^{-\left(\frac{\nu+1}{2}\right)}, \quad \gamma(\nu) := \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}. \tag{3}$$

Let $\theta_0 := (\alpha_0, \nu_0, \omega_0)'$, where $\alpha_0, \omega_0 > 0$ and $\nu_0 > 2$, denote the true parameter value and such that no stationary solution exists, that is

$$E[\ln(\alpha_0 z_t^2)] \geq 0, \tag{4}$$

see Nelson (1990) and Bougerol and Picard (1992).

In terms of estimation, note that the ML estimator of the scale parameter ω in the nonstationary case is not consistent. This is equivalent to the nonstationary Gaussian QML case in Jensen and Rahbek (2004a), and, as there, we derive the results for estimation of α and ν with ω fixed at an arbitrary value, see Francq and Zakoïan (2012) for further considerations on this aspect.

Thus, with $\hat{\alpha}$ and $\hat{\nu}$ denoting the maximizers of $L_T(\theta)$ in (2) for any arbitrary and positive ω, we can state the following theorem for $\sqrt{T}$ asymptotic inference on α and ν in the nonstationary case:

Theorem 1 *Assume that (4) holds, then for arbitrary $\omega > 0$,*

$$\sqrt{T}(\hat{\alpha} - \alpha_0, \hat{\nu} - \nu_0) \xrightarrow{w} N(0, \Sigma^{-1}),$$

where the positive definite Σ is given by

$$\Sigma = \begin{pmatrix} \Sigma_{\alpha\alpha} & \Sigma_{\alpha\nu} \\ \Sigma_{\alpha\nu} & \Sigma_{\nu\nu} \end{pmatrix} = \begin{pmatrix} \frac{\nu_0}{2(\nu_0+3)\alpha_0^2} & \frac{3}{\alpha_0(\nu_0-2)(\nu_0+1)(\nu_0+3)} \\ \frac{3}{\alpha_0(\nu_0-2)(\nu_0+1)(\nu_0+3)} & \frac{1}{4}\left(\psi'\left(\frac{\nu_0}{2}\right) - \psi'\left(\frac{\nu_0+1}{2}\right)\right) + \frac{6}{(\nu_0-2)^2(\nu_0+1)(\nu_0+3)} \end{pmatrix}. \tag{5}$$

The proof of Theorem 1 is given in Appendix A.

Remark 2.1 The functional form of the t_ν-likelihood function is by definition different from the Gaussian likelihood considered in Jensen and Rahbek (2004a,b), with the important implication that properties of transformations of the Beta-distribution can be exploited in the proof of Theorem 1. Specifically, the following quantities appear repeatedly in the derivatives of (the log of) the t_ν-likelihood function,

$$z_{1t}^* := \frac{(\nu_0+1)z_t^2}{(\nu_0-2)+z_t^2} - 1, \; z_{2t}^* := \frac{(\nu_0+1)z_t^2}{\left[(\nu_0-2)+z_t^2\right]^2} \text{ and } z_{3t}^* := -\left(\tfrac{1}{2}\ln(1+\tfrac{z_t^2}{\nu_0-2}) - \tfrac{\partial \ln\gamma(\nu_0)}{\partial\nu}\right). \quad (6)$$

Observe that, and as used in the proof of Lemma A.5 in the appendix, it holds that $z_{1t}^* = [(\nu_0+1)\eta_t - 1]$, $z_{2t}^* = \eta_t(1-\eta_t)(\nu_0+1)/(\nu_0-2)$, and $z_{3t}^* = [\partial \ln\gamma(\nu_0)/\partial\nu + (1/2)\ln(1-\eta_t)]$, where η_t has a Beta$(1/2, \nu_0/2)$-distribution, and results for transformations of the Beta-distribution applied to z_{it}^*, $i = 1, 2, 3$, are then used to derive the explicit form of Σ in Theorem 1.

Note furthermore, that of the key quantities in (6) alone the equivalent of z_{1t}^* appears in the Gaussian (non-)stationary QML case, while z_{2t}^* and z_{3t}^* are specific to the t_ν-case. Moreover, in the Gaussian case the equivalent of z_{1t}^* takes the form $z_t^2 - 1$ corresponding to ν_0 tending to infinity and the t_ν-distribution approaching the Gaussian. Finally, note that unlike for the Gaussian case, z_{1t}^* is bounded by a constant.

Remark 2.2 Observe that only $\nu_0 > 2$, or $Ez_t^2 < \infty$, is required, which contrasts the QML estimation theory for stationary ARCH models where $Ez_t^4 < \infty$ is required corresponding to $\nu_0 > 4$; see Berkes and Horváth (2003,2004) for a discussion on general requirements for QML estimation of ARCH models.

Remark 2.3 The theorem generalizes the result for Gaussian ML estimation in Jensen and Rahbek (2004a,b); observe in particular that as $\nu_0 \to \infty$, such that the t_{ν_0} distribution tends to the standard Gaussian, the asymptotic variance of $\hat{\alpha}$, as given by $V_\alpha := \left(\Sigma_{\alpha\alpha} - \Sigma_{\alpha\nu}^2/\Sigma_{\nu\nu}\right)^{-1}$, tends to $2\alpha_0^2$ which is identical to the limiting variance of the Gaussian ML estimator.

Remark 2.4 An important result is that also the degrees of freedom MLE $\hat{\nu}$ is consistent and asymptotically Gaussian at the $\sqrt{T}$ rate. In particular so as the t_ν likelihood expansions in the ν direction are entirely different from the α direction and hence require different arguments when compared to the Gaussian ML theory.

Remark 2.5 Note the simple explicit form of the individual entries in Σ. One implication is that Σ is consistently estimated by $\hat{\Sigma}$ defined as Σ with $\hat{\alpha}, \hat{\nu}$ replacing α_0 and ν_0.

Remark 2.6 Results for consistency and asymptotic normality of the Gaussian QML estimator allow for innovations to depart from Gaussianity. Likewise, one could be interested in studying the properties of the t_ν QML estimator, that is, allow the innovations to have a non-t_ν-distribution while maximizing the t_ν likelihood function. However, as thoroughly discussed in Fan, Qi, and Xiu (2014), even for the stationary case the estimators would be inconsistent. To provide estimation theory in this case either very different model assumptions are needed such that σ_t^2 is no longer interpretable as a conditional variance as in Berkes and Horváth (2004), or the estimation procedure needs to be changed entirely, such as for the three-step type estimator in Fan, Qi, and Xiu (2014). Note that, Francq and Zakoïan (2014) informally discuss consistency of the three-step estimator for the nonstationary ARCH, and point at the open issue similar to here of the asymptotic distribution of the estimator for the nonstationary case.

2.1 Extension to t_ν-GARCH

We here present the results for the t_ν-GARCH model which generalizes the t_ν-ARCH model in (1). In short, the t_ν-GARCH model is given by

$$x_t = \sigma_t z_t, \quad \sigma_t^2 = \omega + \alpha x_{t-1}^2 + \beta \sigma_{t-1}^2,$$

for $t = 1, 2, ..., T$, with the initial value x_0 fixed in the statistical analysis, and $\{z_t\}_{t=1,2,...}$ is an i.i.d. t_ν-distributed, $\nu > 2$, sequence. Moreover, with the initial variance $\sigma_0^2 := \gamma > 0$, the parameters are given by $\theta := (\alpha, \beta, \nu, \omega, \gamma)'$, $\alpha, \beta, \omega > 0$, and the log-likelihood function is given by (2) with $\sigma_t^2(\theta) := \omega + \alpha x_{t-1}^2 + \beta \sigma_{t-1}^2(\theta)$. Let $\theta_0 := (\alpha_0, \beta_0, \nu_0, \omega_0, \gamma_0)'$, where $\alpha_0, \beta_0, \omega_0, \gamma_0 > 0$ and $\nu_0 > 2$, denote the true parameter values and such that no stationary solution exists, i.e.

$$E[\ln(\alpha_0 z_t^2 + \beta_0)] \geq 0. \tag{7}$$

We are now in position to state the equivalent of Theorem 1 for the GARCH case for arbitrary scale parameter ω and initial value γ, noting that these are inconsistently estimated:

Theorem 2 *Assume that (7) holds, then for arbitrary $\omega > 0$ and $\gamma > 0$,*

$$\sqrt{T}(\hat{\alpha} - \alpha_0, \hat{\beta} - \beta_0, \hat{\nu} - \nu_0) \xrightarrow{w} N(0, \Omega^{-1}),$$

with the positive definite Ω given in (B.1) in the appendix.

The proof of Theorem 2 is given in Appendix B.

3 Conclusion

A full theory for ML estimation in the nonstationary t_ν-ARCH and t_ν-GARCH models has been provided. We close by conjecturing that by similar arguments – but with added notational complexity – it appears possible to extend the results to allow for different functional forms of the conditional variance σ_t^2, including t_ν-(G)ARCH models of general order, possibly asymmetric.

References

Aknouche, A. (2014): "Estimation and strict stationarity testing of ARCH processes based on weighted least squares," *Mathematical Methods of Statistics*, 23(2), 81–102.

Berkes, I., and L. Horváth (2003): "The rate of consistency of the quasi-maximum likelihood estimator," *Statistics & Probability Letters*, 61(2), 133–143.

——— (2004): "The efficiency of the estimators of the parameters in GARCH processes," *The Annals of Statistics*, 32(2), 633–655.

Berkes, I., L. Horváth, and P. Kokoszka (2003): "GARCH processes: Structure and estimation," *Bernoulli*, 9(2), 201–227.

Bougerol, P., and N. Picard (1992): "Stationarity of GARCH processes and of some nonnegative time series," *Journal of Econometrics*, 52(1-2), 115–127.

Brown, B. M. (1971): "Martingale Central Limit Theorems," *The Annals of Mathematical Statistics*, 42(1), 59–66.

Davis, P. J. (1964): "Gamma function and related functions," in *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, ed. by M. Abramowitz, and I. A. Stegun, chap. 6, pp. 253–294. Dover Publications, New York, NY.

Engle, R. F. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, 50(4), 987–1007.

Fan, J., L. Qi, and D. Xiu (2014): "Quasi-maximum likelihood estimation of GARCH models with heavy-tailed likelihoods," *Journal of Business & Economic Statistics*, 32(2), 178–191.

Francq, C., and J.-M. Zakoïan (2012): "Strict stationarity testing and estimation of explosive and stationary generalized autoregressive conditional heteroscedasticity models," *Econometrica*, 80(2), 821–861.

——— (2013): "Inference in nonstationary asymmetric GARCH models," *The Annals of Statistics*, 41(4), 1970–1998.

——— (2014): "Comment," Journal of Business & Economic Statistics, 32(2), 198-201.

Hall, P., and Q. Yao (2003): "Inference in ARCH and GARCH models with heavy-tailed errors," *Econometrica*, 71(1), 285–317.

Johnson, N., A. Kemp, and S. Kotz (1995): *Univariate Continuous Distributions (vol.2)*, Wiley Series in Probability and Statistics. Wiley, New York, NY, 2 edn.

Klüppelberg, C., A. Lindner, and R. Maller (2004): "A continuous-time GARCH process driven by a Lévy process: stationarity and second-order behaviour," *Journal of Applied Probability*, 41(3), 601–622.

Kristensen, D., and A. Rahbek (2005): "Asymptotics of the QMLE for a class of ARCH(q) models," *Econometric Theory*, 21, 946–961.

Lee, S.-W., and B. E. Hansen (1994): "Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator," *Econometric Theory*, 10(1), 29–52.

Linton, O., J. Pan, and H. Wang (2010): "Estimation for a nonstationary semi-strong GARCH(1,1) model with heavy-tailed errors," *Econometric Theory*, 26, 1–28.

Lumsdaine, R. L. (1996): "Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models," *Econometrica*, 64(3), 575–596.

Nelson, D. B. (1990): "Stationarity and persistence in the GARCH(1,1) model," *Econometric Theory*, 6, 318–334.

Straumann, D. (2005): *Estimation in Conditionally Heteroscedastic Time Series Models*, Lecture Notes in Statistics. Springer, Heidelberg.

Weiss, A. A. (1986): "Asymptotic theory for ARCH models: Estimation and testing," *Econometric Theory*, 2(1), 107–131.

Appendix

For the appendix we introduce some additional notation: Let $\stackrel{a.s.}{\rightarrow}$, $\stackrel{L_p}{\rightarrow}$ denote almost sure and L_p convergence respectively, $\|\cdot\|$ the Euclidean norm, and finally $\mathbf{1}(\cdot)$ denotes the indicator function.

Note also that full expressions for all first-, second-, and third-order derivatives of the log-likelihood contribution, $l_t(\theta)$ in (2), are stated in Appendix C.

A Proof of Theorem 1

We prove Theorem 1 in two steps: First, we consider the case of ω fixed at ω_0. Next, we extend to the case of arbitrary ω.

With $\omega = \omega_0$ and $L_T(\theta)$ defined in (2), the result follows by Lemmas A.1-A.3 which provide convergence results for the score and the observed information, and establish uniform bounds for the third-order derivatives of the log-likelihood function, thereby establishing the regularity conditions for general asymptotic inference in Jensen and Rahbek (2004a, Lemma 1).

With ω arbitrary, the result follows by Lemma A.4.

A.1 Lemmas A.1-A.4

Consider first the score:

Lemma A.1 *With $L_T(\theta)$ defined in (2), the score evaluated at θ_0 is given by $\mathcal{S}_T := T^{-1/2}\sum_{t=1}^{T}(s_{t,\alpha}, s_{t,\nu})'$, with*

$$s_{t,\alpha} := \partial l_t(\theta_0)/\partial\alpha = \frac{1}{2}\frac{z_{1t}^* x_{t-1}^2}{\omega_0+\alpha_0 x_{t-1}^2} \quad \text{and} \quad s_{t,\nu} := \partial l_t(\theta_0)/\partial\nu = \frac{z_{1t}^*}{2(\nu_0-2)} + z_{3t}^*, \tag{A.1}$$

with z_{1t}^* and z_{3t}^* defined in (6). Under the nonstationarity condition in (4), as $T \to \infty$, it holds that $\mathcal{S}_T \stackrel{w}{\rightarrow} N(0,\Sigma)$, where Σ is defined in (5).

Consider next the observed information.

Lemma A.2 *With $L_T(\theta)$ defined in (2), define the observed information evaluated at θ_0 by*

$$\mathcal{I}_T := T^{-1}\sum_{t=1}^{T}\begin{pmatrix} i_{t,\alpha\alpha} & i_{t,\alpha\nu} \\ i_{t,\alpha\nu} & i_{t,\nu\nu}\end{pmatrix},$$

where $i_{t,\alpha\alpha} := -\partial^2 l_t(\theta_0)/\partial\alpha^2$, $i_{t,\alpha\nu} := -\partial^2 l_t(\theta_0)/\partial\alpha\partial\nu$ *and* $i_{t,\nu\nu} := -\partial^2 l_t(\theta_0)/\partial\nu^2$ *are defined in terms of* z_{1t}^* *and* z_{2t}^* *in (6). Under the nonstationarity condition in (4),* $\mathcal{I}_T \xrightarrow{P} \Sigma$, *with* Σ *given in (5).*

For the third-order derivatives the following uniform bounds hold.

LEMMA A.3 *With* $L_T(\theta)$ *defined in (2), for any* $\omega > 0$, *the third-order derivatives of the log-likelihood contributions, i.e.* $\partial^3 l_t(\theta)/\partial\alpha^3$, $\partial^3 l_t(\theta)/\partial\alpha^2\partial\nu$, $\partial^3 l_t(\theta)/\partial\nu^3$, *and* $\partial^3 l_t(\theta)/\partial\nu^2\partial\alpha$, *are uniformly bounded by a constant, c, for* $\lambda \in \mathcal{N}$, *where* $\mathcal{N} := [\alpha_L, \alpha_U] \times [\nu_L, \nu_U]$. *Here* $\alpha_U := \alpha_0 + \delta_\alpha$, $\alpha_L := \alpha_0 - \delta_\alpha$, $\nu_U := \nu_0 + \delta_\nu$ *and* $\nu_L := \nu_0 - \delta_\nu$ *for some* $\delta_\alpha, \delta_\nu > 0$, *and such that* $\alpha_L > 0$ *and* $\nu_L > 2$.

Finally, consider the case of arbitrary scale parameter ω:

LEMMA A.4 *With* $L_T(\theta)$ *defined in (2) and with* ω *arbitrary, there exist* α_L, α_U, ν_L, *and* ν_U, *satisfying* $\alpha_L < \alpha_0 < \alpha_U$ *and* $\nu_L < \nu_0 < \nu_U$, *such that*

$$\left\| \frac{1}{T^{1/2}} \sum_{t=1}^T \left(\frac{\partial l_t(\alpha_0,\nu_0,\omega)}{\partial(\alpha,\nu)'} - \frac{\partial l_t(\theta_0)}{\partial(\alpha,\nu)'} \right) \right\| \xrightarrow{P} 0$$

and

$$\sup_{(\alpha,\nu)' \in \mathcal{N}} \left\| \frac{1}{T} \sum_{t=1}^T \left(\frac{\partial^2 l_t(\alpha,\nu,\omega)}{\partial(\alpha,\nu)'\partial(\alpha,\nu)} - \frac{\partial^2 l_t(\alpha,\nu,\omega_0)}{\partial(\alpha,\nu)'\partial(\alpha,\nu)} \right) \right\| \xrightarrow{P} 0,$$

where $\mathcal{N} := [\alpha_L, \alpha_U] \times [\nu_L, \nu_U]$.

A.2 PROOF OF LEMMAS A.1-A.4

PROOF OF LEMMA A.1: First, observe that by definition the of $s_{t,\alpha}$ and $s_{t,v}$ in (A.1), for any $t \geq 1$,

$$|s_{t,\alpha}| \leq \tfrac{\nu_0+1}{2\alpha_0} \text{ and } |s_{t,\nu}| \leq |z_{3t}^*| + \tfrac{\nu_0+1}{2(\nu_0-2)}, \tag{A.2}$$

and in particular, $E|s_{t,j}| < \infty$ for $j = \alpha, \nu$. Lemma A.5 implies next that $(s_{t,\alpha}, s_{t,\nu})'$ is a martingale difference sequence with respect to $\mathcal{F}_t = \sigma(x_t, x_{t-1}, ..., x_0)$. Using Lemmas A.5 and A.7, we find the individual entries in Σ as follows:

$$\tfrac{1}{T} \sum_{t=1}^T E(s_{t,\alpha}^2 | \mathcal{F}_{t-1}) = \tfrac{1}{4} E\left[z_{1t}^{*2}\right] \tfrac{1}{T} \sum_{t=1}^T \left(\tfrac{x_{t-1}^2}{\omega_0 + \alpha_0 x_{t-1}^2} \right)^2 \xrightarrow{P} \tfrac{\nu_0}{2(\nu_0+3)\alpha_0^2} = \Sigma_{\alpha\alpha} > 0. \tag{A.3}$$

Similarly,

$$\tfrac{1}{T} \sum_{t=1}^T E(s_{t,\nu}^2 | \mathcal{F}_{t-1}) = E(s_{t,\nu}^2) = E[z_{3t}^{*2}] + \tfrac{1}{4(\nu_0-2)^2} E[z_{1t}^{*2}] + \left(\tfrac{1}{\nu_0-2} \right) E[z_{1t}^* z_{3t}^*] \tag{A.4}$$

$$= \tfrac{1}{4}\left[\psi'\left(\tfrac{\nu_0}{2}\right) - \psi'\left(\tfrac{\nu_0+1}{2}\right)\right] + \tfrac{6}{(\nu_0-2)^2(\nu_0+1)(\nu_0+3)} = \Sigma_{\nu\nu} > 0,$$

and

$$\tfrac{1}{T}\sum_{t=1}^{T} E(s_{t,\alpha}s_{t,\nu}|\mathcal{F}_{t-1}) = \tfrac{1}{2}E\left[\tfrac{z_{1t}^{*2}}{2(\nu_0-2)} + z_{1t}^{*}z_{3t}^{*}\right]\tfrac{1}{T}\sum_{t=1}^{T}\left(\tfrac{x_{t-1}^2}{\omega_0+\alpha_0 x_{t-1}^2}\right) \xrightarrow{P} \Sigma_{\alpha\nu}. \tag{A.5}$$

It holds that $\det(\Sigma) \geq 0$, and with equality if and only if $z_{1t}^{*} = cz_{3t}^{*}$ almost surely for some $c \in \mathbb{R}$. This is ruled out by the definition of z_{1t}^{*} and z_{3t}^{*}, and we have that $\det(\Sigma) > 0$. As $\Sigma_{\alpha\alpha}, \Sigma_{\nu\nu} > 0$, it follows that Σ is also positive definite. We may therefore conclude that for any $\phi := (\phi_1, \phi_2)' \in \mathbb{R}^2 \setminus \{(0,0)'\}$,

$$\lim_{T\to\infty}\tfrac{1}{T}\sum_{t=1}^{T} E\left[(\phi_1 s_{t,\alpha} + \phi_2 s_{t,\nu})^2|\mathcal{F}_{t-1}\right] \xrightarrow{P} \phi'\Sigma\phi > 0. \tag{A.6}$$

Turning to the Lindeberg condition, using (A.2) it follows that for any $\delta > 0$, and any $(\phi_1, \phi_2)' \in \mathbb{R}^2$,

$$\tfrac{1}{T}\sum_{t=1}^{T} E\left((\phi_1 s_{t,\alpha} + \phi_2 s_{t,\nu})^2 1\{|\phi_1 s_{t,\alpha} + \phi_2 s_{t,\nu}| > \sqrt{T}\delta\}\right) \to 0. \tag{A.7}$$

This establishes the CLT in Brown (1971). $\square$

Proof of Lemma A.2: Using (C.5) in Appendix C, we have that

$$\tfrac{1}{T}\sum_{t=1}^{T} i_{t,\alpha\alpha} = \tfrac{1}{T}\sum_{t=1}^{T}\tfrac{1}{2}\left[z_{1t}^{*} + z_{2t}^{*}(\nu_0-2)\right]\left(\tfrac{x_{t-1}^2}{\omega_0+\alpha_0 x_{t-1}^2}\right)^2.$$

By the strong law of large numbers for i.i.d. processes, $\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T}|\tfrac{1}{2}\left[z_{1t}^{*} + z_{2t}^{*}(\nu_0-2)\right]|$ is almost surely bounded by some finite constant c. Moreover, for any fixed t, it holds by the Borel-Cantelli lemma that $T^{-1}\tfrac{1}{2}\left[z_{1t}^{*} + z_{2t}^{*}(\nu_0-2)\right] \xrightarrow{a.s.} 0$. Then by applying Toeplitz' lemma together with Lemmas A.5 and A.7,

$$\tfrac{1}{T}\sum_{t=1}^{T} i_{t,\alpha\alpha} \xrightarrow{P} \tfrac{1}{2}\left\{E[z_{1t}^{*}] + E[z_{2t}^{*}](\nu_0-2)\right\}\tfrac{1}{\alpha_0^2} = \tfrac{\nu_0}{2\alpha_0^2(\nu_0+3)} = \Sigma_{\alpha\alpha}.$$

By similar arguments, and using (C.8) and (C.6),

$$\begin{aligned}\tfrac{1}{T}\sum_{t=1}^{T} i_{t,\alpha\nu} &= \tfrac{1}{T}\sum_{t=1}^{T} -\tfrac{1}{2}\left\{(\nu_0+1)^{-1}(z_{1t}^{*}+1) - z_{2t}^{*}\right\}\left(\tfrac{x_{t-1}^2}{\omega_0+\alpha_0 x_{t-1}^2}\right) \\ &\xrightarrow{P} -\tfrac{1}{2}\left\{(\nu_0+1)^{-1}E[z_{1t}^{*}+1] - E[z_{2t}^{*}]\right\}\tfrac{1}{\alpha_0} = \Sigma_{\alpha\nu},\end{aligned}$$

and

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} i_{t,\nu\nu} &= \frac{1}{T}\sum_{t=1}^{T}\left\{-\frac{\partial^2 \ln\gamma(\nu)}{\partial\nu\partial\nu} - \frac{z_{1t}^*+1}{(\nu_0-2)(\nu_0+1)} + \frac{z_{1t}^*}{2(\nu_0-2)^2} + \frac{z_{2t}^*}{2(\nu_0-2)}\right\} \\
&\xrightarrow{P} -\frac{\partial^2 \ln\gamma(\nu)}{\partial\nu\partial\nu} - \frac{E[z_{1t}^*]+1}{(\nu_0-2)(\nu_0+1)} + \frac{E[z_{1t}^*]}{2(\nu_0-2)^2} + \frac{E[z_{2t}^*]}{2(\nu_0-2)} \\
&= \frac{1}{4}\left[\psi'\left(\frac{\nu_0}{2}\right) - \psi'\left(\frac{\nu_0+1}{2}\right)\right] - \frac{1}{(\nu_0-2)(\nu_0+1)} + \frac{\nu_0}{(3+\nu_0)(\nu_0-2)^2} = \Sigma_{\nu\nu}.
\end{aligned}
$$

□

Proof of Lemma A.3: From (C.14), for any $\omega > 0$,

$$
\sup_{(\alpha,\nu)'\in\mathcal{N}}\left|\frac{\partial^3 l_t(\alpha,\nu,\omega)}{\partial\alpha^3}\right| \le \left[(\nu_U+1) + \frac{(\nu_U+1)(\nu_U-2)}{(\nu_L-2)} + \frac{(\nu_U+1)(\nu_U-2)^2}{(\nu_L-2)^2} + 1\right]\frac{1}{\alpha_L^3} \le c < \infty.
$$

Likewise, using (C.15), (C.21), and (C.22),

$$
\sup_{(\alpha,\nu)'\in\mathcal{N}}\left|\frac{\partial^3 l_t(\alpha,\nu,\omega)}{\partial\alpha^2\partial\nu}\right| \le \left[\frac{1}{2} + \frac{(\nu_U-2)}{2(\nu_L-2)} + \frac{(\nu_U+1)(\nu_U-2)}{(\nu_L-2)^2}\right]\frac{1}{\alpha_L^2} \le c < \infty,
$$

$$
\sup_{(\alpha,\nu)'\in\mathcal{N}}\left|\frac{\partial^3 l_t(\alpha,\nu,\omega)}{\partial\nu^3}\right| \le \sup_{\nu\in[\nu_L,\nu_U]}\left|\frac{\partial^3 \ln\gamma(\nu)}{\partial\nu^3}\right| + \frac{3}{(\nu_L-2)^2} + \frac{4(\nu_U+1)}{(\nu_L-2)^3} \le c < \infty,
$$

and

$$
\sup_{(\alpha,\nu)'\in\mathcal{N}}\left|\frac{\partial^3 l_t(\alpha,\nu,\omega)}{\partial\nu^2\partial\alpha}\right| \le \left[\frac{(\nu_U+1)}{(\nu_L-2)^2} + \frac{1}{\nu_L-2}\right]\frac{1}{\alpha_L} \le c < \infty.
$$

□

Proof of Lemma A.4: Choose $\alpha_L > 0$ and $\nu_L > 2$, and define $\omega_L := \min(\omega_0, \omega)$ and $\omega_U := \max(\omega_0, \omega)$. Using Taylor expansions, it suffices to show that

$$
\sup_{\omega_L\le\omega\le\omega_U}\left\|\frac{1}{T^{1/2}}\sum_{t=1}^{T}\frac{\partial^2 l_t(\alpha_0,\nu_0,\omega)}{\partial(\alpha,\nu)'\partial\omega}\right\| \xrightarrow{P} 0 \tag{A.8}
$$

and

$$
\sup_{(\alpha,\nu)'\in\mathcal{N}}\sup_{\omega_L\le\omega\le\omega_U}\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^3 l_t(\theta)}{\partial(\alpha,\nu)'\partial(\alpha,\nu)\partial\omega}\right\| \xrightarrow{P} 0. \tag{A.9}
$$

From (C.10),

$$
\begin{aligned}
\sup_{\omega_L\le\omega\le\omega_U}\left|\frac{\partial^2 l_t(\alpha_0,\nu_0,\omega)}{\partial\alpha\partial\omega}\right| \quad &\le \quad \sup_{\omega_L\le\omega\le\omega_U}\frac{1}{2}\left|1 - \frac{(\nu_0+1)x_t^2/\sigma_t^2(\alpha_0,\nu_0,\omega)}{(\nu_0-2)+x_t^2/\sigma_t^2(\alpha_0,\nu_0,\omega)} - \frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\alpha_0,\nu_0,\omega)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\alpha_0,\nu_0,\omega)\right]^2}\right| \\
&\quad \times \sup_{\omega_L\le\omega\le\omega_U}\left|\left(\frac{x_{t-1}^2}{\sigma_t^2(\alpha_0,\nu_0,\omega)}\right)\left(\frac{\sigma_t^2(\alpha_0,\nu_0,\omega_0)}{\sigma_t^2(\alpha_0,\nu_0,\omega)}\right)\left(\frac{1}{\sigma_t^2(\theta_0)}\right)\right| \\
&\le \quad \frac{1}{2}\left[1 + (\nu_0+1) + \frac{(\nu_0+1)(\nu_0-2)}{(\nu_0-2)}\right]\frac{1}{\alpha_0}\left(\frac{\omega_0}{\omega_L}+1\right)\left(\frac{1}{\sigma_t^2(\theta_0)}\right),
\end{aligned}
$$

and an application of Lemma A.8 and Jensen and Rahbek (2004a, Lemma 11) then gives that $\sup_{\omega_L \leq \omega \leq \omega_U} T^{-1/2} \sum_{t=1}^{T} |\partial^2 l_t(\alpha_0, \nu_0, \omega)/\partial\alpha\partial\omega| \xrightarrow{P} 0$. Similar arguments, and using (C.12), yield that $\sup_{\omega_L \leq \omega \leq \omega_U} T^{-1/2} \sum_{t=1}^{T} |\partial^2 l_t(\alpha_0, \nu_0, \omega)/\partial\nu\partial\omega| \xrightarrow{P} 0$, and we have that (A.8) holds. Turning to the proof of (A.9), it holds by (C.23) that

$$\begin{aligned}
&\sup_{(\alpha,\nu)' \in \mathcal{N}} \sup_{\omega_L \leq \omega \leq \omega_U} \left| \tfrac{\partial^3 l_t(\theta)}{\partial\alpha^2\partial\omega} \right| \\
\leq\ & \sup_{(\alpha,\nu)' \in \mathcal{N}} \sup_{\omega_L \leq \omega \leq \omega_U} \left| \tfrac{(\nu+1)(\nu-2)^2 x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3} + \tfrac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} + \tfrac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2} - 1 \right| \\
& \times \sup_{(\alpha,\nu)' \in \mathcal{N}} \sup_{\omega_L \leq \omega \leq \omega_U} \left| \left(\tfrac{x_{t-1}^2}{\sigma_t^2(\theta)} \right)^2 \left(\tfrac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \right) \left(\tfrac{1}{\sigma_t^2(\theta_0)} \right) \right| \\
\leq\ & \left(\tfrac{(\nu_U+1)(\nu_U-2)^2}{(\nu_L-2)^2} + (\nu_U+1) + \tfrac{(\nu_U+1)(\nu_U-2)}{(\nu_L-2)} + 1 \right) \tfrac{1}{\alpha_L^2} \left(\tfrac{\alpha_0}{\alpha_L} + \tfrac{\omega_0}{\omega_L} \right) \left(\tfrac{1}{\sigma_t^2(\theta_0)} \right),
\end{aligned}$$

so another application of Lemma A.8 and Jensen and Rahbek (2004a, Lemma 11) yields that $\sup_{(\alpha,\nu)' \in \mathcal{N}} \sup_{\omega_L \leq \omega \leq \omega_U} T^{-1} \sum_{t=1}^{T} |\partial^3 l_t(\theta)/\partial\alpha^2\partial\omega| \xrightarrow{P} 0$. Similar arguments, and using (C.26) and (C.28), give that $\sup_{(\alpha,\nu)' \in \mathcal{N}} \sup_{\omega_L \leq \omega \leq \omega_U} T^{-1} \sum_{t=1}^{T} |\partial^3 l_t(\theta)/\partial\alpha\partial\nu\partial\omega| \xrightarrow{P} 0$ and $\sup_{(\alpha,\nu)' \in \mathcal{N}} \sup_{\omega_L \leq \omega \leq \omega_U} T^{-1} \sum_{t=1}^{T} |\partial^3 l_t(\theta)/\partial\nu^2\partial\omega| \xrightarrow{P} 0$, and we conclude that (A.9) holds. $\square$

A.3 Auxiliary Lemmas: t_ν-ARCH

We state here the expectation and (co)variances of z_{it}^*, $i = 1, 2, 3$, defined in (6). Moreover, we state important convergence results for ratios of the type $x_{t-1}^{2m}/(\omega_0 + \alpha_0 x_{t-1}^2)^k$, with k and m nonnegative integers. These properties are repeatedly used in the proofs.

Lemma A.5 *With z_{1t}^*, z_{2t}^*, and z_{3t}^* defined in (6),*

(i) $E[z_{1t}^*] = E[z_{3t}^*] = 0$ *and* $E[z_{2t}^*] = \nu_0/[(\nu_0+3)(\nu_0-2)]$,

(ii) $E[z_{1t}^{*2}] = 2\nu_0/(\nu_0+3)$ *and* $E[z_{3t}^{*2}] = [\psi'(\nu_0/2) - \psi'((\nu_0+1)/2)]/4$,

(iii) $E[z_{1t}^* z_{3t}^*] = -(\nu_0+1)^{-1}$.

Proof of Lemma A.5: Notice that $z_t = \sqrt{(\nu_0-2)/\nu_0}\,\tilde{z}_t$, where $\tilde{z}_t$ has a Student's t-distribution with ν degrees of freedom. It holds that $\eta_t := z_t^2/((\nu_0-2) + z_t^2) = (\tilde{z}_t^2/\nu_0)/(1 + \tilde{z}_t^2/\nu_0)$ has a Beta$(1/2, \nu_0/2)$-distribution, see e.g. Johnson, Kemp, and Kotz (1995, p.327). Hence, $z_{1t}^* = [(\nu_0+1)\eta_t - 1]$, $z_{2t}^* = \eta_t(1-\eta_t)(\nu_0+1)/(\nu_0-2)$, and $z_{3t}^* = [\partial \ln\gamma(\nu_0)/\partial\nu + (1/2)\ln(1-\eta_t)]$. The results then follow by using the results for the Beta-distribution, listed in Lemma A.6 below. $\square$

Lemma A.6 *Assume that the random variable X is Beta-distributed with shape parameters $p, q > 0$, i.e. Beta(p, q). Then*

(i) $E[X] = p/(p+q)$ *and* $E[X^2] = p(p+q)^{-1}(p+1)(p+q+1)^{-1}$,

(ii) $E[\ln(1-X)] = \psi(q) - \psi(p+q)$ *and* $E[\ln(1-X)]^2 = \psi'(q) - \psi'(p+q) + (\psi(q) - \psi(p+q))^2$,

(iii) $E[X \ln(1-X)] = [\psi(q) - \psi(p+q) - (p+q)^{-1}]p(p+q)^{-1}$.

Proof of Lemma A.6: The results in (i) are well-known, see for example Johnson, Kemp, and Kotz (1995, p.217). The results in (ii)-(iii) follow by using the density function of the Beta-distribution, and by repeated use of the identity $\frac{\partial}{\partial y} x^y = (\ln x)\, x^y$, $x > 0$. $\square$

Lemma A.7 *Assume that (4) holds. For $m, k \in \mathbb{N} \cup \{0\}$, $m \leq k$, and with the ratios $\gamma_t\,(m,k)$ defined by*

$$\gamma_t\,(m,k) := \tfrac{x_{t-1}^{2m}}{(\omega_0 + \alpha_0 x_{t-1}^2)^k}, \tag{A.10}$$

it holds that $\gamma_t\,(m,k) \xrightarrow{P} \alpha_0^{-m} \mathbf{1}\,(m = k)$, as $t \to \infty$. Moreover, $\frac{1}{T}\sum_{t=1}^{T} \gamma_t\,(m,k) \xrightarrow{P} \alpha_0^{-m} \mathbf{1}\,(m = k)$ as $T \to \infty$.

Proof of Lemma A.7: The results hold as $x_{t-1}^2 \xrightarrow{P} \infty$ under (4) by Theorem 2.1.b in Klüppelberg, Lindner, and Maller (2004). $\square$

Lemma A.8 *Under the nonstationarity condition in (4) there exists a $\rho < 1$ such that for $t \geq 1$,*

$$E\left[\tfrac{1}{\sigma_t^2(\theta_0)}\right] \leq \omega_0^{-1} \rho^{t-1}.$$

Proof of Lemma A.8: Notice that for any $t \geq 1$,

$$\sigma_t^2(\theta_0) = x_0^2 \prod_{i=1}^{t-1} \alpha_0 z_{t-i}^2 + \omega_0 \left[1 + \sum_{k=1}^{t-1} \prod_{i=1}^{k} \alpha_0 z_{t-i}^2\right] \geq \omega_0 \prod_{i=1}^{t-1} \alpha_0 z_{t-i}^2, \tag{A.11}$$

with the conventions $\prod_{i=1}^{0} = 1$ and $\sum_{k=1}^{0} = 0$. When (4) holds, $1 \leq \exp(E[\ln(\alpha_0 z_t^2)]) < E[\alpha_0 z_t^2] (= \alpha_0)$, where the strict inequality holds by Jensen's inequality and the fact that the exponential function is strictly convex and that $\ln(\alpha_0 z_t^2)$ is non-degenerate. By another application of Jensen's inequality (for strictly concave functions), we then have that

$$E\left[1/(\alpha_0 z_t^2)\right] < 1. \tag{A.12}$$

Combining (A.11) and (A.12) yields the result. $\square$

B The t_ν-GARCH Case

We proceed as for the ARCH case by considering first the case of fixed (ω, γ) at (ω_0, γ_0), and next allow for arbitrary (ω, γ). Thus with $(\omega, \gamma) = (\omega_0, \gamma_0)$ and $L_T(\theta)$ defined in (2), the result follows by Lemmas B.1–B.3 establishing Jensen and Rahbek (2004a, Lemma 1). With ω and γ arbitrary, apply Lemma B.4, as in Jensen and Rahbek (2004a, Sections 4.2-4.3).

B.1 Lemmas B.1-B.4

Lemma B.1 *The score evaluated at θ_0 is given by $\mathcal{S}_T := T^{-1/2}\sum_{t=1}^T (s_{t,\alpha}, s_{t,\beta}, s_{t,\nu})'$, with*

$$s_{t,\alpha} := \tfrac{z_{1t}^*}{2}\tfrac{\partial\sigma_t^2(\theta_0)/\partial\alpha}{\sigma_t^2(\theta_0)},\; s_{t,\beta} := \tfrac{z_{1t}^*}{2}\tfrac{\partial\sigma_t^2(\theta_0)/\partial\beta}{\sigma_t^2(\theta_0)},\; \textit{and } s_{t,\nu} := \tfrac{z_{1t}^*}{2(\nu_0-2)} + z_{3t}^*.$$

Under the nonstationarity condition in (7), $\mathcal{S}_T \xrightarrow{w} N(0, \Omega)$, where the positive definite Ω is given by

$$\Omega = \begin{pmatrix} \Omega_{\alpha\alpha} & \Omega_{\alpha\beta} & \Omega_{\alpha\nu} \\ \Omega_{\alpha\beta} & \Omega_{\beta\beta} & \Omega_{\beta\nu} \\ \Omega_{\alpha\nu} & \Omega_{\beta\nu} & \Omega_{\nu\nu} \end{pmatrix}, \tag{B.1}$$

where

$$\begin{aligned}
\Omega_{\alpha\alpha} &= \tfrac{\nu_0}{2(\nu_0+3)\alpha_0^2},\quad \Omega_{\beta\beta} = \tfrac{\nu_0(1+\mu_1)\mu_2}{2(\nu_0+3)\beta_0^2(1-\mu_1)(1-\mu_2)},\quad \mu_i = E[(\beta_0/(\alpha_0 z_t^2 + \beta_0))^i],\ i = 1, 2,\\
\Omega_{\nu\nu} &= \tfrac{1}{4}\left[\psi'\left(\tfrac{\nu_0}{2}\right) - \psi'\left(\tfrac{\nu_0+1}{2}\right)\right] + \tfrac{6}{(\nu_0-2)^2(\nu_0+1)(\nu_0+3)},\quad \Omega_{\alpha\beta} = \tfrac{\nu_0\mu_1}{2(\nu_0+3)\alpha_0\beta_0(1-\mu_1)},\\
\Omega_{\alpha\nu} &= \tfrac{3}{(\nu_0-2)(\nu_0+1)(\nu_0+3)\alpha_0} \quad \textit{and} \quad \Omega_{\beta\nu} = \tfrac{3\mu_1}{(\nu_0-2)(\nu_0+1)(\nu_0+3)\alpha_0\beta_0(1-\mu_1)}.
\end{aligned}$$

Lemma B.2 *Define the observed information evaluated at θ_0 by*

$$\mathcal{I}_T := T^{-1}\sum_{t=1}^T \begin{pmatrix} i_{t,\alpha\alpha} & i_{t,\alpha\beta} & i_{t,\alpha\nu} \\ i_{t,\alpha\beta} & i_{t,\beta\beta} & i_{t,\beta\nu} \\ i_{t,\alpha\nu} & i_{t,\beta\nu} & i_{t,\nu\nu} \end{pmatrix},$$

where $i_{t,\alpha\alpha} := -\partial^2 l_t(\theta_0)/\partial\alpha^2$, $i_{t,\beta\beta} := -\partial^2 l_t(\theta_0)/\partial\beta^2$ $i_{t,\nu\nu} := -\partial^2 l_t(\theta_0)/\partial\nu^2$, $i_{t,\alpha\nu} := -\partial^2 l_t(\theta_0)/\partial\alpha\partial\nu$, $i_{t,\alpha\beta} := -\partial^2 l_t(\theta_0)/\partial\alpha\partial\beta$, and $i_{t,\beta\nu} := -\partial^2 l_t(\theta_0)/\partial\beta\partial\nu$. Under the nonstationarity condition in (7), $\mathcal{I}_T \xrightarrow{P} \Omega$, with Ω defined in (B.1).

In the following we define the neighborhood, $\mathcal{N}(\theta_0)$, around θ_0 as

$$\mathcal{N}(\theta_0) := \{\theta : \alpha_L \le \alpha \le \alpha_U,\ \beta_L \le \beta \le \beta_U,\ \omega_L \le \omega \le \omega_U,\ \gamma_L \le \gamma \le \gamma_U, \nu_L \le \nu \le \nu_U\}, \tag{B.2}$$

for some $\alpha_L < \alpha_0 < \alpha_U$, $\beta_L < \beta_0 < \beta_U$, $\omega_L < \omega_0 < \omega_U$, $\gamma_L < \gamma_0 < \gamma_U$, and $\nu_L < \nu_0 < \nu_U$.

Lemma B.3 *With $(\lambda_1, \lambda_2, \lambda_3) := (\alpha, \beta, \nu)$, there exists a neighborhood, $\mathcal{N}(\theta_0)$ as in (B.2) such that*

$$\max_{h,i,j=1,2,3} \sup_{\theta\in\mathcal{N}(\theta_0)} \left| \frac{\partial^3 L_T(\theta)}{\partial\lambda_h\partial\lambda_i\partial\lambda_j} \right| \leq \tilde{w}_T, \tag{B.3}$$

where $0 \leq \tilde{w}_T \xrightarrow{P} c \in (0, \infty)$.

Next, let $\lambda := (\alpha, \beta, \nu)'$ and define the neighborhood, $\mathcal{N}(\lambda_0)$, around $\lambda_0 := (\alpha_0, \beta_0, \gamma_0)'$ as

$$\mathcal{N}(\lambda_0) := \{\lambda = (\alpha, \beta, \nu)' : \alpha_L \leq \alpha \leq \alpha_U,\ \beta_L \leq \beta \leq \beta_U,\ \nu_L \leq \nu \leq \nu_U\}, \tag{B.4}$$

for some $\alpha_L < \alpha_0 < \alpha_U$, $\beta_L < \beta_0 < \beta_U$, and $\nu_L < \nu_0 < \nu_U$.

Lemma B.4 *Under the nonstationarity condition (7), there exists a neighborhood, $\mathcal{N}(\lambda_0)$, as in (B.4) such that for any fixed $\omega, \gamma > 0$,*

$$\left\| T^{-1/2} \sum_{t=1}^{T} \left(\frac{\partial l_t(\lambda_0, \omega, \gamma)}{\partial\lambda} - \frac{\partial l_t(\theta_0)}{\partial\lambda} \right) \right\| \xrightarrow{P} 0$$

and

$$\sup_{\lambda\in\mathcal{N}(\lambda_0)} \left\| T^{-1} \sum_{t=1}^{T} \left(\frac{\partial^2 l_t(\lambda, \omega, \gamma)}{\partial\lambda\partial\lambda'} - \frac{\partial^2 l_t(\lambda, \omega_0, \gamma_0)}{\partial\lambda\partial\lambda'} \right) \right\| \xrightarrow{P} 0.$$

B.2 Proofs of Lemmas B.1-B.4:

Proof of Lemma B.1: By Lemmas A.5 and B.5, $(s_{t,\alpha}, s_{t,\beta}, s_{t,\nu})'$ is a martingale difference with respect to $\mathcal{F}_t := \sigma(y_t, y_{t-1}, ..., y_0)$, $y_t := (x_t, \sigma_t^2(\theta_0))'$. With $u_{\alpha t}$ defined in (B.11) in Lemma B.5, it holds that

$$\frac{1}{T} \sum_{t=1}^{T} E(s_{t,\alpha}^2 | \mathcal{F}_{t-1}) = \frac{1}{4} E\left[z_{1t}^{*2}\right] \frac{1}{T} \sum_{t=1}^{T} \left(\frac{\partial\sigma_t^2(\theta_0)/\partial\alpha}{\sigma_t^2(\theta_0)} \right)^2 =$$
$$\frac{\nu_0}{2(\nu_0+3)} \left\{ \frac{1}{T} \sum_{t=1}^{T} \left[\left(\frac{\partial\sigma_t^2(\theta_0)/\partial\alpha}{\sigma_t^2(\theta_0)} \right) - u_{\alpha t}^2 \right] + \frac{1}{T} \sum_{t=1}^{T} u_{\alpha t}^2 \right\} \xrightarrow{P} \frac{\nu_0}{2(\nu_0+3)} E[u_{\alpha t}^2] = \Omega_{\alpha\alpha},$$

where we have used Lemma B.5 and the ergodic theorem. Likewise, with $u_{\beta t}$ defined in (B.12),

$$\frac{1}{T} \sum_{t=1}^{T} E(s_{t,\beta}^2 | \mathcal{F}_{t-1}) \xrightarrow{P} \frac{\nu_0}{2(\nu_0+3)} E[u_{\beta t}^2] = \frac{\nu_0}{2(\nu_0+3)} \frac{(1+\mu_1)\mu_2}{\beta_0^2(1-\mu_1)(1-\mu_2)} = \Omega_{\beta\beta},$$

and

$$\frac{1}{T}\sum_{t=1}^{T}E(s_{t,\alpha}s_{t,\beta}|\mathcal{F}_{t-1})\xrightarrow{P}\frac{\nu_0}{2(\nu_0+3)}E[u_{\alpha t}u_{\beta t}]=\frac{\nu_0}{2(\nu_0+3)}\frac{\mu_1}{\alpha_0\beta_0(1-\mu_1)}=\Omega_{\alpha\beta}.$$

Similar to the proof of Lemma 1,

$$\frac{1}{T}\sum_{t=1}^{T}E(s_{t,\nu}^2|\mathcal{F}_{t-1})=\frac{1}{4}\left[\psi'\left(\frac{\nu_0}{2}\right)-\psi'\left(\frac{\nu_0+1}{2}\right)\right]+\frac{6}{(\nu_0-2)^2(\nu_0+1)(\nu_0+3)}=\Omega_{\nu\nu},$$

$$\frac{1}{T}\sum_{t=1}^{T}E(s_{t,\alpha}s_{t,\nu}|\mathcal{F}_{t-1})\xrightarrow{P}\frac{1}{2}E\left[\frac{z_{1t}^{*2}}{2(\nu_0-2)}+z_{1t}^{*}z_{3t}^{*}\right]E[u_{\alpha t}]=\frac{3}{(\nu_0-2)(\nu_0+1)(\nu_0+3)}\alpha_0^{-1}=\Omega_{\alpha\nu},$$

and

$$\frac{1}{T}\sum_{t=1}^{T}E(s_{t,\beta}s_{t,\nu}|\mathcal{F}_{t-1})\xrightarrow{P}\frac{1}{2}E\left[\frac{z_{1t}^{*2}}{2(\nu_0-2)}+z_{1t}^{*}z_{3t}^{*}\right]E[u_{\beta t}]=\frac{3}{(\nu_0-2)(\nu_0+1)(\nu_0+3)}\frac{\mu_1}{\alpha_0\beta_0(1-\mu_1)}=\Omega_{\beta\nu}.$$

By construction, Ω is positive semi-definite, and next we seek to show that it is in fact positive definite. Notice that $\Omega=\mathrm{Var}[(z_{1t}^{*}u_{\alpha t}/2,z_{1t}^{*}u_{\beta t}/2,z_{1t}^{*}/(2\,(\nu_0-2))+z_{3t}^{*})']$, and it hence suffices to show that we cannot find a constant vector $\phi:=(\phi_1,\phi_2,\phi_3)'\in\mathbb{R}^3\setminus\{(0,0,0)'\}$ such that

$$\phi'(z_{1t}^{*}u_{\alpha t}/2,z_{1t}^{*}u_{\beta t}/2,z_{1t}^{*}/(2\,(\nu_0-2))+z_{3t}^{*})=0\quad\text{a.s.}$$

The proof follows by contradiction. First, suppose that $(\phi_1,\phi_2)\neq(0,0)$ and $\phi_3=0$, which means that $\phi_1z_{1t}^{*}u_{\alpha t}/2+\phi_2z_{1t}^{*}u_{\beta t}/2=0$ a.s. Since $P(z_{1t}^{*}=0)=0$, $\phi_1u_{\alpha t}+\phi_2u_{\beta t}=0$ a.s., but this is clearly ruled out since $u_{\alpha t}$ and $u_{\beta t}$ are linearly independent. Next, suppose that $(\phi_1,\phi_2)=(0,0)$ and $\phi_3\neq0$. Then $\phi_3(z_{1t}^{*}/(2\,(\nu_0-2))+z_{3t}^{*})=0$ a.s., which is ruled out since $P(z_{1t}^{*}/(2\,(\nu_0-2))+z_{3t}^{*}=0)=0$. So it must hold that $(\phi_1,\phi_2)\neq(0,0)$ and $\phi_3\neq0$. Using again that $P(z_{1t}^{*}=0)=0$, $\phi_1u_{\alpha t}/2+\phi_2u_{\beta t}/2=\phi_3(1/(2\,(\nu_0-2))+z_{3t}^{*}/z_{1t}^{*})$ a.s. This is ruled out by the fact that z_{3t}^{*}/z_{1t}^{*} is non-degenerate and independent of $(u_{\alpha t},u_{\beta t})$. We conclude that Ω is positive definite, and hence for any $\phi=(\phi_1,\phi_2,\phi_3)'\in\mathbb{R}^3\setminus\{(0,0,0)'\}$,

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E\left[(\phi_1s_{t,\alpha}+\phi_2s_{t,\beta}+\phi_3s_{t,\nu})^2|\mathcal{F}_{t-1}\right]\xrightarrow{P}\phi'\Omega\phi>0.$$

Turning to the Lindeberg condition, using Lemma B.5 it follows that for any $\delta>0$, and any $(\phi_1,\phi_2,\phi_3)'\in\mathbb{R}^3$,

$$\frac{1}{T}\sum_{t=1}^{T}E\left((\phi_1s_{t,\alpha}+\phi_2s_{t,\beta}+\phi_3s_{t,\nu})^2 1\{|\phi_1s_{t,\alpha}+\phi_2s_{t,\nu}|>\sqrt{T}\delta\}\right)\to0.$$

This establishes the CLT in Brown (1971). $\square$

Proof of Lemma B.2: From (C.4), we have that

$$
\begin{aligned}
T^{-1}\sum_{t=1}^{T} i_{t,\beta\beta} &= T^{-1}\sum_{t=1}^{T}\left\{\tfrac{1}{2}\left[z_{1t}^{*}+(\nu_0-2)z_{2t}^{*}\right]\left(\frac{\partial\sigma_t^2(\theta_0)/\partial\beta}{\sigma_t^2(\theta_0)}\right)^2+\tfrac{1}{2}z_{1t}^{*}\frac{\partial^2\sigma_t^2(\theta_0)/\partial\beta^2}{\sigma_t^2(\theta_0)}\right\} \qquad \text{(B.5)}\\
&= T^{-1}\sum_{t=1}^{T}\tfrac{1}{2}\left[z_{1t}^{*}+(\nu_0-2)z_{2t}^{*}\right]\left[\left(\frac{\partial\sigma_t^2(\theta_0)/\partial\beta}{\sigma_t^2(\theta_0)}\right)^2-u_{\beta t}^2\right]+T^{-1}\sum_{t=1}^{T}\tfrac{1}{2}z_{1t}^{*}u_{\beta\beta t}\\
&\quad +T^{-1}\sum_{t=1}^{T}\tfrac{1}{2}\left[z_{1t}^{*}+(\nu_0-2)z_{2t}^{*}\right]u_{\beta t}^2+T^{-1}\sum_{t=1}^{T}\tfrac{1}{2}z_{1t}^{*}\left(\frac{\partial^2\sigma_t^2(\theta_0)/\partial\beta^2}{\sigma_t^2(\theta_0)}-u_{\beta\beta t}\right),
\end{aligned}
$$

where $u_{\beta t}$ and $u_{\beta\beta t}$ are defined in Lemmas B.5 and B.6, respectively. By the strong law of large numbers for i.i.d. processes, Lemma B.5, and Toeplitz's lemma,

$$
T^{-1}\sum_{t=1}^{T}\tfrac{1}{2}\left[z_{1t}^{*}+(\nu_0-2)z_{2t}^{*}\right]\left[\left(\frac{\partial\sigma_t^2(\theta_0)/\partial\beta}{\sigma_t^2(\theta_0)}\right)^2-u_{\beta t}^2\right]\xrightarrow{P}0. \qquad \text{(B.6)}
$$

Likewise, using Lemma B.6 instead of Lemma B.5,

$$
T^{-1}\sum_{t=1}^{T}\tfrac{1}{2}z_{1t}^{*}\left(\frac{\partial^2\sigma_t^2(\theta_0)/\partial\beta^2}{\sigma_t^2(\theta_0)}-u_{\beta\beta t}\right)\xrightarrow{P}0. \qquad \text{(B.7)}
$$

By Lemmas B.5 and B.6 and the ergodic theorem,

$$
\begin{aligned}
&T^{-1}\sum_{t=1}^{T}\tfrac{1}{2}\left[z_{1t}^{*}+(\nu_0-2)z_{2t}^{*}\right]u_{\beta t}^2+T^{-1}\sum_{t=1}^{T}\tfrac{1}{2}z_{1t}^{*}u_{\beta\beta t}\\
&\xrightarrow{a.s.}\tfrac{1}{2}E\{\left[z_{1t}^{*}+(\nu_0-2)z_{2t}^{*}\right]u_{\beta t}^2+z_{1t}^{*}u_{\beta\beta t}\}\\
=\;&\tfrac{1}{2}E\left[z_{1t}^{*}+(\nu_0-2)z_{2t}^{*}\right]E[u_{\beta t}^2]+\tfrac{1}{2}E[z_{1t}^{*}]E[u_{\beta\beta t}]\\
=\;&\tfrac{1}{2}(\nu_0-2)E\left[z_{2t}^{*}\right]E[u_{\beta t}^2]=\tfrac{\nu_0}{2(\nu_0+3)}\tfrac{(1+\mu_1)\mu_2}{\beta_0^2(1-\mu_1)(1-\mu_2)}=\Omega_{\beta\beta}, \qquad \text{(B.8)}
\end{aligned}
$$

where we have used that (z_{1t}^{*},z_{2t}^{*}) is independent of $(u_{\beta t},u_{\beta\beta t})$. By combining (B.5)-(B.8), we obtain

$$
T^{-1}\sum_{t=1}^{T}i_{t,\beta\beta}\xrightarrow{P}\Omega_{\beta\beta}.
$$

Using similar arguments together with (C.5)-(C.9), we conclude that $\mathcal{I}_T\xrightarrow{P}\Omega$. □

Proof of Lemma B.3: From (C.13) it holds that,

$$
\begin{aligned}
\sup_{\theta\in\mathcal{N}(\theta_0)}\left|\frac{\partial^3 l_t(\theta)}{\partial\beta^3}\right| &\le \tfrac{3}{2}\left[1+(\nu_U+1)+\tfrac{(\nu-2)(\nu+1)}{(\nu_L-2)}\right]\sup_{\theta\in\mathcal{N}(\theta_0)}\left|\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right|\sup_{\theta\in\mathcal{N}(\theta_0)}\left|\frac{\partial^2\sigma_t^2(\theta)/\partial\beta^2}{\sigma_t^2(\theta)}\right|\\
&+\left[\tfrac{(\nu_U+1)(\nu_U-2)^2}{(\nu_L-2)^2}+(\nu_U+1)+\tfrac{(\nu_U-2)(\nu_U+1)}{(\nu_L-2)}+1\right]\sup_{\theta\in\mathcal{N}(\theta_0)}\left|\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\lambda)}\right|^3\\
&+\tfrac{1}{2}\left[1+(\nu_U+1)\right]\sup_{\theta\in\mathcal{N}(\theta_0)}\left|\frac{\partial^3\sigma_t^2(\theta)/\partial\beta^3}{\sigma_t^2(\lambda)}\right|
\end{aligned}
$$

$$\leq \tfrac{3}{2}\left[1+(\nu_U+1)+\tfrac{(\nu-2)(\nu+1)}{(\nu_L-2)}\right]\bar{u}_{\beta t}(\beta_L,\beta_U)\bar{u}_{\beta\beta t}(\beta_L,\beta_U)$$
$$+\left[\tfrac{(\nu_U+1)(\nu_U-2)^2}{(\nu_L-2)^2}+(\nu_U+1)+\tfrac{(\nu_U-2)(\nu_U+1)}{(\nu_L-2)}+1\right]\bar{u}^3_{\beta t}(\beta_L,\beta_U)$$
$$+\tfrac{1}{2}\left[1+(\nu_U+1)\right]\bar{u}_{\beta\beta\beta t}(\beta_L,\beta_U)=:\tilde{w}_t,$$

where we have used Lemma B.8, and where $\bar{u}_{\beta t}(\beta_L,\beta_U)$, $\bar{u}_{\beta\beta t}(\beta_L,\beta_U)$, and $\bar{u}_{\beta\beta\beta t}(\beta_L,\beta_U)$ are defined in Lemma B.8. Another application of Lemma B.8 yields that $\{\tilde{w}_t\}$ is ergodic and $\tilde{w}_t$ is integrable, so by the ergodic theorem,

$$\sup_{\theta\in\mathcal{N}(\theta_0)}\left|\frac{\partial^3 L_T(\theta)}{\partial\beta^3}\right|\leq\frac{1}{T}\sum_{t=1}^{T}\sup_{\theta\in\mathcal{N}(\theta_0)}\left|\frac{\partial^3 l_t(\theta)}{\partial\beta^3}\right|\leq\frac{1}{T}\sum_{t=1}^{T}\tilde{w}_t\xrightarrow{a.s.}E[\tilde{w}_t]=c<\infty.$$

Similar arguments applied to the rest of the third-order derivatives of the log-likelihood function, stated in (C.14)-(C.22), yield (B.3). □

Proof of Lemma B.4: We choose $\alpha_L,\beta_L>0$ and $\nu_L>2$, and define $\omega_L:=\min(\omega_0,\omega)$ and $\omega_U:=\max(\omega_0,\omega)$, and, likewise, $\gamma_L:=\min(\gamma_0,\gamma)$ and $\gamma_U:=\max(\gamma_0,\gamma)$. Using Taylor expansions, it suffices to show that

$$\sup_{\substack{\omega_L\leq\omega\leq\omega_U\\ \gamma_L\leq\gamma\leq\gamma_U}}\left\|\frac{1}{T^{1/2}}\sum_{t=1}^{T}\frac{\partial^2 l_t(\lambda_0,\omega,\gamma)}{\partial\lambda\partial\psi}\right\|\xrightarrow{P}0 \tag{B.9}$$

and

$$\sup_{\lambda\in\mathcal{N}(\lambda_0)}\sup_{\substack{\omega_L\leq\omega\leq\omega_U\\ \gamma_L\leq\gamma\leq\gamma_U}}\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^3 l_t(\theta)}{\partial\lambda\partial\lambda'\partial\psi}\right\|\xrightarrow{P}0, \tag{B.10}$$

for $\psi=\omega,\gamma$. From (C.10),

$$\begin{aligned}\sup_{\substack{\omega_L\leq\omega\leq\omega_U\\ \gamma_L\leq\gamma\leq\gamma_U}}\left|\frac{\partial^2 l_t(\lambda_0,\omega,\gamma)}{\partial\alpha\partial\omega}\right| &\leq \sup_{\substack{\omega_L\leq\omega\leq\omega_U\\ \gamma_L\leq\gamma\leq\gamma_U}}\frac{1}{2}\left|1-\frac{(\nu_0+1)x_t^2/\sigma_t^2(\lambda_0,\omega,\gamma)}{(\nu_0-2)+x_t^2/\sigma_t^2(\lambda_0,\omega,\gamma)}-\frac{(\nu_0+1)(\nu_0-2)x_t^2/\sigma_t^2(\lambda_0,\omega,\gamma)}{\left[(\nu_0-2)+x_t^2/\sigma_t^2(\lambda_0,\omega,\gamma)\right]^2}\right|\\ &\quad\times\sup_{\substack{\omega_L\leq\omega\leq\omega_U\\ \gamma_L\leq\gamma\leq\gamma_U}}\left|\left(\frac{\partial\sigma_t^2(\lambda_0,\omega,\gamma)/\partial\alpha}{\sigma_t^2(\lambda_0,\omega,\gamma)}\right)\left(\frac{\partial\sigma_t^2(\lambda_0,\omega,\gamma)/\partial\omega}{\sigma_t^2(\lambda_0,\omega,\gamma)}\right)\right|\\ &\leq \tfrac{1}{2}(1+2(\nu_0+1))\bar{u}_{\alpha t}(\beta_L,\beta_U)\kappa_2[2+(\beta_0-\beta_L)u_{1t}(\beta_0,\beta_L)]r_{1\omega t},\end{aligned}$$

where we have used Lemmas B.8 and B.10, and where $\bar{u}_{\alpha t}(\beta_L,\beta_U)$ and κ_2 are defined in Lemma B.8 and $r_{1\omega t}$ is defined in Lemma B.10. An application of Lemmas B.8 and B.10 together with Jensen and Rahbek (2004a, Lemma 11) gives that

$$\sup_{\substack{\omega_L\leq\omega\leq\omega_U\\ \gamma_L\leq\gamma\leq\gamma_U}}T^{-1/2}\sum_{t=1}^{T}\left|\partial^2 l_t(\lambda_0,\omega,\gamma)/\partial\alpha\partial\omega\right|\xrightarrow{P}0.$$

Likewise,

$$\sup_{\substack{\omega_L\le\omega\le\omega_U\\ \gamma_L\le\gamma\le\gamma_U}}\left|\frac{\partial^2 l_t(\lambda_0,\omega,\gamma)}{\partial\alpha\partial\gamma}\right| \le \sup_{\substack{\omega_L\le\omega\le\omega_U\\ \gamma_L\le\gamma\le\gamma_U}}\frac{1}{2}\left|1-\frac{(\nu_0+1)x_t^2/\sigma_t^2(\lambda_0,\omega,\gamma)}{(\nu_0-2)+x_t^2/\sigma_t^2(\lambda_0,\omega,\gamma)}-\frac{(\nu_0+1)(\nu_0-2)x_t^2/\sigma_t^2(\lambda_0,\omega,\gamma)}{\left[(\nu_0-2)+x_t^2/\sigma_t^2(\lambda_0,\omega,\gamma)\right]^2}\right|$$

$$\times\sup_{\substack{\omega_L\le\omega\le\omega_U\\ \gamma_L\le\gamma\le\gamma_U}}\left|\left(\frac{\partial\sigma_t^2(\lambda_0,\omega,\gamma)/\partial\alpha}{\sigma_t^2(\lambda_0,\omega,\gamma)}\right)\left(\frac{\partial\sigma_t^2(\lambda_0,\omega,\gamma)/\partial\gamma}{\sigma_t^2(\lambda_0,\omega,\gamma)}\right)\right|$$

$$\le \tfrac{1}{2}(1+2(\nu_0+1))\bar{u}_{\alpha t}(\beta_L,\beta_U)\kappa_2[2+(\beta_0-\beta_L)u_{1t}(\beta_0,\beta_L)]\frac{r_{\gamma t}}{\gamma_0},$$

where $r_{\gamma t}$ is defined in Lemma B.10. By arguments similar to the ones above,

$$\sup_{\substack{\omega_L\le\omega\le\omega_U\\ \gamma_L\le\gamma\le\gamma_U}} T^{-1/2}\sum_{t=1}^{T}\left|\partial^2 l_t(\lambda_0,\omega,\gamma)/\partial\alpha\partial\gamma\right| \xrightarrow{P} 0.$$

Likewise, similar arguments can be applied to (C.11) and (C.12) in order to conclude that (B.9) holds. Turning to the proof of (B.10), it holds by (C.23) that

$$\sup_{\lambda\in\mathcal{N}(\lambda_0)}\sup_{\substack{\omega_L\le\omega\le\omega_U\\ \gamma_L\le\gamma\le\gamma_U}}\left|\frac{\partial^3 l_t(\theta)}{\partial\alpha^2\partial\omega}\right|$$

$$\le \sup_{\lambda\in\mathcal{N}(\lambda_0)}\sup_{\substack{\omega_L\le\omega\le\omega_U\\ \gamma_L\le\gamma\le\gamma_U}}\left|\frac{(\nu+1)(\nu-2)^2x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}+\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}+\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}-1\right|$$

$$\times\sup_{\lambda\in\mathcal{N}(\lambda_0)}\sup_{\substack{\omega_L\le\omega\le\omega_U\\ \gamma_L\le\gamma\le\gamma_U}}\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)^2\left(\frac{\partial\sigma_t^2(\theta)/\partial\omega}{\sigma_t^2(\theta)}\right),$$

$$\le \left(\frac{(\nu_U+1)(\nu_U-2)^2}{(\nu_L-2)^2}+(\nu_U+1)+\frac{(\nu_U+1)(\nu_U-2)}{(\nu_L-2)}+1\right)$$
$$\times\bar{u}_{\alpha t}^2(\beta_L,\beta_U)\kappa_2[2+(\beta_0-\beta_L)u_{1t}(\beta_0,\beta_L)]r_{1\omega t},$$

and, likewise,

$$\sup_{\lambda\in\mathcal{N}(\lambda_0)}\sup_{\substack{\omega_L\le\omega\le\omega_U\\ \gamma_L\le\gamma\le\gamma_U}}\left|\frac{\partial^3 l_t(\theta)}{\partial\alpha^2\partial\gamma}\right| \le \left(\frac{(\nu_U+1)(\nu_U-2)^2}{(\nu_L-2)^2}+(\nu_U+1)+\frac{(\nu_U+1)(\nu_U-2)}{(\nu_L-2)}+1\right)$$

$$\times\bar{u}_{\alpha t}^2(\beta_L,\beta_U)\kappa_2[2+(\beta_0-\beta_L)u_{1t}(\beta_0,\beta_L)]\frac{r_{\gamma t}}{\gamma_0},$$

so another application of application of Lemmas B.8 and B.10 together with Jensen and Rahbek (2004a, Lemma 11) yields that

$$\sup_{\lambda\in\mathcal{N}(\lambda_0)}\sup_{\substack{\omega_L\le\omega\le\omega_U\\ \gamma_L\le\gamma\le\gamma_U}} T^{-1}\sum_{t=1}^{T}\left|\partial^3 l_t(\theta)/\partial\alpha^2\partial\psi\right| \xrightarrow{P} 0.$$

Similar arguments applied to (C.24)-(C.28) yield (B.10). □

B.3 Auxiliary Lemmas: t_ν-GARCH

Lemma B.5 *Define*

$$u_{\alpha t} := \sum_{j=1}^{\infty} \beta_0^{-1} z_{t-j}^2 \prod_{k=1}^{j} \tfrac{\beta_0}{\alpha_0 z_{t-k}^2 + \beta_0} \tag{B.11}$$

and

$$u_{\beta t} := \sum_{j=1}^{\infty} \beta_0^{-1} \prod_{k=1}^{j} \tfrac{\beta_0}{\alpha_0 z_{t-k}^2 + \beta_0}. \tag{B.12}$$

Then the sequences $\{u_{\alpha t}\}$ and $\{u_{\beta t}\}$ are strictly stationary and ergodic,

$$0 \leq \tfrac{\partial \sigma_t^2(\theta_0)/\partial \alpha}{\sigma_t^2(\theta_0)} \leq u_{\alpha t} \quad \text{and} \quad 0 \leq \tfrac{\partial \sigma_t^2(\theta_0)/\partial \beta}{\sigma_t^2(\theta_0)} \leq u_{\beta t},$$

and for any $p \geq 1$

$$E[u_{\alpha t}^p] < \infty \quad \text{and} \quad E[u_{\beta t}^p] < \infty.$$

In particular,

$$\begin{aligned} E[u_{\alpha t}] &= \alpha_0^{-1},\ E[u_{\alpha t}^2] = \alpha_0^{-2},\ E[u_{\beta t}] = \tfrac{\mu_1}{\beta_0(1-\mu_1)}, \\ E[u_{\beta t}^2] &= \tfrac{(1+\mu_1)\mu_2}{\beta_0^2(1-\mu_1)(1-\mu_2)},\ \text{and}\ E[u_{\alpha t} u_{\beta t}] = \tfrac{\mu_1}{\alpha_0 \beta_0 (1-\mu_1)}, \end{aligned}$$

where $\mu_i = E[(\beta_0/(\alpha_0 z_t^2 + \beta_0))^i]$, $i = 1, 2$.
Under the nonstationarity condition (7),

$$\tfrac{\partial \sigma_t^2(\theta_0)/\partial \alpha}{\sigma_t^2(\theta_0)} - u_{\alpha t} \xrightarrow{L_p} 0 \text{ as } t \to \infty, \quad \tfrac{\partial \sigma_t^2(\theta_0)/\partial \beta}{\sigma_t^2(\theta_0)} - u_{\beta t} \xrightarrow{L_p} 0 \text{ as } t \to \infty,$$

$$\tfrac{1}{T}\sum_{t=1}^{T}\left[\left(\tfrac{\partial \sigma_t^2(\theta_0)/\partial \alpha}{\sigma_t^2(\theta_0)}\right)^2 - (u_{\alpha t})^2\right] \xrightarrow{L_p} 0, \quad \tfrac{1}{T}\sum_{t=1}^{T}\left[\left(\tfrac{\partial \sigma_t^2(\theta_0)/\partial \beta}{\sigma_t^2(\theta_0)}\right)^2 - (u_{\beta t})^2\right] \xrightarrow{L_p} 0,$$

$$\left(\tfrac{\partial \sigma_t^2(\theta_0)/\partial \alpha}{\sigma_t^2(\theta_0)}\right)\left(\tfrac{\partial \sigma_t^2(\theta_0)/\partial \beta}{\sigma_t^2(\theta_0)}\right) - u_{\alpha t} u_{\beta t} \xrightarrow{L_p} 0 \text{ as } t \to \infty, \tag{B.13}$$

and

$$\tfrac{1}{T}\sum_{t=1}^{T}\left[\left(\tfrac{\partial \sigma_t^2(\theta_0)/\partial \alpha}{\sigma_t^2(\theta_0)}\right)\left(\tfrac{\partial \sigma_t^2(\theta_0)/\partial \beta}{\sigma_t^2(\theta_0)}\right) - u_{\alpha t} u_{\beta t}\right] \xrightarrow{L_p} 0 \tag{B.14}$$

for all $p \geq 1$.

Proof of Lemma B.5: The results hold by arguments similar to the ones given in the proofs of Lemmas 3 and 4 in Jensen and Rahbek (2004a)[1]. In particular, notice that $u_{\alpha t}$ and $u_{\beta t}$ are measurable functions of the strictly stationary and ergodic

[1]In Jensen and Rahbek (2004a, Proof of Lemma 4), the convergence in display (25), and in the display immediately before, holds in probability in the case where the nonstationarity condition holds with equality, see Theorem 2.1.a in Klüppelberg, Lindner, and Maller (2004). The conclusion of their Lemma 4 is however still valid, since it is only needed that the convergence in (25) holds in L_1. This holds as $\prod_{k=1}^{j}(\beta_0/(\alpha_0 z_{t-k}^2 + \beta_0) - \beta_0^j h_{t-j}/h_t)$ is uniformly integrable.

process $\{z_t^2\}$, and due to the fact that $u_{\alpha t}$ and $u_{\beta t}$ are integrable, they are almost surely finite. Hence, $\{u_{\alpha t}\}$ and $\{u_{\beta t}\}$ are strictly stationary and ergodic. For deriving the second-order moments of $(u_{\alpha t}, u_{\beta_t})$ it is used that $u_{\alpha t} = \alpha_0^{-1}$ a.s., see Jensen and Rahbek (2004a, p.1218). The convergence in (B.13) and (B.14) hold by observing that $[(\partial\sigma_t^2(\theta_0)/\partial\alpha)/\sigma_t^2(\theta_0)][(\partial\sigma_t^2(\theta_0)/\partial\beta)/\sigma_t^2(\theta_0)] - u_{\alpha t}u_{\beta t} = [(\partial\sigma_t^2(\theta_0)/\partial\alpha)/\sigma_t^2(\theta_0) - u_{\alpha t}][(\partial\sigma_t^2(\theta_0)/\partial\beta)/\sigma_t^2(\theta_0) - u_{\beta t}] + [(\partial\sigma_t^2(\theta_0)/\partial\alpha)/\sigma_t^2(\theta_0) - u_{\alpha t}]u_{\beta t} + [(\partial\sigma_t^2(\theta_0)/\partial\beta)/\sigma_t^2(\theta_0) - u_{\beta t}]u_{\alpha t}$. $\square$

LEMMA B.6 *Let*

$$u_{\alpha\beta t} := \sum_{j=1}^{\infty}(j-1)\beta_0^{-1}\tfrac{z_{t-j}^2}{\alpha_0 z_{t-j}^2+\beta_0}\prod_{k=1}^{j-1}\tfrac{\beta_0}{\alpha_0 z_{t-k}^2+\beta_0}, \quad \textit{and} \quad u_{\beta\beta t} := 2\sum_{j=1}^{\infty}(j-1)\beta_0^{-2}\prod_{k=1}^{j}\tfrac{\beta_0}{\alpha_0 z_{t-k}^2+\beta_0}.$$

Then the sequences $\{u_{\alpha\beta t}\}$ and $\{u_{\beta\beta t}\}$ are strictly stationary and ergodic,

$$0 \le \tfrac{\partial^2\sigma_t^2(\theta_0)/\partial\alpha\partial\beta}{\sigma_t^2(\theta_0)} \le u_{\alpha\beta t} \quad \textit{and} \quad 0 \le \tfrac{\partial^2\sigma_t^2(\theta_0)/\partial\beta^2}{\sigma_t^2(\theta_0)} \le u_{\beta\beta t},$$

and for any $p \ge 1$, $E[u_{\alpha\beta t}^p] < \infty$ and $E[u_{\beta\beta t}^p] < \infty$. Moreover, under the nonstationarity condition (7), as $t \to \infty$, for all $p \ge 1$,

$$\tfrac{\partial^2\sigma_t^2(\theta_0)/\partial\alpha/\partial\beta}{\sigma_t^2(\theta_0)} - u_{\alpha\beta t} \xrightarrow{L_p} 0 \quad \textit{and} \quad \tfrac{\partial^2\sigma_t^2(\theta_0)/\partial\beta^2}{\sigma_t^2(\theta_0)} - u_{\beta\beta t} \xrightarrow{L_p} 0.$$

PROOF OF LEMMA B.6: The proof follows by arguments similar to the ones given in the proof of Lemma B.5. $\square$

LEMMA B.7 *With $a, b > 0$, define*

$$u_{mt}(a,b) = m\sum_{j=1}^{\infty}a^{j-m}\prod_{n=1}^{m-1}(j-n)\prod_{k=1}^{j}\tfrac{1}{\alpha_0 z_{t-k}^2+b}, \quad m = 1,2,3,4,$$

with the convention that $\prod_{n=1}^{0} = 1$. For any $p \ge 1$ there exists β_L and β_U, $\beta_L < \beta_0 < \beta_U$, such that $\{u_{mt}(\beta_0,\beta_L)\}$ and $\{u_{mt}(\beta_U,\beta_0)\}$ are strictly stationary and ergodic with

$$E[u_{mt}^p(\beta_0,\beta_L)] < \infty \quad \textit{and} \quad E[u_{mt}^p(\beta_U,\beta_0)] < \infty.$$

PROOF OF LEMMA B.7: See Jensen and Rahbek (2004a, Lemma 3). $\square$

LEMMA B.8 *For $a, b > 0$, define*

$$\bar{u}_{\beta t}(a,b) \;=\; \kappa_1[u_{1t}(\beta_0,a) + u_{1t}(b,\beta_0) + \tfrac{1}{2}(b-\beta_0)u_{2t}(b,\beta_0)],$$

$$
\begin{aligned}
\bar{u}_{\beta\beta t}(a,b) &= \kappa_1[u_{2t}(\beta_0,a)+u_{2t}(b,\beta_0)+\tfrac{1}{3}(b-\beta_0)u_{3t}(b,\beta_0)],\\
\bar{u}_{\beta\beta\beta t}(a,b) &= \kappa_1[u_{3t}(\beta_0,a)+u_{3t}(b,\beta_0)+\tfrac{1}{4}(b-\beta_0)u_{4t}(b,\beta_0)],\\
\bar{u}_{\alpha t}(a,b) &= \kappa_2[2+(\beta_0-a)u_{1t}(\beta_0,a)]\sum_{j=1}^{\infty}\left(\tfrac{z_{t-j}^2}{\alpha_0 z_{t-j}^2+\beta_0}\right)\prod_{k=1}^{j-1}\tfrac{b}{\alpha_0 z_{t-k}^2+\beta_0},\\
\bar{u}_{\alpha\beta t}(a,b) &= \kappa_2[2+(\beta_0-a)u_{1t}(\beta_0,a)]\sum_{j=2}^{\infty}(j-1)a^{-1}\left(\tfrac{z_{t-j}^2}{\alpha_0 z_{t-j}^2+\beta_0}\right)\prod_{k=1}^{j-1}\tfrac{b}{\alpha_0 z_{t-k}^2+\beta_0},\\
\bar{u}_{\alpha\beta\beta t}(a,b) &= \kappa_2[2+(\beta_0-a)u_{1t}(\beta_0,a)]\sum_{j=3}^{\infty}(j-1)(j-2)a^{-2}\left(\tfrac{z_{t-j}^2}{\alpha_0 z_{t-j}^2+\beta_0}\right)\prod_{k=1}^{j-1}\tfrac{b}{\alpha_0 z_{t-k}^2+\beta_0},
\end{aligned}
$$

where $u_{mt}(a,b)$, $m=1,2,3,4$, is defined in Lemma B.7, and the constants κ_i, $i=1,2$, are given by

$$
\kappa_1 = \tfrac{\max(\alpha_U/\alpha_0,\gamma_U/\gamma_0,\omega_U/\omega_0)}{\min(\alpha_L/\alpha_0,\gamma_L/\gamma_0,\omega_L/\omega_0)} \quad \textit{and} \quad \kappa_2 = \min(\alpha_L/\alpha_0,\gamma_L/\gamma_0,\omega_L/\omega_0).
$$

For any $p \geq 1$, there exists a neighborhood, $\mathcal{N}(\theta_0)$, as in (B.2) such that

$$
\begin{aligned}
\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)} &\leq \bar{u}_{\beta t}(\beta_L,\beta_U), && \text{(B.15)}\\
\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial^2\sigma_t^2(\theta)/\partial\beta^2}{\sigma_t^2(\theta)} &\leq \bar{u}_{\beta\beta t}(\beta_L,\beta_U), && \text{(B.16)}\\
\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial^3\sigma_t^2(\theta)/\partial\beta^3}{\sigma_t^2(\theta)} &\leq \bar{u}_{\beta\beta\beta t}(\beta_L,\beta_U), && \text{(B.17)}\\
\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)} &\leq \bar{u}_{\alpha t}(\beta_L,\beta_U), && \text{(B.18)}\\
\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial^2\sigma_t^2(\theta)/\partial\alpha\partial\beta}{\sigma_t^2(\theta)} &\leq \bar{u}_{\alpha\beta t}(\beta_L,\beta_U), && \text{(B.19)}\\
\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial^3\sigma_t^2(\theta)/\partial\alpha\partial\beta^2}{\sigma_t^2(\theta)} &\leq \bar{u}_{\alpha\beta\beta t}(\beta_L,\beta_U). && \text{(B.20)}
\end{aligned}
$$

and such that the process $\{u_t\}$, where

$$
u_t := [\bar{u}_{\beta t}(\beta_L,\beta_U),\bar{u}_{\beta\beta t}(\beta_L,\beta_U),\bar{u}_{\beta\beta\beta t}(\beta_L,\beta_U),\bar{u}_{\alpha t}(\beta_L,\beta_U),\bar{u}_{\alpha\beta t}(\beta_L,\beta_U),\bar{u}_{\alpha\beta\beta t}(\beta_L,\beta_U)]',
$$

is ergodic with $E[\|u_t\|^p] < \infty$.

Proof of Lemma B.8: First, observe that (B.15)-(B.17) and $E[\bar{u}_{\beta t}^p(\beta_L,\beta_U)]$, $E[\bar{u}_{\beta\beta t}^p(\beta_L,\beta_U)]$, $E[\bar{u}_{\beta\beta\beta t}^p(\beta_L,\beta_U)] < \infty$ follow by Lemma B.7 together with Jensen and Rahbek (2004a, Lemmas 7 and 9). Next, notice that

$$
\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)} = \frac{\sum_{j=1}^t \beta^{j-1}x_{t-j}^2}{\sigma_t^2(\theta)} = \frac{\sum_{j=1}^t \beta^{j-1}z_{t-j}^2\tilde{\sigma}_{t-j}^2(\theta_0)}{\sigma_t^2(\theta)} = \sum_{j=1}^t \beta^{j-1}z_{t-j}^2\frac{\tilde{\sigma}_{t-j}^2(\theta_0)}{\sigma_t^2(\theta_0)}\frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)}
$$

$$
\begin{aligned}
&= \sum_{j=1}^{t} \beta^{j-1} z_{t-j}^2 \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \prod_{k=1}^{j} \frac{\tilde{\sigma}_{t-k}^2(\theta_0)}{\tilde{\sigma}_{t-k+1}^2(\theta_0)} = \sum_{j=1}^{t} \beta^{j-1} z_{t-j}^2 \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \prod_{k=1}^{j} \frac{\tilde{\sigma}_{t-k}^2(\theta_0)}{\omega_0+(\alpha_0 z_{t-k}^2+\beta_0)\tilde{\sigma}_{t-k}^2(\theta_0)} \\
&\leq \sum_{j=1}^{t} \beta^{j-1} z_{t-j}^2 \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \prod_{k=1}^{j} \frac{1}{\alpha_0 z_{t-k}^2+\beta_0} = \sum_{j=1}^{t} \left(\frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)}\right) \frac{z_{t-j}^2}{\alpha_0 z_{t-j}^2+\beta_0} \prod_{k=1}^{j-1} \frac{\beta}{\alpha_0 z_{t-k}^2+\beta_0}.
\end{aligned}
$$

From Jensen and Rahbek (2004a, Lemmas 7 and 9), $\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \leq \kappa_2[2 + (\beta_0 - \beta_L)u_{1t}(\beta_0, \beta_L)]$, and hence

$$
\begin{aligned}
&\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)} \leq \kappa_2[2 + (\beta_0 - \beta_L)u_{1t}(\beta_0, \beta_L)] \sum_{j=1}^{t} \left(\frac{z_{t-j}^2}{\alpha_0 z_{t-j}^2+\beta_0}\right) \prod_{k=1}^{j-1} \frac{\beta_U}{\alpha_0 z_{t-k}^2+\beta_0} \\
&\leq \kappa_2[2 + (\beta_0 - \beta_L)u_{1t}(\beta_0, \beta_L)] \sum_{j=1}^{\infty} \left(\frac{z_{t-j}^2}{\alpha_0 z_{t-j}^2+\beta_0}\right) \prod_{k=1}^{j-1} \frac{\beta_U}{\alpha_0 z_{t-k}^2+\beta_0} = \bar{u}_{\alpha t}(\beta_L, \beta_U).
\end{aligned}
$$

Similar to Jensen and Rahbek (2004a, Proof of Lemma 3), it holds that for any $p \geq 1$ there exits a $\beta_U > \beta_0$ such that

$$
E\left[\left(\frac{\beta_U}{\alpha_0 z_{t-k}^2+\beta_0}\right)^p\right] < 1 \text{ and } \frac{z_{t-j}^2}{\alpha_0 z_{t-j}^2+\beta_0} < \frac{1}{\alpha_0} \quad \text{a.s.} \tag{B.21}
$$

Combining (B.21) with Lemma B.7 yields that for any $p \geq 1$, there exists a neighborhood $\mathcal{N}(\theta_0)$ such that $E[\bar{u}_{\alpha t}^p(\beta_L, \beta_U)] < \infty$. Similar arguments yield (B.19) and (B.20) and $E[\bar{u}_{\alpha\beta t}^p(\beta_L, \beta_U)], E[\bar{u}_{\alpha\beta\beta t}^p(\beta_L, \beta_U)] < \infty$. Lastly, u_t is a measurable function of the ergodic process $\{z_t^2\}$, and due to the fact that u_t is almost surely finite (element-wise), we conclude that $\{u_t\}$ is ergodic. □

Lemma B.9 *Under the nonstationarity condition (7),*

$$
E[(\alpha_0 z_t^2 + \beta_0)^{-1}] < 1.
$$

Proof of Lemma B.9: Similar to the proof of Lemma A.8, and due to (7), it holds that $1 \leq \exp(E[\ln(\alpha_0 z_t^2 + \beta_0)]) < E[\alpha_0 z_t^2 + \beta_0]$ $(= \alpha_0 + \beta_0)$, where the strict inequality holds by Jensen's inequality and the fact that the exponential function is strictly convex and that $\ln(\alpha_0 z_t^2 + \beta_0)$ is non-degenerate. We then have that $1/E[\alpha_0 z_t^2 + \beta_0] < 1$, so by an application of Jensen's inequality for concave functions, $E[(\alpha_0 z_t^2 + \beta_0)^{-1}] < 1$. □

Lemma B.10 *Under the nonstationarity condition (7), there exists a neighborhood, $\mathcal{N}(\theta_0)$, as in (B.2) such that*

$$
\begin{aligned}
\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial\sigma_t^2(\theta)/\partial\omega}{\sigma_t^2(\theta)} &\leq \kappa_2[2 + (\beta_0 - \beta_L)u_{1t}(\beta_0, \beta_L)]r_{1\omega t}, \\
\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial^2\sigma_t^2(\theta)/\partial\beta\partial\omega}{\sigma_t^2(\theta)} &\leq \kappa_2[2 + (\beta_0 - \beta_L)u_{1t}(\beta_0, \beta_L)]r_{2\omega t}t,
\end{aligned}
$$

$$
\begin{aligned}
\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial^3\sigma_t^2(\theta)/\partial\beta^2\partial\omega}{\sigma_t^2(\theta)} &\leq \kappa_2[2+(\beta_0-\beta_L)u_{1t}(\beta_0,\beta_L)]r_{3\omega t}t^2, \\
\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial\sigma_t^2(\theta)/\partial\gamma}{\sigma_t^2(\theta)} &\leq \kappa_2[2+(\beta_0-\beta_L)u_{1t}(\beta_0,\beta_L)]\tfrac{r_{\gamma t}}{\gamma_0}, \\
\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial^2\sigma_t^2(\theta)/\partial\beta\partial\gamma}{\sigma_t^2(\theta)} &\leq \kappa_2[2+(\beta_0-\beta_L)u_{1t}(\beta_0,\beta_L)]\tfrac{r_{\gamma t}}{\beta_0\gamma_0}t, \\
\sup_{\theta\in\mathcal{N}(\theta_0)} \frac{\partial^2\sigma_t^2(\theta)/\partial\beta\partial\gamma}{\sigma_t^2(\theta)} &\leq \kappa_2[2+(\beta_0-\beta_L)u_{1t}(\beta_0,\beta_L)]\tfrac{r_{\gamma t}}{\beta_0\gamma_0}t(t-1),
\end{aligned}
$$

where κ_2 is defined in Lemma B.8 and $u_{1t}(\beta_0,\beta_L)$ is defined in Lemma B.7. Moreover, with $i=1,2,3$,

$$
r_{i\omega t} := \gamma_0^{-1}\left[(\beta_u-1)^{-i}\prod_{k=1}^{t}\tfrac{\beta_u}{\alpha_0 z_{t-k}^2+\beta_0}\mathbf{1}(\beta_0\geq 1)+(1-\beta_u)^{-i}\prod_{k=1}^{t}\tfrac{1}{\alpha_0 z_{t-k}^2+\beta_0}\mathbf{1}(\beta_0<1)\right]
$$

and

$$
r_{\gamma t} := \prod_{k=1}^{t}\tfrac{\beta_u}{\alpha_0 z_{t-k}^2+\beta_0},
$$

satisfying

$$
\begin{aligned}
E[r_{i\omega t}] &= \gamma_0^{-1}(\beta_u-1)^{-i}\rho^t,\ \textit{when } \beta_0\geq 1, \\
E[r_{i\omega t}] &= \gamma_0^{-1}(1-\beta_u)^{-i}\rho^t,\ \textit{when } \beta_0< 1 \\
E[r_{\gamma t}] &= \rho^t,
\end{aligned}
$$

with some $\rho<1$.

PROOF OF LEMMA B.10: The results follow from Jensen and Rahbek (2004a, Lemmas 7,9,12 and 13) and Lemma B.9. Notice that Lemma B.9 implies that Jensen and Rahbek (2004a, Proposition 1) holds for $p=1$, even when the nonstationarity condition (7) holds with equality. Thereby it is easily concluded that $r_{i\omega t}$ has exponentially decreasing mean for the case where $\beta_0<1$. □

C Likelihood Derivatives

With $l_t(\theta)$ the log-likelihood contribution defined in (2), its first-, second-, and third-order derivatives are given as follows.

C.1 First-order derivatives

$$
\frac{\partial l_t(\theta)}{\partial\alpha} = \frac{1}{2}\left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}-1\right]\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}, \tag{C.1}
$$

$$\frac{\partial l_t(\theta)}{\partial \beta} = \frac{1}{2}\left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - 1\right]\frac{\partial \sigma_t^2(\theta)/\partial \beta}{\sigma_t^2(\theta)}, \tag{C.2}$$

$$\frac{\partial l_t(\theta)}{\partial \nu} = \frac{\partial \ln \gamma(\nu)}{\partial \nu} - \frac{1}{2}\ln\left(1 + \frac{x_t^2/\sigma_t^2(\theta)}{\nu-2}\right) + \frac{1}{2(\nu-2)}\left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - 1\right]. \tag{C.3}$$

C.2 Second-order derivatives

$$\begin{aligned}\frac{\partial^2 l_t(\theta)}{\partial \beta^2} &= \frac{1}{2}\left[1 - \frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - \frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial \sigma_t^2(\theta)/\partial \beta}{\sigma_t^2(\theta)}\right)^2 \\ &+\frac{1}{2}\left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - 1\right]\frac{\partial^2 \sigma_t^2(\theta)/\partial \beta^2}{\sigma_t^2(\theta)},\end{aligned} \tag{C.4}$$

$$\frac{\partial^2 l_t(\theta)}{\partial \alpha^2} = \frac{1}{2}\left[1 - \frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - \frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial \sigma_t^2(\theta)/\partial \alpha}{\sigma_t^2(\theta)}\right)^2, \tag{C.5}$$

$$\begin{aligned}\frac{\partial^2 l_t(\theta)}{\partial \nu^2} &= \frac{\partial^2 \ln \gamma(\nu)}{\partial \nu \partial \nu} + \frac{1}{(\nu-2)}\frac{x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - \frac{1}{2(\nu-2)^2}\left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - 1\right] \\ &- \left(\frac{1}{2(\nu-2)}\right)\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2},\end{aligned} \tag{C.6}$$

$$\begin{aligned}\frac{\partial^2 l_t(\theta)}{\partial \alpha \partial \beta} &= \frac{1}{2}\left[1 - \frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - \frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial \sigma_t^2(\theta)/\partial \alpha}{\sigma_t^2(\theta)}\right)\left(\frac{\partial \sigma_t^2(\theta)/\partial \beta}{\sigma_t^2(\theta)}\right) \\ &+\frac{1}{2}\left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - 1\right]\frac{\partial^2 \sigma_t^2(\theta)/\partial \beta \partial \alpha}{\sigma_t^2(\theta)},\end{aligned} \tag{C.7}$$

$$\frac{\partial^2 l_t(\theta)}{\partial \alpha \partial \nu} = \frac{1}{2}\left[\frac{x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - \frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\frac{\partial \sigma_t^2(\theta)/\partial \alpha}{\sigma_t^2(\theta)}, \tag{C.8}$$

$$\frac{\partial^2 l_t(\theta)}{\partial \beta \partial \nu} = \frac{1}{2}\left[\frac{x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - \frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\frac{\partial \sigma_t^2(\theta)/\partial \beta}{\sigma_t^2(\theta)}. \tag{C.9}$$

With ψ denoting either ω or γ,

$$\frac{\partial^2 l_t(\theta)}{\partial \alpha \partial \psi} = \frac{1}{2}\left[1 - \frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - \frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial \sigma_t^2(\theta)/\partial \alpha}{\sigma_t^2(\theta)}\right)\left(\frac{\partial \sigma_t^2(\theta)/\partial \psi}{\sigma_t^2(\theta)}\right), \tag{C.10}$$

$$\begin{aligned}\frac{\partial^2 l_t(\theta)}{\partial \beta \partial \psi} &= \frac{1}{2}\left[1 - \frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - \frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial \sigma_t^2(\theta)/\partial \psi}{\sigma_t^2(\theta)}\right)\left(\frac{\partial \sigma_t^2(\theta)/\partial \beta}{\sigma_t^2(\theta)}\right) \\ &+\frac{1}{2}\left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - 1\right]\frac{\partial^2 \sigma_t^2(\theta)/\partial \beta \partial \psi}{\sigma_t^2(\theta)},\end{aligned} \tag{C.11}$$

$$\frac{\partial^2 l_t(\theta)}{\partial \nu \partial \psi} = \frac{1}{2}\left[\frac{x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - \frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial \sigma_t^2(\theta)/\partial \psi}{\sigma_t^2(\theta)}\right). \tag{C.12}$$

C.3 Third-order derivatives

$$\begin{aligned}\frac{\partial^3 l_t(\theta)}{\partial\beta^3} &= \frac{3}{2}\left[1-\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}-\frac{(\nu-2)(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta^2}{\sigma_t^2(\theta)}\right)\\ &+\left[\frac{(\nu+1)(\nu-2)^2x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}+\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}+\frac{(\nu-2)(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}-1\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)^3\\ &-\frac{1}{2}\left[1-\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}\right]\left(\frac{\partial^3\sigma_t^2(\theta)/\partial\beta^3}{\sigma_t^2(\theta)}\right),\end{aligned}\tag{C.13}$$

$$\begin{aligned}\frac{\partial^3 l_t(\theta)}{\partial\alpha^3} &= \left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}+\frac{(\nu-2)(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)^3\\ &+\left[\frac{(\nu+1)(\nu-2)^2x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}-1\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)^3,\end{aligned}\tag{C.14}$$

$$\begin{aligned}\frac{\partial^3 l_t(\theta)}{\partial\nu^3} &= \frac{\partial^3\ln\gamma(\nu)}{\partial\nu^3}-\left(\frac{3}{2(\nu-2)^2}\right)\frac{x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}-\left(\frac{3}{2(\nu-2)}\right)\frac{x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\\ &+\frac{1}{(\nu-2)^3}\left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}-1\right]+\left(\frac{1}{\nu-2}\right)^2\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\\ &+\left(\frac{1}{\nu-2}\right)\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3},\end{aligned}\tag{C.15}$$

$$\begin{aligned}\frac{\partial^3 l_t(\theta)}{\partial\alpha^2\partial\beta} &= \left[\frac{(\nu+1)(\nu-2)^2x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}+\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)^2\\ &-\left[1-\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)^2\\ &+\left[1-\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}-\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta\partial\alpha}{\sigma_t^2(\theta)}\right)\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right),\end{aligned}\tag{C.16}$$

$$\begin{aligned}\frac{\partial^3 l_t(\theta)}{\partial\alpha\partial\beta^2} &= \left[\frac{(\nu+1)(\nu-2)^2x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}+\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)^2\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)\\ &-\left[1-\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)^2\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)\\ &+\frac{1}{2}\left[1-\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}-\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta^2}{\sigma_t^2(\theta)}\right)\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)\\ &+\left[1-\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}-\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta\partial\alpha}{\sigma_t^2(\theta)}\right)\\ &+\frac{1}{2}\left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}-1\right]\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta^2\partial\alpha}{\sigma_t^2(\theta)}\right),\end{aligned}\tag{C.17}$$

$$\frac{\partial^3 l_t(\theta)}{\partial\beta^2\partial\nu} = \frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)^2-\frac{1}{2}\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta^2}{\sigma_t^2(\theta)}\right)$$

$$+\ \frac{1}{2}\left[\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta^2}{\sigma_t^2(\theta)}\right)-\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)^2\right]\frac{x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}$$

$$-\ \frac{1}{2}\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)^2\frac{(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}, \tag{C.18}$$

$$\frac{\partial^3 l_t(\theta)}{\partial\alpha\partial\beta\partial\nu}=\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)-\frac{1}{2}\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\alpha\partial\beta}{\sigma_t^2(\theta)}\right)$$

$$+\ \frac{1}{2}\left[\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\alpha\partial\beta}{\sigma_t^2(\theta)}\right)-\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)\right]\frac{x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}$$

$$-\ \frac{1}{2}\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)\frac{(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}, \tag{C.19}$$

$$\frac{\partial^3 l_t(\theta)}{\partial\nu^2\partial\beta}=\left\{\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}-\frac{x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right\}\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right), \tag{C.20}$$

$$\begin{aligned}\frac{\partial^3 l_t(\theta)}{\partial\alpha^2\partial\nu} &= \frac{1}{2}\left[-\frac{x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}-\frac{(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)^2\\ &\quad+\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)^2,\end{aligned} \tag{C.21}$$

$$\frac{\partial^3 l_t(\theta)}{\partial\nu^2\partial\alpha}=\left\{\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}-\frac{x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right\}\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right). \tag{C.22}$$

With ψ denoting either ω or γ,

$$\begin{aligned}\frac{\partial^3 l_t(\theta)}{\partial\alpha^2\partial\psi} &= \left[\frac{(\nu+1)(\nu-2)^2x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}+\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right)\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)^2\\ &\quad-\left[1-\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right)\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)^2,\end{aligned} \tag{C.23}$$

$$\begin{aligned}\frac{\partial^3 l_t(\theta)}{\partial\beta^2\partial\psi} &= \left[\frac{(\nu+1)(\nu-2)^2x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}+\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)^2\left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right)\\ &\quad-\left[1-\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)^2\left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right)\\ &\quad+\frac{1}{2}\left[1-\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}-\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta^2}{\sigma_t^2(\theta)}\right)\left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right)\\ &\quad+\left[1-\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}-\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta\partial\psi}{\sigma_t^2(\theta)}\right)\\ &\quad+\frac{1}{2}\left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}-1\right]\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta^2\partial\psi}{\sigma_t^2(\theta)}\right),\end{aligned} \tag{C.24}$$

$$\frac{\partial^3 l_t(\theta)}{\partial\alpha\partial\beta\partial\psi} = \left[\frac{(\nu+1)(\nu-2)^2x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3}+\frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right]\left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right)\left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right)\left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right)$$

$$
\begin{aligned}
&- \left[1 - \frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)}\right] \left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right) \left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right) \left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right) \\
&+ \frac{1}{2} \left[1 - \frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - \frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right] \\
&\times \left[\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\alpha\partial\beta}{\sigma_t^2(\theta)}\right) \left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right) + \left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right) \left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta\partial\psi}{\sigma_t^2(\theta)}\right)\right] \\
&+ \frac{1}{2} \left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} - 1\right] \left(\frac{\partial^3\sigma_t^2(\theta)/\partial\alpha\partial\beta\partial\psi}{\sigma_t^2(\theta)}\right),
\end{aligned} \tag{C.25}
$$

$$
\begin{aligned}
\frac{\partial^3 l_t(\theta)}{\partial\alpha\partial\nu\partial\psi} &= \frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3} \left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right) \left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right) \\
&\quad - \frac{1}{2} \left[\frac{x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} + \frac{(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right] \left(\frac{\partial\sigma_t^2(\theta)/\partial\alpha}{\sigma_t^2(\theta)}\right) \left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right),
\end{aligned} \tag{C.26}
$$

$$
\begin{aligned}
\frac{\partial^3 l_t(\theta)}{\partial\beta\partial\nu\partial\psi} &= \frac{(\nu+1)(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3} \left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right) \left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right) - \frac{1}{2} \frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2} \left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta\partial\psi}{\sigma_t^2(\theta)}\right) \\
&+ \frac{1}{2} \left[\left(\frac{\partial^2\sigma_t^2(\theta)/\partial\beta\partial\psi}{\sigma_t^2(\theta)}\right) - \left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right) \left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right)\right] \frac{x_t^2/\sigma_t^2(\theta)}{(\nu-2)+x_t^2/\sigma_t^2(\theta)} \\
&- \frac{1}{2} \left(\frac{\partial\sigma_t^2(\theta)/\partial\beta}{\sigma_t^2(\theta)}\right) \left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right) \frac{(\nu-2)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2},
\end{aligned} \tag{C.27}
$$

$$
\frac{\partial^3 l_t(\theta)}{\partial\nu^2\partial\psi} = \left[\frac{(\nu+1)x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^3} - \frac{x_t^2/\sigma_t^2(\theta)}{\left[(\nu-2)+x_t^2/\sigma_t^2(\theta)\right]^2}\right] \left(\frac{\partial\sigma_t^2(\theta)/\partial\psi}{\sigma_t^2(\theta)}\right). \tag{C.28}
$$