ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

http://www.econometricsociety.org/

Econometrica, Vol. 80, No. 2 (March, 2012), 821–861

STRICT STATIONARITY TESTING AND ESTIMATION OF EXPLOSIVE AND STATIONARY GENERALIZED AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTICITY MODELS

CHRISTIAN FRANCQ
University Lille 3 (EQUIPPE), 59653 Villeneuve d'Ascq cedex, France and CREST

JEAN-MICHEL ZAKOÏAN
CREST, 92245 Malakoff Cedex, France and University Lille 3 (EQUIPPE)

Econometrica, Vol. 80, No. 2 (March, 2012), 821–861

STRICT STATIONARITY TESTING AND ESTIMATION OF EXPLOSIVE AND STATIONARY GENERALIZED AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTICITY MODELS

BY CHRISTIAN FRANCQ AND JEAN-MICHEL ZAKOÏAN[1]

This paper studies the asymptotic properties of the quasi-maximum likelihood estimator of (generalized autoregressive conditional heteroscedasticity) GARCH(1, 1) models without strict stationarity constraints and considers applications to testing problems. The estimator is unrestricted in the sense that the value of the intercept, which cannot be consistently estimated in the explosive case, is not fixed. A specific behavior of the estimator of the GARCH coefficients is obtained at the boundary of the stationarity region, but, except for the intercept, this estimator remains consistent and asymptotically normal in every situation. The asymptotic variance is different in the stationary and nonstationary situations, but is consistently estimated with the same estimator in both cases. Tests of strict stationarity and nonstationarity are proposed. The tests developed for the classical GARCH(1, 1) model are able to detect nonstationarity in more general GARCH models. A numerical illustration based on stock indices and individual stock returns is proposed.

KEYWORDS: GARCH model, inconsistency of estimators, nonstationarity, quasi-maximum likelihood estimation.

1. INTRODUCTION

TESTING FOR STRICT STATIONARITY is an important issue in the context of financial time series. A standard assumption is that the prices are nonstationary while the returns (or log returns) are stationary. Numerous econometric tools, such as the unit root tests, have been introduced for testing the nonstationarity of prices. For the log returns, the most widely used models are arguably generalized autoregressive conditional heteroscedasticity (GARCH) models introduced by Engle (1982) and Bollerslev (1986). No econometric tools are available for testing strict stationarity in the GARCH framework. The main aim of this paper is to develop such tools. The problem is nonstandard because, contrary to stationarity in linear time series models, which solely depends on the lag polynomials, the strict stationarity condition for GARCH models has a nonexplicit form, involving the distribution of the underlying independent and identically distributed (i.i.d.) sequence.

The asymptotic properties of the quasi-maximum likelihood estimator (QMLE) for classical GARCH models have been extensively studied. Lumsdaine (1996) proved that the local QMLE is consistent and asymptotically

[1]We are most thankful to a co-editor and three referees for their constructive comments and suggestions. We would like to thank Youri A. Davydov, Siegfried Hörmann, Lajos Horváth, and several seminar and conference audiences for helpful discussions. We are also grateful to the Agence Nationale de la Recherche (ANR), which supported this work via the project ECONOM&RISK (ANR 2010 blanc 1804 03).

 DOI: 10.3982/ECTA9405

normal (CAN) in the GARCH(1, 1) case. These results were extended to the GARCH(p, q) model, under less stringent conditions, by Berkes, Horváth, and Kokoszka (2003) and Francq and Zakoïan (2004) (see also the references therein). For valid inference based on those results, strict stationarity must hold. Thus, from the point of view of the validity of the asymptotic results for the QMLE, strict stationarity testing in GARCH models is also an important issue. Surprisingly, this issue has not been addressed in the literature, to the best of our knowledge.

1.1. *Modes of Divergence in the Nonstationary Case*

The complexity of the statistical problem arises from the specificities of the probabilistic framework, even for the simplest GARCH model. To fix ideas, consider the GARCH(1, 1) model given by

$$\text{(1.1)} \qquad \begin{aligned} &\epsilon_t = \sqrt{h_t}\eta_t, \quad t = 1, 2, \dots, \\ &h_t = \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 h_{t-1}, \end{aligned}$$

with initial values ϵ_0 and $h_0 \geq 0$, where $\omega_0 > 0$, $\alpha_0, \beta_0 \geq 0$, and (η_t) is a sequence of independent and identically distributed (i.i.d.) variables such that $E\eta_1 = 0$, $E\eta_1^2 = 1$, and $P(\eta_1^2 = 1) < 1$. The top Lyapunov exponent associated to this model (see Bougerol and Picard (1992)) is given by

$$\gamma_0 = E \log a_0(\eta_1), \quad a_0(x) = \alpha_0 x^2 + \beta_0.$$

The necessary and sufficient condition for the existence of a strictly stationary solution to (1.1) is (by Nelson (1990))

$$\text{(1.2)} \qquad \gamma_0 < 0.$$

More precisely, if (1.2) holds, we have

$$\text{(1.3)} \qquad h_t - h_{t,\infty} \to 0 \quad \text{almost surely (a.s.)} \quad \text{as} \quad t \to \infty,$$

where

$$\text{(1.4)} \qquad h_{t,\infty} = \lim_{n\to\infty} \uparrow h_{t,n}, \quad h_{t,n} = \omega_0 \left(1 + \sum_{k=1}^{n-1} a_0(\eta_{t-1}) \cdots a_0(\eta_{t-k}) \right).$$

In particular, the integrated GARCH model, obtained when $\alpha_0 + \beta_0 = 1$, satisfies the condition (1.2).[2] Let us now turn to the nonstationary case, for which it is necessary to consider $\gamma_0 > 0$ and $\gamma_0 = 0$ separately. Under the assumption

$$\text{(1.5)} \qquad \gamma_0 > 0,$$

[2]Berkes, Horváth, and Kokoszka (2005) studied the asymptotic behavior of the GARCH(1, 1) process when $\alpha_0 + \beta_0$ approaches 1 as the sample size increases.

$h_t \to \infty$ almost surely as $t \to \infty$, as shown by Nelson (1990). In this case, the increasing sequence $h_{t,n}$ goes to infinity almost surely as $n \to \infty$, by the Cauchy root test. The case $\gamma_0 = 0$ is much more intricate. By the Chung–Fuchs theorem, it can be seen that $h_{t,n}$ goes to infinity almost surely as $n \to \infty$. However, the a.s. convergence of h_t to infinity may not hold when $\gamma_0 = 0$. Actually, Klüppelberg, Lindner, and Maller (2004) (see also Goldie and Maller (2000)) showed that

$$\text{(1.6)} \qquad \text{when} \quad \gamma_0 = 0, \quad h_t \to \infty \quad \text{in probability}$$

instead of almost surely in the case $\gamma_0 > 0$.[3] The astonishing difficulties encountered in the case $\gamma_0 = 0$ are related to the fact that the sequence $h_t = h_{t,t} + a_0(\eta_{t-1}) \cdots a_0(\eta_0) h_0$ does not increase with t.

1.2. *The Econometric Problem*

We consider the QMLE, which is the commonly used estimator for autoregressive conditional heteroscedasticity (ARCH) models. Denote by $\theta = (\omega, \alpha, \beta)'$ the GARCH(1, 1) parameter and define the QMLE as any measurable solution of

$$\text{(1.7)} \qquad \hat{\theta}_n = (\hat{\omega}_n, \hat{\alpha}_n, \hat{\beta}_n)' = \arg\min_{\theta \in \Theta} \frac{1}{n} \sum_{t=1}^{n} \ell_t(\theta), \quad \ell_t(\theta) = \frac{\epsilon_t^2}{\sigma_t^2(\theta)} + \log \sigma_t^2(\theta),$$

where Θ is a compact subset of $(0, \infty)^3$ that contains the true value $\theta_0 = (\omega_0, \alpha_0, \beta_0)'$ and where $\sigma_t^2(\theta) = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2(\theta)$ for $t = 1, \ldots, n$ (with initial values for ϵ_0^2 and $\sigma_0^2(\theta)$). The rescaled residuals are defined by $\hat{\eta}_t = \eta_t(\hat{\theta}_n)$, where $\eta_t(\theta) = \epsilon_t / \sigma_t(\theta)$ for $t = 1, \ldots, n$.

To construct a test of the strict stationarity assumption, we establish the asymptotic distribution of the statistic

$$\text{(1.8)} \qquad \hat{\gamma}_n = \frac{1}{n} \sum_{t=1}^{n} \log(\hat{\alpha}_n \hat{\eta}_t^2 + \hat{\beta}_n).$$

To study the asymptotic properties of the test, it is necessary to analyze the asymptotic behavior of the QMLE when $\gamma_0 \geq 0$. Jensen and Rahbek (2004a, 2004b) were the first researchers to establish an asymptotic theory for estimators of nonstationary GARCH.[4] However, they only considered a constrained QMLE of (α_0, β_0) (in the sense that the value of the intercept is fixed) that is consistent in the nonstationary case, but is inconsistent in the stationary case.

[3]Klüppelberg, Lindner, and Maller (2004) noted that the arguments given by Nelson (1990) for the a.s. convergence fail when $\gamma_0 = 0$.

[4]See Linton, Pan, and Wang (2010) for extensions in the case of non-i.i.d. errors.

Instead, we use the standard (unconstrained) QMLE. We complete the well known results in the case $\gamma_0 < 0$ by establishing the consistency and asymptotic normality of the QMLE of (α_0, β_0), the only components that matter for our testing problem, in the cases $\gamma_0 > 0$ and $\gamma_0 = 0$. When $\gamma_0 > 0$, the estimator $(\hat{\alpha}_n, \hat{\beta}_n)$ is shown to be strongly consistent and it turns out that its asymptotic distribution is simpler than in the strict stationarity case, and is given by

$$\sqrt{n}(\hat{\alpha}_n - \alpha_0, \hat{\beta}_n - \beta_0)' \xrightarrow{d} \mathcal{N}\{0, (\kappa_\eta - 1)\mathcal{I}^{-1}\} \quad \text{as} \quad n \to \infty,$$

where $\xrightarrow{d}$ stands for the convergence in distribution, $\kappa_\eta = E\eta_1^4$, and $\mathcal{I}$ is a matrix which has an explicit form and does not depend on ω_0. When $\gamma_0 = 0$, the QMLE of (α_0, β_0) is shown to be weakly consistent with the same asymptotic normal distribution as in the case $\gamma_0 > 0$. The asymptotic variances of $(\hat{\alpha}_n, \hat{\beta}_n)$ when $\gamma_0 \geq 0$ and when $\gamma_0 < 0$ do not coincide, but we propose an estimator which is consistent in both situations. This is in accordance with similar results for autoregressive models with random coefficients derived by Aue and Horváth (2011).

Even if the QMLE of (α_0, β_0) is consistent in every situation, we show that the QMLE of ω_0 is only consistent in the stationary case. For this reason, it is important to test the sign of γ_0.

The rest of the paper is organized as follows. Section 2 is devoted to the asymptotic properties of the QMLE. In Section 3, we first consider the problem of testing the value of (α_0, β_0) without any stationarity restriction. Then we consider strict stationarity testing. The asymptotic distributions of two tests are studied when the null assumption is either the stationarity or the nonstationarity. We also consider testing stationarity in more general GARCH-type models. Numerical illustrations are provided in Section 5. In particular, the stationarity of 11 major stock returns is analyzed. Section 6 concludes. Proofs and complementary results are collected in the Appendix. Replication files may be found in the online supplement (Francq and Zakoïan (2012)).

2. ASYMPTOTIC PROPERTIES OF THE QMLE

In this paper, we consider the standard QMLE, which is the commonly used estimator for GARCH models.

2.1. *Consistency and Asymptotic Normality of* $\hat{\alpha}_n$ *and* $\hat{\beta}_n$

The following result completes those already established in the stationary case, which we recall for convenience. The asymptotic distribution in the case $\gamma_0 = 0$ is treated separately. To handle initial values, we introduce the following notation. For any asymptotically stationary process $(X_t)_{t\geq 0}$, let $E_\infty(X_t) = \lim_{t\to\infty} E(X_t)$ provided this limit exists.[5]

[5]For instance, for the process (ϵ_t), we have $E_\infty(\epsilon_t^2) = \omega_0/(1 - \alpha_0 - \beta_0)$ when $\alpha_0 + \beta_0 < 1$.

THEOREM 2.1: *For the* GARCH(1, 1) *model* (1.1), *the QMLE defined in* (1.7) *satisfies the following properties*.

(i) *When* $\gamma_0 < 0$ *for* Θ *such that* $\forall\theta \in \Theta$, $\beta < 1$, *then*

$$\hat{\alpha}_n \to \alpha_0, \quad \hat{\beta}_n \to \beta_0 \quad \text{and} \quad \hat{\omega}_n \to \omega_0 \quad \text{a.s.} \quad \text{as} \quad n \to \infty.$$

(ii) *When* $\gamma_0 > 0$, *if* $P(\eta_1 = 0) = 0$, *then*

$$\hat{\alpha}_n \to \alpha_0 \quad \text{and} \quad \hat{\beta}_n \to \beta_0 \quad \text{a.s.} \quad \text{as} \quad n \to \infty.$$

(iii) *When* $\gamma_0 = 0$, *if* $P(\eta_1 = 0) = 0$, *for* Θ *such that* $\forall\theta \in \Theta$, $\beta < \|1/a_0(\eta_1)\|_p^{-1}$ *for some* $p > 1$, *then*

$$\hat{\alpha}_n \to \alpha_0 \quad \text{and} \quad \hat{\beta}_n \to \beta_0 \quad \text{in probability} \quad \text{as} \quad n \to \infty.$$

(iv) *When* $\gamma_0 < 0$, $\kappa_\eta = E\eta_1^4 \in (1, \infty)$, θ_0 *belongs to the interior* $\overset{\circ}{\Theta}$ *of* Θ, *and for* Θ *such that* $\forall\theta \in \Theta$, $\beta < 1$, *then*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathcal{N}\{0, (\kappa_\eta - 1)\mathcal{J}^{-1}\} \quad \text{as} \quad n \to \infty \tag{2.1}$$

and

$$\mathcal{J} = E_\infty\left(\frac{1}{h_t^2}\frac{\partial\sigma_t^2}{\partial\theta}\frac{\partial\sigma_t^2}{\partial\theta'}(\theta_0)\right). \tag{2.2}$$

(v) *When* $\gamma_0 > 0$, $\kappa_\eta \in (1, \infty)$, $E|\log\eta_1^2| < \infty$, *and* $\theta_0 \in \overset{\circ}{\Theta}$, *then*

$$\sqrt{n}(\hat{\alpha}_n - \alpha_0, \hat{\beta}_n - \beta_0)' \xrightarrow{d} \mathcal{N}\{0, (\kappa_\eta - 1)\mathcal{I}^{-1}\} \quad \text{as} \quad n \to \infty, \tag{2.3}$$

where

$$\mathcal{I} = \begin{pmatrix} \dfrac{1}{\alpha_0^2} & \dfrac{\nu_1}{\alpha_0\beta_0(1-\nu_1)} \\ \dfrac{\nu_1}{\alpha_0\beta_0(1-\nu_1)} & \dfrac{(1+\nu_1)\nu_2}{\beta_0^2(1-\nu_1)(1-\nu_2)} \end{pmatrix}$$
$$\text{with} \quad \nu_i = E\left(\frac{\beta_0}{\alpha_0\eta_1^2 + \beta_0}\right)^i.$$

To obtain the asymptotic distribution of $(\hat{\alpha}_n, \hat{\beta}_n)$ in the case $\gamma_0 = 0$, we need an additional assumption on the distribution of η_t^2. Let $Z_t = \alpha_0\eta_t^2 + \beta_0$. Note that $\gamma_0 = E\log Z_t = 0$ entails $EZ_t \geq 1$ by Jensen's inequality and, thus, in view of the independence, $E(1 + Z_{t-1} + Z_{t-1}Z_{t-2} + \cdots + Z_{t-1}\cdots Z_1) \geq t$. We introduce the following assumption.

ASSUMPTION A: *When t tends to infinity*,

$$E\left(\frac{1}{1+Z_1+Z_1Z_2+\cdots+Z_1\cdots Z_{t-1}}\right)=o\left(\frac{1}{\sqrt{t}}\right).$$

Note that Assumption A is obviously satisfied when $\eta_t=\pm 1$ with equal probabilities and for $\alpha_0+\beta_0=1$, since the expectation is then equal to $1/t$. It can be shown that the expectation involved in Assumption A is of order $(\log t)/\sqrt{t}$ for any distribution such that $E|\log Z_1|^2<\infty$ (details are available from the authors). This assumption is used to prove that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{1}{h_t}\to 0 \quad \text{as} \quad n\to\infty \tag{2.4}$$

in L^1 when $\gamma_0=0$.[6]

THEOREM 2.2: *Suppose that the assumptions of Theorem 2.1*(iii) *hold*, *in particular*, $\gamma_0=0$. *Then if* $\theta_0\in\overset{\circ}{\Theta}$, $\kappa_\eta\in(1,\infty)$, $E|\log\eta_1^2|<\infty$, *and Assumption A is satisfied*, *the QMLE* $(\hat{\alpha}_n,\hat{\beta}_n)$ *is asymptotically normal and its asymptotic distribution is given by* (2.3).

2.2. *Estimating the Asymptotic Variance of $(\hat{\alpha}_n,\hat{\beta}_n)$ Without Assuming Stationarity*

In view of (2.1) and (2.2), when $\gamma_0<0$, the asymptotic distribution of the QMLE $(\hat{\alpha}_n,\hat{\beta}_n)$ of (α_0,β_0) is given by

$$\sqrt{n}(\hat{\alpha}_n-\alpha_0,\hat{\beta}_n-\beta_0)'\xrightarrow{d}\mathcal{N}\{0,(\kappa_\eta-1)\mathcal{I}_*^{-1}\} \quad \text{as} \quad n\to\infty, \tag{2.5}$$

with

$$\mathcal{I}_*=\mathcal{J}_{\alpha\beta,\alpha\beta}-\mathcal{J}_{\alpha\beta,\omega}\mathcal{J}_{\omega,\omega}^{-1}\mathcal{J}_{\omega,\alpha\beta} \tag{2.6}$$

and

$$\mathcal{J}_{\omega,\omega}=E_\infty\left(\frac{1}{h_t^2}\frac{\partial\sigma_t^2}{\partial\omega}\frac{\partial\sigma_t^2}{\partial\omega}(\theta_0)\right),$$

$$\mathcal{J}_{\alpha\beta,\alpha\beta}=E_\infty\left(\frac{1}{h_t^2}\frac{\partial\sigma_t^2}{\partial(\alpha,\beta)'}\frac{\partial\sigma_t^2}{\partial(\alpha,\beta)}(\theta_0)\right),$$

[6]In the case $\gamma_0>0$, (2.4) holds a.s. from Proposition A.1(i) and Assumption A is not needed.

and

$$\mathcal{J}_{\omega,\alpha\beta} = \mathcal{J}'_{\alpha\beta,\omega} = E_\infty\left(\frac{1}{h_t^2}\frac{\partial\sigma_t^2}{\partial\omega}\frac{\partial\sigma_t^2}{\partial(\alpha,\beta)}(\theta_0)\right).$$

Letting

$$\widehat{\mathcal{J}}_{\alpha\beta,\alpha\beta} = \frac{1}{n}\sum_{t=1}^{n}\frac{1}{\sigma_t^4(\hat{\theta}_n)}\frac{\partial\sigma_t^2}{\partial(\alpha,\beta)'}\frac{\partial\sigma_t^2}{\partial(\alpha,\beta)}(\hat{\theta}_n)$$

and defining $\widehat{\mathcal{J}}_{\alpha\beta,\omega}$, $\widehat{\mathcal{J}}_{\omega,\omega}$, and $\widehat{\mathcal{J}}_{\omega,\alpha\beta}$ accordingly, it can be shown that

$$\widehat{\mathcal{I}}_* = \widehat{\mathcal{J}}_{\alpha\beta,\alpha\beta} - \widehat{\mathcal{J}}_{\alpha\beta,\omega}\widehat{\mathcal{J}}_{\omega,\omega}^{-1}\widehat{\mathcal{J}}_{\omega,\alpha\beta}$$

is a strongly consistent estimator of $\mathcal{I}_*$ in the stationary case $\gamma_0 < 0$. The following result shows that this estimator is also a consistent estimator of the asymptotic variance of $(\hat{\alpha}_n, \hat{\beta}_n)$ in the nonstationary case $\gamma_0 \geq 0$.

THEOREM 2.3: *Let the assumptions* (i)–(iii) *of Theorem* 2.1 *hold, assume* $\kappa_\eta \in (1,\infty)$, *and let* $\hat{\kappa}_\eta = n^{-1}\sum_{t=1}^{n}\hat{\eta}_t^4$, *where* $\hat{\eta}_t = \epsilon_t/\sigma_t(\hat{\theta}_n)$.

(i) *When* $\gamma_0 < 0$, *we have* $\hat{\kappa}_\eta \to \kappa_\eta$ *and* $\widehat{\mathcal{I}}_* \to \mathcal{I}_*$ *a.s. as* $n \to \infty$.

(ii) *When* $\gamma_0 > 0$, *we have* $\hat{\kappa}_\eta \to \kappa_\eta$ *and* $\widehat{\mathcal{I}}_* \to \mathcal{I}$ *a.s.*

(iii) *When* $\gamma_0 = 0$, *we have* $\hat{\kappa}_\eta \to \kappa_\eta$ *and if* A *is satisfied*, $\widehat{\mathcal{I}}_* \to \mathcal{I}$ *in probability*.

In any case, $(\hat{\kappa}_\eta - 1)\widehat{\mathcal{I}}_*^{-1}$ *is a consistent estimator of the asymptotic variance of the QMLE of* (α_0, β_0).

2.3. *Inconsistency of* $\hat{\omega}_n$ *When* $\gamma_0 > 0$

The previous results do not give any insight on the asymptotic behavior of the QMLE of ω_0. From the proof of Theorem 2.1, it can be shown that the log-likelihood is completely flat in the direction where (α_0, β_0) is fixed and ω_0 varies. More precisely, we have

$$\Lambda_n\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\ell_t(\theta_0) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\begin{pmatrix}0 & 0\\ 0 & \mathcal{I}\end{pmatrix}\right\}, \quad \Lambda_n = \begin{pmatrix}\lambda_n & 0\\ 0 & n^{-1/2}I_2\end{pmatrix}$$

for any sequence λ_n tending to zero as $n \to \infty$.

Thus, in general, as noted by Jensen and Rahbek (2004b), there is no consistent estimator of ω_0. Indeed, we have the following result.

PROPOSITION 2.1: *Consider the* GARCH(1,1) *model* (1.1) *with* $\eta_t \sim \mathcal{N}(0,1)$. *Assume that* Θ *contains two arbitrarily close points* $\theta_1 = (\omega_1, \alpha_1, \beta_1)$ *and* $\theta_1^* = (\omega_1^*, \alpha_1, \beta_1)$ *such that* $E\log(\alpha_1\eta_t^2 + \beta_1) > 0$ *and* $\omega_1 \neq \omega_1^*$. *There exists no consistent estimator of* $\theta_0 \in \Theta$.

The inconsistency of $\hat{\omega}_n$ is illustrated via simulations in Francq and Zakoïan (2010, p. 150).

2.4. *A Constrained QMLE of* (α_0, β_0)

Whereas the asymptotic behavior of the QMLE $(\hat{\alpha}_n, \hat{\beta}_n)$ is independent of ω_0 when $\gamma_0 > 0$ and the QMLE of ω_0 is generally inconsistent in view of Proposition 2.1, it seems natural to avoid estimating ω_0. To this aim, a *constrained* QMLE of (α_0, β_0), in which the first component of θ is fixed to an arbitrary value ω, can be introduced. The estimator

$$(2.7)\qquad (\hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega)) = \arg \min_{(\alpha,\beta)\in\Theta_{\alpha,\beta}} \frac{1}{n}\sum_{t=1}^{n} \ell_t(\omega, \alpha, \beta)$$

was studied by Jensen and Rahbek (2004b). They proved that when $\gamma_0 > 0$, (2.3) continues to hold when the *global* QMLE $(\hat{\alpha}_n, \hat{\beta}_n)$ is replaced by the *local* and *constrained* QMLE $(\hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega))$. In Appendix A.3, we prove that, under the assumptions of Theorem 2.2, in particular, $\gamma_0 = 0$,

$$(2.8)\qquad \sqrt{n}(\hat{\alpha}_n^c(\omega) - \alpha_0, \hat{\beta}_n^c(\omega) - \beta_0) \xrightarrow{d} \mathcal{N}\{0, (\kappa_\eta - 1)\mathcal{I}^{-1}\} \quad \text{as} \quad n \to \infty.$$

However, the next result shows that the restricted QMLE of (α_0, β_0) is generally inconsistent in the stationary case.

PROPOSITION 2.2: *Let* (ϵ_t) *be a stationary solution of the* GARCH(1, 1) *model with parameters* ω_0, α_0, *and* β_0, *such that* $E\epsilon_t^4 < \infty$. *If* $\omega \neq \omega_0$, *then*

$$(2.9)\qquad (\hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega)) \quad \textit{does not converge in probability to} \quad (\alpha_0, \beta_0).$$

On the contrary, Theorem 2.1 and Theorem 2.2 show that

(2.10) the QMLE of (α_0, β_0) is always CAN

(under Assumption A when $\gamma_0 = 0$).

3. TESTING

The consequence of Theorem 2.3, from a practical point of view, is extremely important. It means that we can get confidence intervals or tests for (α_0, β_0) without assuming stationarity/nonstationarity.

Before considering strict stationarity testing, we start with tests on the GARCH parameters.

3.1. *Testing the GARCH Coefficients*

First consider a testing problem of the form

$$H_0: a\alpha_0 + b\beta_0 \leq c \quad \text{against} \quad H_1: a\alpha_0 + b\beta_0 > c, \tag{3.1}$$

where a, b, and c are given numbers. A case of particular interest is $a = b = c = 1$, because $E\epsilon_t^2 < \infty$ if and only if $\alpha_0 + \beta_0 < 1$. Note, however, that we do not impose any constraint on a, b, and c, so some values of θ_0 satisfying H_0 may correspond to nonstationary GARCH models. A direct consequence of Theorems 2.1–2.3 is the following result, in which Φ denotes the $\mathcal{N}(0, 1)$ cumulative distribution function. Let $\underline{\alpha} \in (0, 1)$.

COROLLARY 3.1: *Assume that $\theta_0 \in \overset{\circ}{\Theta}$ and the assumptions of Theorem 2.3 hold. For the testing problem (3.1), the test defined by the critical region*

$$\mathrm{C}^{\alpha^*} = \left\{ T_n^{\alpha^*} := \frac{\sqrt{n}(a\hat{\alpha}_n + b\hat{\beta}_n - c)}{\sqrt{(\hat{\kappa}_\eta - 1)(a, b)\widehat{\mathcal{I}}_*^{-1}(a, b)'}} > \Phi^{-1}(1 - \underline{\alpha}) \right\} \tag{3.2}$$

has the asymptotic significance level $\underline{\alpha}$ and is consistent.

The following result, showing that no consistent test exists for ω_0, is used to prove the inconsistency of any estimator of this parameter.

PROPOSITION 3.1: *Consider the* GARCH$(1, 1)$ *model* (1.1) *with $\eta_t \sim \mathcal{N}(0, 1)$. Let $\theta_1 = (\omega_1, \alpha_1, \beta_1)$ and $\theta_1^* = (\omega_1^*, \alpha_1, \beta_1)$ be two points of Θ such that $E\log(\alpha_1\eta_t^2 + \beta_1) > 0$ and $\omega_1 \neq \omega_1^*$. When ω_1 and ω_1^* are sufficiently close, there exists no consistent test for $H_0: \theta_0 = \theta_1$ against $H_1: \theta_0 = \theta_1^*$ at the asymptotic level $\underline{\alpha} \in (0, 1/2)$.*

Propositions 2.1 and 3.1 show that no asymptotically valid inference on ω_0 can be done in the nonstationary case. It is thus of interest to test whether a given series is stationary or not.

3.2. *Strict Stationarity Testing*

Consider the strict stationarity testing problems

$$H_0: \gamma_0 < 0 \quad \text{against} \quad H_1: \gamma_0 \geq 0 \tag{3.3}$$

and

$$H_0: \gamma_0 \geq 0 \quad \text{against} \quad H_1: \gamma_0 < 0, \tag{3.4}$$

where $\gamma_0 = E\log(\alpha_0\eta_1^2 + \beta_0)$. These hypotheses are not of the form (3.1) because γ_0 not only depends on α_0 and β_0, but also on the unknown distribution of η_1. The following result gives the asymptotic distribution of the empirical estimator of γ_0 defined by (1.8) under either the stationarity or the nonstationarity conditions.

THEOREM 3.1: *Assume the following conditions are satisfied*: *Theorem* 2.1(iv) *in the case* $\gamma_0 < 0$, *Theorem* 2.1(v) *in the case* $\gamma_0 > 0$, *and Theorem* 2.2 (*and*, *in particular*, *Assumption* A) *when* $\gamma_0 = 0$. *Moreover*, *assume that* $E|a_0(\eta_1)|^2 < \infty$. *Let* $u_t = \log a_0(\eta_t) - \gamma_0$ *and* $\sigma_u^2 = Eu_t^2$. *Then*

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \xrightarrow{d} \mathcal{N}(0, \sigma_\gamma^2) \quad \textit{as} \quad n \to \infty, \tag{3.5}$$

where

$$\sigma_\gamma^2 = \begin{cases} \sigma_u^2 + (\kappa_\eta - 1)\{a'\mathcal{J}^{-1}a - (1-\nu_1)^2\} & \textit{when } \gamma_0 < 0, \\ \sigma_u^2 & \textit{when } \gamma_0 \geq 0, \end{cases}$$

with $a = (0, (1-\nu_1)/\alpha_0, \nu_1/\beta_0)'$ *and* $\nu_1 = E\{\beta_0/a_0(\eta_1)\}$.

It can be seen from the proof that the term in brackets in the first expression of σ_γ^2 is positive, showing that the asymptotic variance of $\hat{\gamma}_n$ is larger in the stationary case than in the nonstationary case. The next result provides an estimator of σ_γ^2 which is consistent in every situation (explosive and stationary). It allows to construct a confidence interval for the top Lyapunov exponent. Let

$$\hat{\sigma}_u^2 = \frac{1}{n}\sum_{t=1}^{n}\{\log(\hat{\alpha}_n\hat{\eta}_t^2 + \hat{\beta}_n)\}^2 - \left\{\frac{1}{n}\sum_{t=1}^{n}\log(\hat{\alpha}_n\hat{\eta}_t^2 + \hat{\beta}_n)\right\}^2,$$

$$\hat{\sigma}_\gamma^2 = \hat{\sigma}_u^2 + (\hat{\kappa}_\eta - 1)\{\hat{a}'\widehat{\mathcal{J}}^{-1}\hat{a} - (1-\hat{\nu}_1)^2\},$$

$$\hat{a} = \left(0, \frac{1-\hat{\nu}_1}{\hat{\alpha}_n}, \frac{\hat{\nu}_1}{\hat{\beta}_n}\right)', \quad \hat{\nu}_1 = \frac{1}{n}\sum_{t=1}^{n}\frac{\hat{\beta}_n}{\hat{\alpha}_n\hat{\eta}_t^2 + \hat{\beta}_n}.$$

COROLLARY 3.2: *Under the assumptions of Theorem* 3.1,

$$\hat{\sigma}_\gamma^2 \to \sigma_\gamma^2 \quad \textit{in probability (and a.s. when } \gamma_0 \neq 0) \quad \textit{as} \quad n \to \infty. \tag{3.6}$$

Therefore, *at the asymptotic level* $\underline{\alpha} \in (0, 1)$, *a confidence interval for* γ_0 *is*

$$\left[\hat{\gamma}_n - \frac{\hat{\sigma}_\gamma}{\sqrt{n}}\Phi^{-1}(1-\underline{\alpha}/2), \hat{\gamma}_n + \frac{\hat{\sigma}_\gamma}{\sqrt{n}}\Phi^{-1}(1-\underline{\alpha}/2)\right].$$

The following result provides asymptotic critical regions for the strict stationarity testing problems.

COROLLARY 3.3: *Let the assumptions of Theorem 3.1 hold. For the testing problem (3.3), the test defined by the (stationary (ST)) critical region*

$$\text{(3.7)} \qquad \mathrm{C}^{\mathrm{ST}} = \left\{ T_n := \sqrt{n}\frac{\hat{\gamma}_n}{\hat{\sigma}_u} > \Phi^{-1}(1-\underline{\alpha}) \right\}$$

has its asymptotic significance level bounded by $\underline{\alpha}$, has an asymptotic probability of rejection $\underline{\alpha}$ under $\gamma_0 = 0$, and is consistent for all $\gamma_0 > 0$.

For the testing problem (3.4), the test defined by the (nonstationary (NS)) critical region

$$\text{(3.8)} \qquad \mathrm{C}^{\mathrm{NS}} = \{T_n < \Phi^{-1}(\underline{\alpha})\}$$

has its asymptotic significance level bounded by $\underline{\alpha}$, has an asymptotic probability of rejection $\underline{\alpha}$ under $\gamma_0 = 0$, and is consistent for all $\gamma_0 < 0$.

4. TESTING NONSTATIONARITY IN NONLINEAR GARCH

In this section, we study the behavior of the stationarity tests in Section 3.2 when the data are generated by the GARCH-type model

$$\text{(4.1)} \qquad \begin{aligned} &\epsilon_t = \sqrt{h_t}\eta_t, \quad t = 1, 2, \ldots, \\ &h_t = \omega(\eta_{t-1}) + b_0(\eta_{t-1})h_{t-1}, \end{aligned}$$

with an initial value h_0, under the same assumptions on (η_t) as in model (1.1). In this model, $\omega : \mathbb{R} \to [\underline{\omega}, \overline{\omega}]$ and $b_0 : \mathbb{R} \to [\underline{b}, +\infty)$ for some $\underline{\omega}, \overline{\omega}, \underline{b} > 0$. It is assumed that b_0 is decreasing over $(-\infty, 0]$ and increasing over $[0, +\infty)$. This model belongs to the so-called class of augmented GARCH models (see Hörmann (2008)) and encompasses many classes of GARCH(1, 1) models introduced in the literature: for instance, with constant $\omega(\cdot)$, the standard GARCH(1, 1) when $b_0(x) = \alpha_0 x^2 + \beta_0$, and the GJR model (Glosten, Jaganathan, and Runkle (1993)) when $b_0(x) = \alpha_1(\max\{x, 0\})^2 + \alpha_2(\min\{x, 0\})^2 + \beta_0$. It can be shown that if $E\max\{0, \log b_0(\eta_t)\} < \infty$, then

$$\text{(4.2)} \qquad \Gamma := E\log b_0(\eta_t) < 0$$

is a necessary and sufficient condition for the strict stationarity of this model (see, e.g., Francq and Zakoïan (2006)). Our aim is to test strict stationarity without estimating the nonparametric model (4.1). We shall see that, surprisingly, the tests developed for the standard GARCH(1, 1) model still work in this framework.

We still consider the statistic $\hat{\gamma}_n$ defined by (1.8), where $\hat{\theta}_n$ is the estimator (1.7) of the standard GARCH(1, 1) parameter, but the observations are generated by the augmented GARCH(1, 1) model (4.1) instead of the standard GARCH(1, 1).

PROPOSITION 4.1: *Let* $\epsilon_1, \ldots, \epsilon_n$ *denote observations from model* (4.1). *Assume* $0 < E|\log \eta_1^2|^2 < \infty$ *and* $E|\log b_0(\eta_1)|^2 < \infty$.

If $\Gamma > 0$, *then, under regularity conditions that imply the existence and uniqueness of a pseudo-true value* $(\alpha^*, \beta^*) = \lim_{n\to\infty}(\hat{\alpha}_n, \hat{\beta}_n)$ *a.s.*,

$$\hat{\gamma}_n \to \Gamma, \quad \textit{and} \quad \hat{\sigma}_u^2 \to \sigma_*^2 \quad \textit{a.s.} \quad \textit{for some constant} \quad \sigma_*^2 > 0.$$

If $\Gamma < 0$, *then, under regularity conditions that imply the strong consistency of* $\hat{\theta}_n$ *to the unique pseudo-true value*,

$$(\omega^*, \alpha^*, \beta^*)' = \arg\min_{\theta\in\Theta} E\left\{\frac{\epsilon_t^2}{\sigma_t^2(\theta)} + \log \sigma_t^2(\theta)\right\}$$

and if $\operatorname{Var}\log \epsilon_t^2 < \infty$, *we have, for some* Γ^*,

$$\hat{\gamma}_n \to \Gamma^* < 0, \quad \textit{and} \quad \hat{\sigma}_u^2 \to \operatorname{Var}\log\left\{\alpha^* \frac{\epsilon_t^2}{\sigma_t^2(\theta^*)} + \beta^*\right\} > 0 \quad \textit{a.s.}$$

REMARK 4.1: In the ARCH(1) case, under the condition $E\{a(\eta_1)/\eta_1^2\} < \infty$, the pseudo-true value is $\alpha^* = E(a(\eta_1)/\eta_1^2)$ when $\Gamma > 0$.

Thus, the (non)stationarity tests developed in the standard GARCH(1, 1) case lead, asymptotically, to the right decision, even if the GARCH(1, 1) model is misspecified, except in the limit case where $\Gamma = 0$. More precisely, we have the following result.

COROLLARY 4.1: *Let the assumptions of Proposition* 4.1 *hold*.

If $\Gamma > 0$, *then*

$$P(\mathrm{C}^{\mathrm{NS}}) \to 0 \quad \textit{and} \quad P(\mathrm{C}^{\mathrm{ST}}) \to 1,$$

where C^{ST} *and* C^{ST} *are defined in Corollary* 3.3.

If $\Gamma < 0$, *then*

$$P(\mathrm{C}^{\mathrm{ST}}) \to 0 \quad \textit{and} \quad P(\mathrm{C}^{\mathrm{NS}}) \to 1.$$

5. NUMERICAL ILLUSTRATIONS

Before illustrating our asymptotic results for the tests, we study the behavior of the QMLE in finite samples.

5.1. *Finite Sample Properties of the QMLE*

From Theorems 2.1 and 2.2 and Proposition 2.1, we know that $(\hat{\alpha}_n, \hat{\beta}_n)$ is always CAN, whereas $\hat{\omega}_n$ is only consistent in the stationary case. In the Monte

TABLE I

BIAS (MEAN ERRORS) AND MSE (MEAN SQUARED ERRORS) FOR THE QMLE OF A GARCH(1, 1), WITH $\eta_t \sim \mathcal{N}(0, 1)$[a]

	2nd ($\gamma_0 = -0.180$)			ST ($\gamma_0 = -0.038$)			NS ($\gamma_0 = 0.078$)		
	ω	α	β	ω	α	β	ω	α	β
$n = 200$									
Bias	−0.21	0.01	0.01	−0.34	0.01	0.01	−0.51	0.02	0.02
MSE	0.58	0.01	0.01	1.10	0.02	0.02	3.77	0.03	0.03
$n = 4{,}000$									
Bias	0.00	0.00	0.00	−0.03	0.00	0.00	−0.51	0.00	0.00
MSE	0.01	0.00	0.00	0.03	0.00	0.00	4.95	0.00	0.00

[a]The parameters $\theta_0 = (1, 0.3, 0.6)$, $\theta_0 = (1, 0.5, 0.6)$, and $\theta_0 = (1, 0.7, 0.6)$, correspond to second-order stationary (2nd), strict stationary (ST), and nonstationary (NS) models. Bias and MSE are computed over 1,000 independent simulations of length $n = 200$ or 4,000.

Carlo experiments that we conducted, the finite sample behavior of the QMLE is in perfect agreement with these asymptotic results. Table I summarizes a few of these simulation experiments. We simulated 1,000 independent trajectories of size $n = 200$ and $n = 4{,}000$ of GARCH(1, 1) models with $\eta_t \sim \mathcal{N}(0, 1)$ and parameter θ_0 corresponding to a second-order stationary process, a strictly stationary process without second-order moment, and a nonstationary process. Since $\eta_t \sim \mathcal{N}(0, 1)$, the QMLE corresponds to the maximum likelihood estimator (MLE). Similar results were obtained with other distributions for η_t. Concerning the estimation of (α_0, β_0), the results are very similar for the three values of θ_0, confirming that stationarity is not necessary for the estimation of these parameters. By contrast, the column ω in the nonstationary case confirms the asymptotic results of Proposition 2.1 and illustrates the impossibility of estimating the parameter ω_0 with a reasonable accuracy under the nonstationarity condition (1.5). Note that the root mean squared error (RMSE) of estimation of ω_0 even worsens when the sample size increases (3.77 for $n = 200$ and 4.95 for $n = 4{,}000$).

5.2. *Finite Sample Properties of the Tests*

5.2.1. *On Simulated Data*

To assess the performance of the tests developed in Section 3, we simulated $N = 1{,}000$ independent trajectories of size $n = 500$, 2,000, and 4,000 of a GARCH(1, 1) model of the form (1.1), with different values of θ_0 and the standardized Student distribution with 7 degrees of freedom for η_t. The standardized Student distribution is often employed as the distribution of GARCH errors in applied works.

With this distribution, we have $\gamma_0 = 0$ for, in particular, $\alpha_0 = 0.2575$ and $\beta_0 = 0.8$. Results concerning the test of the hypotheses (3.1) with $a = 0$ and

TABLE II

RELATIVE FREQUENCY OF REJECTION (IN %) FOR THE TEST (3.2) OF THE NULL HYPOTHESIS $H_0 : \beta_0 \leq 0.7$ AGAINST $H_1 : \beta_0 > 0.7$[a]

	β_0						
n	0.61	0.64	0.67	0.70	0.73	0.76	0.79
500	3.5	4.3	5.2	8.9	12.6	26.8	49.6
2,000	0.3	0.6	1.8	6.8	18.3	53.1	91.5
4,000	0.2	0.3	1.0	5.5	27.7	76.9	99.0

[a]Here $a = 0$, $b = 1$, and $c = 0.7$ in (3.1). The nominal level is $\underline{\alpha} = 5\%$ when $\alpha_0 = 0.2$. The value $(\alpha_0, \beta_0) = (0.2, 0.7)$ corresponds to a stationary process.

$b = 1$, that is, a test on the value of β_0, are presented in Tables II and III. Note that the test (3.2) behaves similarly when the tested value corresponds to a stationary solution (Table II) or to a nonstationary process (Table III).

We now illustrate the behavior of the strict stationarity tests (3.7) and (3.8), through simulations of the GARCH(1, 1) models with $\beta_0 = 0.8$ and values of α_0 corresponding to $\gamma_0 < 0$ ($\alpha_0 \in \{0.18, 0.20, 0.22\}$), $\gamma_0 = 0$ ($\alpha_0 = 0.2575$), and $\gamma_0 > 0$ ($\alpha_0 \in \{0.28, 0.30, 0.31\}$). Tables IV and V show that, as expected, the frequency of rejection of the C^{ST} test increases with γ_0, while, obviously, that of the C^{NS} test decreases. The rejection frequencies of the two tests approach the nominal level when $\gamma_0 = 0$ and n increases, although it remains far from the theoretical value in Table V. Other simulation experiments (not reported here) reveal that the error of the first kind is better controlled for tests of stationarity of ARCH(1) models, which is not surprising, because the model is simpler.

Now consider testing strict stationarity in a generalized GARCH(1, 1) model using the tests developed for the standard GARCH(1, 1). The generalized GARCH that we considered is a GJR model of the form (4.1) with $b_0(x) = \alpha_1(\max\{x, 0\})^2 + \alpha_2(\min\{x, 0\})^2 + \beta_0$. We keep the same standardized Student distribution for η_t, and we take $\beta_0 = 0.8$ and $\alpha_2 = 0.2575$, whereas α_1 varies in

TABLE III

RELATIVE FREQUENCY OF REJECTION (IN %) FOR THE TEST (3.2) OF THE HYPOTHESIS $H_0 : \beta_0 \leq 0.7$ AGAINST $H_1 : \beta_0 > 0.7$ WHEN $\alpha_0 = 0.5$[a]

	β_0						
n	0.61	0.64	0.67	0.70	0.73	0.76	0.79
500	0.3	0.5	2.8	9.9	25.5	47.7	67.2
2,000	0.0	0.0	0.1	6.2	41.6	81.8	97.0
4,000	0.0	0.0	0.1	6.1	61.0	96.2	99.7

[a]The value $(\alpha_0, \beta_0) = (0.5, 0.7)$ corresponds to a nonstationary process.

TABLE IV

RELATIVE FREQUENCY OF REJECTION (IN %) OF THE TEST (3.7) OF THE STATIONARITY HYPOTHESIS $H_0: \gamma_0 < 0$ FOR THE GARCH(1, 1) MODEL WITH $\beta_0 = 0.8$[a]

	α_0						
n	0.18	0.20	0.22	0.2575	0.28	0.30	0.31
500	0.0	0.0	0.1	7.5	27.8	61.4	75.2
2,000	0.0	0.0	0.0	6.3	67.8	98.6	99.9
4,000	0.0	0.0	0.0	5.3	92.4	100.0	100.0

[a]The nominal level is $\underline{\alpha} = 5\%$. The parameter $\alpha_0 = 0.2575$ corresponds to $\gamma_0 = 0$.

such a way that $\Gamma < 0$ when $\alpha_1 \in \{0.18, 0.20, 0.22\}$, $\Gamma = 0$ when $\alpha_1 = 0.2575$, and $\Gamma > 0$ when $\alpha_1 \in \{0.28, 0.30, 0.31\}$.

Table VI confirms the theoretical result of Section 4. More precisely, for n sufficiently large, the tests give the right conclusion when $\Gamma < 0$ and $\Gamma > 0$. Note that when $\Gamma = 0$, the rejection frequency is close to the nominal 5% level, which is not surprising because the model is a standard GARCH(1, 1) in this case. In general, when the test is applied to nonstandard GARCH models, there is no guarantee that the asymptotic relative frequency of rejection will be close to the nominal asymptotic level of the standard GARCH(1, 1).

5.2.2. *On Real Data*

The strict stationarity tests were then applied to the daily returns of 11 major stock market indices. We considered the CAC, DAX, DJA, DJI, DJT, DJU, FTSE, Nasdaq,[7] Nikkei, SMI, and SP500, from January 2, 1990 to January 22, 2009, except for the indices for which such historical data do not exist. Ta-

TABLE V

RELATIVE FREQUENCY OF REJECTION (IN %) FOR TESTING THE NONSTATIONARITY HYPOTHESIS $H_0: \gamma_0 \geq 0$ WITH THE TEST (3.8) FOR THE GARCH(1, 1) MODEL WITH $\beta_0 = 0.8$[a]

	α_0						
n	0.18	0.20	0.22	0.2575	0.28	0.30	0.31
500	98.3	91.7	69.3	19.8	4.1	0.7	0.4
2,000	100.0	100.0	98.3	11.1	0.1	0.0	0.0
4,000	100.0	100.0	100.0	9.1	0.0	0.0	0.0

[a]The nominal level is $\alpha = 5\%$.

[7]Since the Nasdaq index level was halved on January 3, 1994, one outlier has been eliminated for this series.

TABLE VI

RELATIVE FREQUENCY OF REJECTION (IN %) FOR THE TEST (3.8) OF THE STATIONARITY HYPOTHESIS $H_0 : \alpha < 0$ FOR A GJR MODEL[a]

	α_1						
n	0.18	0.20	0.22	0.2575	0.28	0.30	0.31
500	0.1	0.1	1.1	7.8	15.8	32.7	35.2
2,000	0.0	0.0	0.1	6.6	31.7	65.8	77.4
4,000	0.0	0.0	0.0	5.6	45.1	87.7	96.1

[a]The parameter $\alpha_1 = 0.2575$ corresponds to $\Gamma = 0$. The nominal level is $\alpha = 5\%$.

ble VII displays the test statistics T_n computed on each series. Note that as $n \to \infty$, a.s.

$$T_n = \sqrt{n}\frac{\hat{\gamma}_n - \gamma_0}{\hat{\sigma}_u} + \sqrt{n}\frac{\gamma_0}{\hat{\sigma}_u} \to -\infty$$

when $\gamma_0 < 0$ and $T_n \to +\infty$ when $\gamma_0 > 0$. Because the values of T_n given in Table VII are very small, a nonstationary augmented GARCH(1, 1) model is not plausible, for any of these series.

For individual stock returns, the opposite conclusion can occur as the following examples show. We estimated GARCH(1, 1) models on the daily series of Icagen (NasdaqGM: ICGN), Monarch Community Bancorp (NasdaqCM: MCBF), KV Pharmaceutical (NYSE: KV-A), Community Bankers Trust (AMEX: BTC), and China MediaExpress (NasdaqGS: CCME).[8]

Table VIII shows that for four of these five stocks, the nonstationarity assumption cannot be rejected at any reasonable significance level. Interestingly, nonstationarity can occur with a small or a large ARCH coefficient $\hat{\alpha}_n$. In any case, the value of $\hat{\alpha}_n + \hat{\beta}_n$ does not give clear insight on the possible nonstationarity of the series. As an example, Figure 1 displays the sample path of the

TABLE VII

TEST STATISTIC T_n OF THE STRICT STATIONARITY TESTS (3.7) AND (3.8)[a]

CAC	DAX	DJA	DJI	DJT	DJU	FTSE	Nasdaq	Nikkei	SMI	SP500
−14.5	−15.8	−15.1	−13	−15.1	−14	−10.7	−8.5	−15.4	−23	−11.1

[a]The test statistic is the realization of a random variable which is asymptotically $\mathcal{N}(0, 1)$ distributed when $\gamma_0 = 0$, tends to $-\infty$ under the strict stationarity hypothesis $\gamma_0 < 0$, and tends to $+\infty$ when $\gamma_0 > 0$.

[8]The data range from May 31, 2007, August 28, 2007, March 31, 2006, June 29, 2007, and March 31, 2009, respectively, to February 7, 2011.

TABLE VIII

TEST STATISTIC T_n AND p-VALUES OF THE NONSTATIONARITY TEST (3.8) FOR STOCK RETURNS

	ICGN	MCBF	KV-A	BTC	CCME
n	928	868	1,221	908	469
$\hat{\alpha}_n$	0.581	0.023	0.143	0.508	0.413
$\hat{\beta}_n$	0.696	0.979	0.927	0.765	0.750
T_n	−2.297	0.024	1.120	0.491	0.457
p-value	0.011	0.510	0.869	0.688	0.676

MCBF series. The positive estimated value of γ_0 for this series is in accordance with the seemingly increasing volatility along the sample path.

6. CONCLUSION

This paper develops a unified theory for the inference of both stationary and nonstationary GARCH(1, 1) processes. The practical implications of our results are the following.

(i) If one is interested in inference on (α_0, β_0) in the GARCH(1, 1) model, then stationarity testing is unnecessary. The standard QMLE of (α_0, β_0) is CAN in every (stationary or nonstationary) situation. A key result, allowing one to construct confidence intervals, is that the asymptotic variance of $(\hat{\alpha}_n, \hat{\beta}_n)$ can be estimated without any stationarity restriction.

(ii) If one is interested in a GARCH(1, 1) application in which ω_0 is used (e.g., estimating the variance of today's or tomorrow's conditional distribution

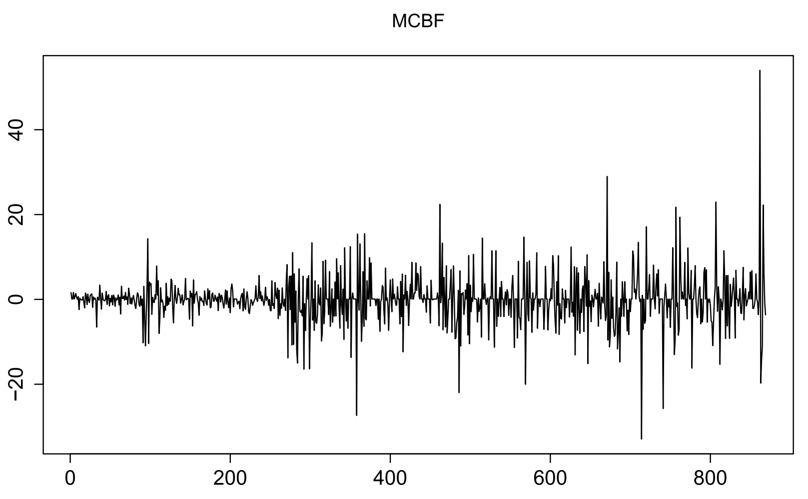

FIGURE 1.—Log returns (in %) of the MCBF stock series.

for derivatives pricing or for computing forecast intervals), then stationarity testing is necessary, because if $\gamma_0 > 0$, the QMLE of ω_0 is not consistent.

(iii) The *constrained* QMLE of (α_0, β_0), which was known to be CAN in the nonstationary case, is inconsistent in the stationary case. As a consequence, this estimator should only be used if nonstationarity is taken for granted.

(iv) Surprisingly, the tests developed for the standard GARCH(1, 1) are able to detect nonstationarity in more general GARCH(1, 1) models.

To conclude, let us mention two possible extensions of this work. First, it would be interesting to know whether the test in this present paper works for detecting stationarity in other volatility models (not restricted to augmented GARCH). Second, it may be worth developing specific stationarity tests for particular augmented GARCH models.

APPENDIX: PROOFS AND COMPLEMENTARY RESULTS

A.1. *Asymptotic Behaviors of* (h_t)

When $\gamma_0 \neq 0$, the asymptotic behavior of the sequences (h_t) (defined by (1.1)) and $(h_{t,\infty})$ (defined by (1.4)) is the same and is easily obtained by the Cauchy rule. When $\gamma_0 = 0$, the asymptotic behavior of $h_{t,\infty}$ can be obtained by the Chung–Fuchs theorem. The behavior of h_t is different in this case and is described in the result below.

PROPOSITION A.1: *For the* GARCH(1, 1) *model* (1.1), *the following properties hold*.

(i) *When* $\gamma_0 > 0$, $h_t \to \infty$ *a.s. at an exponential rate*: *for any* $\rho > e^{-\gamma_0}$,

$$\rho^t h_t \to \infty \quad \textit{and}, \quad \textit{if} \quad E|\log(\eta_1^2)| < \infty,$$
$$\rho^t \epsilon_t^2 \to \infty \quad \textit{a.s.} \quad \textit{as} \quad t \to \infty.$$

(ii) *Klüppelberg, Lindner, and Maller* (2004). *When* $\gamma_0 = 0$,

$$h_t \to \infty \quad \textit{and}, \quad \textit{if} \quad E|\log(\eta_1^2)| < \infty, \quad \epsilon_t^2 \to \infty \quad \textit{in probability}.$$

(iii) *Let* ψ *be a decreasing bijection from* $(0, \infty)$ *to* $(0, \infty)$ *such that* $E\psi(\epsilon_1^2) < \infty$. *When* $\gamma_0 = 0$ *and* $\alpha_0 > 0$,

$$\psi(\epsilon_t^2) \to 0 \quad \textit{and} \quad \psi(h_t) \to 0 \quad \textit{in } L^1. \tag{A.1}$$

PROOF: To prove (i) we note that for $t > 1$,

$$\begin{aligned} h_t &= \omega_0 \left\{ 1 + \sum_{i=1}^{t-1} a_0(\eta_{t-1}) \cdots a_0(\eta_{t-i}) \right\} + a_0(\eta_{t-1}) \cdots a_0(\eta_0) h_0 \\ &\geq \omega_0 \prod_{i=1}^{t-1} a_0(\eta_i). \end{aligned} \tag{A.2}$$

Thus, for any constant $\rho \in (e^{-\gamma_0}, 1)$, we have

$$\liminf_{t\to\infty} \frac{1}{t} \log \rho^t h_t \geq \lim_{t\to\infty} \frac{1}{t}\left\{\log \rho\omega_0 + \sum_{i=1}^{t-1} \log \rho a_0(\eta_i)\right\}$$
$$= E \log \rho a_0(\eta_1) = \log \rho + \gamma_0 > 0.$$

It follows that $\log \rho^t h_t$, and hence $\rho^t h_t$, tends to $+\infty$ a.s. as $n \to \infty$. Now $E|\log \eta_1^2| < \infty$ entails $\log \eta_t^2/t \to 0$ a.s. as $t \to \infty$. Therefore, $\liminf_{t\to\infty} t^{-1} \times \log \rho^t \eta_t^2 h_t \geq E \log \rho a_0(\eta_1) > 0$ and $\rho^t \epsilon_t^2 = \rho^t \eta_t^2 h_t \to +\infty$ a.s. by already given arguments.

The proof of (ii) follows from Klüppelberg, Lindner, and Maller (2004). Their condition $E|\log(\delta + \lambda\varepsilon_1^2)| < \infty$ becomes in our notation $E|\log a_0(\eta_1)| < \infty$, and this condition is satisfied because $E \log^+ a_0(\eta_1) - E \log^- a_0(\eta_1) = \gamma_0 = 0$ and $E \log^+ a_0(\eta_1) \leq \alpha_0 + \beta_0$, where for any real-valued function f, $f^+(x) = \max\{f(x), 0\}$ and $f^-(x) = \max\{-f(x), 0\}$.

To prove (iii), note that since $h_t > \alpha_0 \epsilon_{t-1}^2$ with $\alpha_0 > 0$, we have $\psi(h_t) < \psi^*(\epsilon_{t-1}^2)$, where $\psi^*(x) = \psi(\alpha_0 x)$ satisfies the same assumptions as $\psi(x)$. Therefore, the second convergence in (A.1) follows from the first convergence. It suffices to consider the case $\epsilon_0 = 0$. Note that even if ϵ_t^2 does not increase with probability 1, ϵ_{t+1}^2 is stochastically greater than ϵ_t^2 because

$$\begin{aligned}
\epsilon_{t+1}^2 &= \{\omega_0 + \omega_0 a_0(\eta_t) + \cdots + \omega_0 a_0(\eta_t) \cdots a_0(\eta_2) \\
&\quad + \omega_0 a_0(\eta_t) \cdots a_0(\eta_1)\}\eta_{t+1}^2 \\
&\geq \{\omega_0 + \omega_0 a_0(\eta_t) + \cdots + \omega_0 a_0(\eta_t) \cdots a_0(\eta_2)\}\eta_{t+1}^2 \\
&\overset{d}{=} \epsilon_t^2,
\end{aligned}$$

where $\overset{d}{=}$ stands for equality in distribution. The dominated convergence theorem and (i) and (ii) then entail

$$\begin{aligned}
E\psi(\epsilon_t^2) &= \int_0^\infty P\{\epsilon_t^2 < \psi^{-1}(u)\}\,du \\
&\to \int_0^\infty \lim_{t\to\infty} \downarrow P\{\epsilon_t^2 < \psi^{-1}(u)\}\,du = 0,
\end{aligned}$$

which completes the proof. *Q.E.D.*

A.2. *Asymptotic Normality of the QMLE of* (α_0, β_0)

Let $\underline{\omega} = \inf\{\omega|\theta \in \Theta\}$, $\underline{\alpha} = \inf\{\alpha|\theta \in \Theta\}$, $\underline{\beta} = \inf\{\beta|\theta \in \Theta\}$, $\overline{\omega} = \sup\{\omega|\theta \in \Theta\}$, $\overline{\alpha} = \sup\{\alpha|\theta \in \Theta\}$, and $\overline{\beta} = \sup\{\beta|\theta \in \Theta\}$. Denote by K any constant whose value is unimportant and can change throughout the proofs.

Define the $[0, \infty]$-valued process

$$v_t(\alpha, \beta) = \sum_{j=1}^{\infty} \frac{\alpha\eta_{t-j}^2}{a_0(\eta_{t-j})} \prod_{k=1}^{j-1} \frac{\beta}{a_0(\eta_{t-k})}$$

with the convention $\prod_{k=1}^{j-1} = 1$ when $j \leq 1$. Let $\Theta_0 = \{\theta \in \Theta : \beta < e^{\gamma_0}\}$ and $\Theta_p = \{\theta \in [0, \infty)^3 : \beta < \|1/a_0(\eta_1)\|_p^{-1}\}$.

LEMMA A.1: (i) *When* $\gamma_0 > 0$, *for any* $\theta \in \Theta_0$, *the process* $v_t(\alpha, \beta)$ *is stationary and ergodic. Moreover, for any compact* $\Theta_0^* \subset \Theta_0$,

$$\sup_{\theta \in \Theta_0^*} \left| \frac{\sigma_t^2(\theta)}{h_t} - v_t(\alpha, \beta) \right| \to 0 \quad a.s. \quad as \quad t \to \infty.$$

Finally, for any $\theta \notin \Theta_0$, *it holds that* $\sigma_t^2(\theta)/h_t \to \infty$ *a.s.*

(ii) *When* $\gamma_0 = 0$, *for any* $\theta \in \Theta_p$ *with* $p \geq 1$, *the process* $v_t(\alpha, \beta)$ *is stationary and ergodic. Moreover, for any compact* $\Theta_p^* \subset \Theta_p$,

$$\sup_{\theta \in \Theta_p^*} \left| \frac{\sigma_t^2(\theta)}{h_t} - v_t(\alpha, \beta) \right| \to 0 \quad in\ L^p.$$

PROOF: Without loss of generality, assume that $\sigma_0^2(\theta) = 0$. We then have $\sigma_t^2(\theta) = \sum_{j=1}^t \beta^{j-1}(\omega + \alpha\epsilon_{t-j}^2)$ and

$$\text{(A.3)} \qquad \frac{\sigma_t^2(\theta)}{h_t} = \sum_{j=1}^{t} \beta^{j-1} \left\{ \prod_{k=1}^{j} \frac{h_{t-k}}{h_{t-k+1}} \right\} \frac{\omega + \alpha\epsilon_{t-j}^2}{h_{t-j}} = a_t + b_t,$$

where

$$a_t = \sum_{j=1}^{t} \beta^{j-1} \left\{ \prod_{k=1}^{j} \frac{h_{t-k}}{h_{t-k+1}} \right\} \alpha\eta_{t-j}^2 := \sum_{j=1}^{t} a_{tj} \quad \text{and} \quad b_t = \sum_{j=1}^{t} \beta^{j-1} \frac{\omega}{h_t}.$$

For $\theta \in \Theta_0$, by the Cauchy root test, the series $v_t(\alpha, \beta)$ in a.s. finite. As a measurable function of $\{\eta_u, u < t\}$, the process $v_t(\alpha, \beta)$ is thus stationary and ergodic. We have, for $\overline{\beta}_0 = \sup\{\beta | \theta \in \Theta_0^*\}$, $b_t \leq K\overline{\omega}(t + \overline{\beta}_0^t)/h_t \to 0$ a.s. by Proposition A.1(i). It follows that $\sup_{\theta \in \Theta_0^*} b_t \to 0$ a.s. Now note that

$$\text{(A.4)} \qquad \frac{h_{t-k}}{h_{t-k+1}} = \frac{h_{t-k}}{\omega_0 + a_0(\eta_{t-k})h_{t-k}} \leq \frac{1}{a_0(\eta_{t-k})}.$$

As in the proof of Lemma 4 in Jensen and Rahbek (2004b), for any fixed $t_0 < t$, we thus have

$$0 \leq v_t(\alpha,\beta) - a_t \leq \sum_{j=1}^{t_0}(v_{tj} - a_{tj}) + \sum_{j=t_0+1}^{\infty} v_{tj},$$

where $v_t(\alpha,\beta) = \sum_{j=1}^{\infty} v_{tj}$. In the case $\gamma_0 > 0$, $\sup_{\theta\in\Theta_0^*}(v_{tj} - a_{tj}) \to 0$ a.s. as $t \to \infty$ by Proposition A.1(i). Moreover, the series $\sup_{\theta\in\Theta_0^*}\sum_{j=t_0+1}^{\infty} v_{tj} = \sum_{j=t_0+1}^{\infty} v_{tj}(\overline{\alpha}_0, \overline{\beta}_0)$, with obvious notations, converges by the Cauchy root test. The first convergence in (i) follows.

Now for $\theta \notin \Theta_0$, for any $t_0 < t$, we have, for $\rho > e^{-\gamma_0}$,

$$\frac{\sigma_t^2(\theta)}{h_t} \geq \sum_{j=1}^{t_0} a_{tj} = \sum_{j=1}^{t_0} v_{tj} + o(\rho^t t_0^2) \quad \text{a.s.}$$

as $t \to \infty$ by Proposition A.1(i). The proof of (i) is completed by noting that $\sum_{j=1}^{t_0} v_{tj} \to \infty$ a.s. as $t_0 \to \infty$ by the Cauchy root test when $\beta > e^{\gamma_0}$ and by the Chung–Fuchs theorem when $\beta = e^{\gamma_0}$.

Now we turn to (ii). Since $\|v_t(\alpha,\beta)\|_p < \infty$, $v_t(\alpha,\beta)$ is a.s. finite, and the stationarity and ergodicity follow. By $\gamma_0 = 0 = E\log(\alpha_0\eta_1^2 + \beta_0)^{-p}$ and Jensen's inequality, we have $\|(\alpha_0\eta_1^2 + \beta_0)^{-1}\|_p > 1$. Thus $\theta \in \Theta_p$ entails $\beta < 1$. It follows that $\sup_{\theta\in\Theta_p^*} b_t \to 0$ in L^p using Proposition A.1(iii). Noting that $\|v_{tj}\|_p \leq (\alpha/\alpha_0)\rho^{j-1}$, where $\rho = \|\beta/(\alpha_0\eta_1 + \beta_0)\|_p < 1$, the rest of the proof follows from arguments similar to those used in the proof of (i) with a.s. convergences replaced by L^p convergences. *Q.E.D.*

LEMMA A.2: *If* $\theta \in \Theta_0$, *we have*

$$v_t(\alpha,\beta) = 1 \quad a.s. \quad \text{if and only if} \quad (\alpha,\beta) = (\alpha_0,\beta_0).$$

PROOF: Straightforward algebra shows that

$$v_t(\alpha,\beta)(\alpha_0\eta_{t-1}^2 + \beta_0) = \beta v_{t-1}(\alpha,\beta) + \alpha\eta_{t-1}^2. \tag{A.5}$$

Hence

$$\{v_t(\alpha,\beta) - 1\}(\alpha_0\eta_{t-1}^2 + \beta_0) = \beta v_{t-1}(\alpha,\beta) - \beta_0 + (\alpha - \alpha_0)\eta_{t-1}^2.$$

It follows that $v_t(\alpha,\beta) = 1$ a.s. if and only if (iff) $\beta v_{t-1}(\alpha,\beta) - \beta_0 + (\alpha - \alpha_0)\eta_{t-1}^2 = 0$. By strict stationarity, $v_{t-1}(\alpha,\beta) = 1$ a.s. and we have $\beta - \beta_0 + (\alpha - \alpha_0)\eta_{t-1}^2 = 0$. Because the distribution of η_{t-1}^2 is nondegenerate, the conclusion follows. *Q.E.D.*

PROOF OF THEOREM 2.1(i)–(iii): The result stated in (i) is standard. Consider the case (ii). Note that $(\hat{\omega}_n, \hat{\alpha}_n, \hat{\beta}_n) = \arg\min_{\theta\in\Theta} Q_n(\theta)$, where $Q_n(\theta) = n^{-1}\sum_{t=1}^n \{\ell_t(\theta) - \ell_t(\theta_0)\}$. We have

$$Q_n(\theta) = \frac{1}{n}\sum_{t=1}^n \eta_t^2\left\{\frac{h_t}{\sigma_t^2(\theta)} - 1\right\} + \log\frac{\sigma_t^2(\theta)}{h_t} = O_n(\alpha, \beta) + R_n(\theta),$$

where

$$O_n(\alpha, \beta) = \frac{1}{n}\sum_{t=1}^n \eta_t^2\left\{\frac{1}{v_t(\alpha, \beta)} - 1\right\} + \log v_t(\alpha, \beta)$$

and

$$R_n(\theta) = \frac{1}{n}\sum_{t=1}^n \eta_t^2\left\{\frac{h_t}{\sigma_t^2(\theta)} - \frac{1}{v_t(\alpha, \beta)}\right\} + \log\frac{\sigma_t^2(\theta)}{h_t v_t(\alpha, \beta)}.$$

Lemma A.1(i) entails that if $\theta \notin \Theta_0$, then $Q_n(\theta) \to \infty$ a.s. It thus suffices to consider the case $\theta \in \Theta_0^*$, where Θ_0^* is an arbitrary compact subset of Θ_0. We have by stationarity and ergodicity of $v_t(\alpha, \beta)$ a.s.

$$\lim_{n\to\infty} O_n(\alpha, \beta) = E\left\{\frac{1}{v_1(\alpha, \beta)} - 1 + \log v_1(\alpha, \beta)\right\} \geq 0$$

because $\log x \leq x - 1$ for $x > 0$. The inequality is strict except when $v_1(\alpha, \beta) = 1$ a.s. By Lemma A.2, we thus have $E\{O_n(\alpha, \beta)\} \geq 0$ with equality only if $(\alpha, \beta) = (\alpha_0, \beta_0)$.

To handle $R_n(\theta)$, we prove the following lemma. Let $\Theta_{\alpha,\beta}$ be the compact set of the (α, β)'s such that $(\omega, \alpha, \beta)' \in \Theta$.

LEMMA A.3: *Suppose that $P(\eta_t = 0) = 0$. Then, for any $k > 0$,*

$$E\sup_{(\alpha,\beta)\in\Theta_{\alpha,\beta}}\left(\frac{1}{v_t(\alpha, \beta)}\right)^k < \infty \quad \textit{and} \quad E\sup_{\theta\in\Theta}\left(\frac{h_t}{\sigma_t^2(\theta)}\right)^k < \infty.$$

PROOF: Let $\varepsilon > 0$ such that $p(\varepsilon) := P(|\eta_t| \leq \varepsilon) \in [0, 1)$. If $|\eta_{t-1}| > \varepsilon$, since the sum $v_t(\alpha, \beta)$ is greater than its first term, we have

$$\frac{1}{v_t(\alpha, \beta)} \leq \frac{\alpha_0\eta_{t-1}^2 + \beta_0}{\alpha\eta_{t-1}^2} = \frac{\alpha_0}{\alpha} + \frac{\beta_0}{\alpha\eta_{t-1}^2} \leq \frac{\alpha_0}{\underline{\alpha}} + \frac{\beta_0}{\underline{\alpha}\varepsilon^2} := K(\varepsilon).$$

Now if $|\eta_{t-1}| \leq \varepsilon$ but $|\eta_{t-2}| > \varepsilon$, minorizing the sum $v_t(\alpha, \beta)$ by its second term, we have

$$\frac{1}{v_t(\alpha, \beta)} \leq \left(\frac{\alpha_0}{\alpha} + \frac{\beta_0}{\alpha\eta_{t-2}^2}\right)\frac{a_0(\varepsilon)}{\beta} \leq K(\varepsilon)\frac{a_0(\varepsilon)}{\underline{\beta}}.$$

Continuing in this manner, we can write

$$\sup_{(\alpha,\beta)\in\Theta_{\alpha,\beta}} \frac{1}{v_t(\alpha,\beta)} \le K(\varepsilon)\sum_{i=1}^{\infty}\mathbf{1}_{|\eta_{t-1}|\le\varepsilon}\cdots\mathbf{1}_{|\eta_{t-i+1}|\le\varepsilon}\mathbf{1}_{|\eta_{t-i}|>\varepsilon}\left(\frac{a_0(\varepsilon)}{\underline{\beta}}\right)^{i-1}.$$

Thus, for any integer k,

$$\begin{aligned}&\sup_{(\alpha,\beta)\in\Theta_{\alpha,\beta}}\left(\frac{1}{v_t(\alpha,\beta)}\right)^k\\ &\quad\le\{K(\varepsilon)\}^k\sum_{i=1}^{\infty}\mathbf{1}_{|\eta_{t-1}|\le\varepsilon}\cdots\mathbf{1}_{|\eta_{t-i+1}|\le\varepsilon}\mathbf{1}_{|\eta_{t-i}|>\varepsilon}\left(\frac{a_0(\varepsilon)}{\underline{\beta}}\right)^{k(i-1)}.\end{aligned}$$

It follows that

$$\begin{aligned}&E\sup_{(\alpha,\beta)\in\Theta_{\alpha,\beta}}\left(\frac{1}{v_t(\alpha,\beta)}\right)^k\\ &\quad\le\{K(\varepsilon)\}^k\{1-p(\varepsilon)\}\sum_{i=1}^{\infty}p(\varepsilon)^{i-1}\left(\frac{a_0(\varepsilon)}{\underline{\beta}}\right)^{k(i-1)}.\end{aligned}$$

Noting that $\lim_{\varepsilon\to 0}p(\varepsilon)=0$ and $\lim_{\varepsilon\to 0}a_0(\varepsilon)=\beta_0$, we have $p(\varepsilon)(\frac{a_0(\varepsilon)}{\underline{\beta}})^k<1$ for ε sufficiently small. The first result of the lemma is thus proven.

Similarly, we have for $|\eta_{t-1}|>\varepsilon$,

$$\frac{h_t}{\sigma_t^2(\theta)}\le\frac{\omega_0+a_0(\eta_{t-1})h_{t-1}}{\omega+\alpha h_{t-1}\eta_{t-1}^2+\beta\sigma_{t-1}^2(\theta)}\le\frac{\omega_0}{\underline{\omega}}+\frac{\alpha_0}{\underline{\alpha}}+\frac{\beta_0}{\underline{\alpha}\varepsilon^2}:=H(\varepsilon),$$

and for $|\eta_{t-1}|\le\varepsilon$ and $|\eta_{t-2}|>\varepsilon$,

$$\frac{h_t}{\sigma_t^2(\theta)}\le\frac{\omega_0}{\underline{\omega}}+\frac{a_0(\varepsilon)}{\underline{\beta}}H(\varepsilon).$$

More generally,

$$\sup_{\theta\in\Theta}\frac{h_t}{\sigma_t^2(\theta)}\le V_t, \tag{A.6}$$

where

$$\begin{aligned}V_t=&\sum_{i=1}^{\infty}\mathbf{1}_{|\eta_{t-1}|\le\varepsilon}\cdots\mathbf{1}_{|\eta_{t-i+1}|\le\varepsilon}\mathbf{1}_{|\eta_{t-i}|>\varepsilon}\\ &\times\left(\frac{\omega_0}{\underline{\omega}}\sum_{j=0}^{i-2}\left(\frac{a_0(\varepsilon)}{\underline{\beta}}\right)^j+\left(\frac{a_0(\varepsilon)}{\underline{\beta}}\right)^{i-1}H(\varepsilon)\right).\end{aligned}$$

The conclusion follows by the same arguments as before. *Q.E.D.*

Now we show that

$$\lim_{n\to\infty} \sup_{\theta\in\Theta_0^*} |R_n(\theta)| = 0 \quad \text{a.s.,} \tag{A.7}$$

where Θ_0^* is defined in Lemma A.1. Using Lemma A.1, the absolute value of the first term in $R_n(\theta)$ satisfies, using (A.6),

$$\begin{aligned} \sup_{\theta\in\Theta_0^*}\left|\frac{1}{n}\sum_{t=1}^{n}\eta_t^2\frac{h_t}{\sigma_t^2(\theta)}\left\{v_t(\alpha,\beta)-\frac{\sigma_t^2(\theta)}{h_t}\right\}\frac{1}{v_t(\alpha,\beta)}\right| \\ \leq \frac{K\varepsilon}{n}\sum_{t=1}^{n}\eta_t^2\frac{V_t}{v_t(\underline{\alpha},\underline{\beta})} \end{aligned} \tag{A.8}$$

for any $\varepsilon > 0$, when n is large enough. The right-hand side tends a.s. to $K\varepsilon$ as $n\to\infty$ by the ergodic theorem and Lemma A.3. The second term in $R_n(\theta)$ is handled similarly and (A.7) follows. The proof is completed by standard arguments, invoking the compactness of Θ.

The proof of (iii) is identical, except that the a.s. convergence in (A.7) is replaced by an L^1 convergence. More precisely, by the Hölder inequality and Lemma A.3, the expectation of the left-hand side of (A.8) is bounded by

$$\frac{K}{n}\sum_{t=1}^{n}\left\|\frac{V_t}{v_t(\underline{\alpha},\underline{\beta})}\right\|_q \sup_{\theta\in\Theta_p}\left\|v_t(\alpha,\beta)-\frac{\sigma_t^2(\theta)}{h_t}\right\|_p,$$

which tends to zero by Lemma A.1(ii). *Q.E.D.*

We now need to introduce new $[0,\infty]$-valued processes. Let $a(\eta_t)=\alpha\eta_t^2+\beta$ and

$$d_t^{\alpha}(\alpha,\beta)=\sum_{j=1}^{\infty}\frac{\eta_{t-j}^2}{a(\eta_{t-j})}\prod_{k=1}^{j-1}\frac{\beta}{a(\eta_{t-k})},$$

$$d_t^{\beta}(\alpha,\beta)=\sum_{j=2}^{\infty}\frac{(j-1)\alpha\eta_{t-j}^2}{\beta a(\eta_{t-j})}\prod_{k=1}^{j-1}\frac{\beta}{a(\eta_{t-k})}.$$

LEMMA A.4: *Assume $\gamma_0 \geq 0$ and $E\eta_t^4<\infty$. We have*

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\begin{pmatrix}\dfrac{\partial}{\partial\alpha}\ell_t(\theta_0)\\[2mm] \dfrac{\partial}{\partial\beta}\ell_t(\theta_0)\end{pmatrix}\xrightarrow{d}\mathcal{N}\{0,(\kappa_\eta-1)\mathcal{I}\}\quad \textit{as}\quad n\to\infty.$$

PROOF: Using the Wold–Cramér device, it suffices to show that for all $\lambda = (\lambda_1, \lambda_2)'$, the sequence

$$\nabla_t = \left(\frac{\partial}{\partial\alpha}\ell_t(\theta_0), \frac{\partial}{\partial\beta}\ell_t(\theta_0)\right)\lambda = \frac{1-\eta_t^2}{h_t}\lambda' \begin{pmatrix} \frac{\partial\sigma_t^2}{\partial\alpha}(\theta_0) \\ \frac{\partial\sigma_t^2}{\partial\beta}(\theta_0) \end{pmatrix}$$

satisfies

$$\text{(A.9)} \qquad \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\nabla_t \xrightarrow{d} \mathcal{N}\{0, (\kappa_\eta - 1)\lambda'\mathcal{I}\lambda\}.$$

Since $E\log\beta/a(\eta_1) < 0$, by the Cauchy root test, the processes $d_t^\alpha(\alpha,\beta)$ and $d_t^\beta(\alpha,\beta)$ are stationary and ergodic. Still assuming $\sigma_0^2 = 0$, we have

$$\frac{\partial\sigma_t^2}{\partial\alpha}(\theta) = \sum_{j=1}^{t}\beta^{j-1}\epsilon_{t-j}^2, \qquad \frac{\partial\sigma_t^2}{\partial\beta}(\theta) = \sum_{j=2}^{t}(j-1)\beta^{j-2}(\omega + \alpha\epsilon_{t-j}^2).$$

Thus, using a direct extension of (A.4),

$$\frac{1}{\sigma_t^2(\theta)}\frac{\partial\sigma_t^2}{\partial\alpha}(\theta) = \sum_{j=1}^{t}\beta^{j-1}\left\{\prod_{k=1}^{j}\frac{\sigma_{t-k}^2(\theta)}{\sigma_{t-k+1}^2(\theta)}\right\}\frac{\epsilon_{t-j}^2}{\sigma_{t-j}^2(\theta)} \le d_t^\alpha(\alpha,\beta),$$

$$\frac{1}{\sigma_t^2(\theta)}\frac{\partial\sigma_t^2}{\partial\beta}(\theta) = \sum_{j=2}^{t}(j-1)\beta^{j-2}\left\{\prod_{k=1}^{j}\frac{\sigma_{t-k}^2(\theta)}{\sigma_{t-k+1}^2(\theta)}\right\}\frac{\omega + \alpha\epsilon_{t-j}^2}{\sigma_{t-j}^2(\theta)} \le d_t^\beta(\alpha,\beta).$$

Moreover, we have

$$0 \le d_t^\alpha(\alpha,\beta) - \frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\alpha}(\theta) \le s_{t_0} + r_{t_0},$$

where

$$s_{t_0} = \sum_{j=1}^{t_0}\frac{\eta_{t-j}^2}{a(\eta_{t-j})}\prod_{k=1}^{j-1}\frac{\beta}{a(\eta_{t-k})} - \frac{\epsilon_{t-j}^2}{\sigma_{t-j}^2(\theta)}\prod_{k=1}^{j}\frac{\beta\sigma_{t-k}^2(\theta)}{\sigma_{t-k+1}^2(\theta)},$$

$$r_{t_0} = \sum_{j=t_0+1}^{\infty}\frac{\eta_{t-j}^2}{a(\eta_{t-j})}\prod_{k=1}^{j-1}\frac{\beta}{a(\eta_{t-k})}.$$

For all $p \geq 1$, $\|r_{t_0}\|_p \to 0$ as $t_0 \to \infty$ because $\|\beta/a(\eta_1)\|_p < 1$. Since $\|\eta_1^2/a(\eta_1)\|_p < 1/\alpha$, $\|\beta\sigma_{t-1}^2(\theta)/\sigma_t^2(\theta)\|_p < 1$, and

$$\left\| \frac{\beta}{a(\eta_t)} - \frac{\beta\sigma_{t-1}^2(\theta)}{\sigma_t^2(\theta)} \right\|_p = \left\| \frac{\beta\omega}{a(\eta_t)\sigma_t^2(\theta)} \right\|_p \to 0$$

as $t \to \infty$ by the dominated convergence theorem, $s_{t_0} = s_{t_0}(t)$ converges to 0 in L^p as $t \to \infty$. The same derivations hold true when $d_t^\alpha(\alpha,\beta)$ is replaced by $d_t^\beta(\alpha,\beta)$. It follows that $d_t^\alpha(\alpha,\beta)$ and $d_t^\alpha(\beta)$ have moments of any order, and

$$\text{(A.10)} \quad \left| \frac{1}{\sigma_t^2(\theta)} \frac{\partial\sigma_t^2}{\partial\alpha}(\theta) - d_t^\alpha(\alpha,\beta) \right| \to 0 \quad \text{and}$$
$$\left| \frac{1}{\sigma_t^2(\theta)} \frac{\partial\sigma_t^2}{\partial\beta}(\theta) - d_t^\beta(\alpha,\beta) \right| \to 0$$

in L^p for any $p \geq 1$. Standard computations show that $\mathcal{I} = Ed_1d_1'$, where $d_t' = (d_t^\alpha(\alpha_0,\beta_0), d_t^\beta(\alpha_0,\beta_0))$. The same matrix was obtained by Jensen and Rahbek (2004b).

Using (A.10) at $\theta = \theta_0$ and the ergodic theorem, it then follows that, as $n \to \infty$,

$$\operatorname{Var} \frac{1}{\sqrt{n}} \sum_{t=1}^n \nabla_t = \frac{\kappa_\eta - 1}{n} \sum_{t=1}^n E|\lambda' d_t|^2 + o(1) \to (\kappa_\eta - 1)\lambda'\mathcal{I}\lambda.$$

Moreover, for all $\varepsilon > 0$,

$$\begin{aligned} \frac{1}{n}\sum_{t=1}^n E\nabla_t^2 1_{\{|\nabla_t/\sqrt{n}|>\varepsilon\}} &\leq (\kappa_\eta - 1)E\{|\lambda_1|d_1^\alpha(\beta_0) + |\lambda_2|d_1^\beta(\alpha_0,\beta_0)\}^2 \\ &\quad \times 1_{\{(|1-\eta_1^2|/\sqrt{n})(|\lambda_1|d_1^\alpha(\beta_0)+|\lambda_2|d_1^\beta(\alpha_0,\beta_0))>\varepsilon\}} \\ &\to 0. \end{aligned}$$

We thus obtain (A.9) by the Lindeberg central limit theorem for martingale differences (see Billingsley (1995, p. 476)). *Q.E.D.*

LEMMA A.5: *Let* ϖ *be an arbitrary compact subset of* $[0,\infty)$. *Assume that* $E\log\eta_1^2 < \infty$. *When* $\gamma_0 > 0$, *we have*

$$\text{(A.11)} \quad \sum_{t=1}^\infty \sup_{\theta\in\Theta_0} \left| \frac{\partial}{\partial\omega} \ell_t(\theta) \right| < \infty \quad a.s.,$$

$$\text{(A.12)} \quad \sum_{t=1}^\infty \sup_{\theta\in\Theta_0} \left\| \frac{\partial^2}{\partial\omega\,\partial\theta} \ell_t(\theta) \right\| < \infty \quad a.s.,$$

$$\text{(A.13)}\quad \sup_{\omega\in\varpi}\left|\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2\ell_t(\omega,\alpha_0,\beta_0)}{\partial\theta_{i+1}\partial\theta_{j+1}}-\mathcal{I}(i,j)\right|=o(1)\quad a.s.\quad \textit{for all } i,j\in\{1,2\},$$

$$\text{(A.14)}\quad \frac{1}{n}\sum_{t=1}^{n}\sup_{\theta\in\Theta}\left|\frac{\partial^3}{\partial\theta_i\partial\theta_j\partial\theta_k}\ell_t(\theta)\right|=O(1)\quad a.s.\quad \textit{for all } i,j,k\in\{2,3\}.$$

When $\gamma_0=0$,

$$\text{(A.15)}\quad \sup_{\omega\in\varpi}\left|\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2\ell_t(\omega,\alpha_0,\beta_0)}{\partial\theta_{i+1}\partial\theta_{j+1}}-\mathcal{I}(i,j)\right|=o_P(1)\quad \textit{for all } i,j\in\{1,2\},$$

$$\text{(A.16)}\quad \frac{1}{n}\sum_{t=1}^{n}\sup_{\theta\in\Theta_4}\left|\frac{\partial^3}{\partial\theta_i\partial\theta_j\partial\theta_k}\ell_t(\theta)\right|=O_P(1)\quad \textit{for all } i,j,k\in\{2,3\}.$$

PROOF: Let us first suppose $\gamma_0>0$. Using Proposition A.1(i) and Lemma A.1(i), $\sigma_t^2(\theta)>\alpha\epsilon_{t-1}^2$, and using arguments similar to those used to show $b_t\to 0$ in Lemma A.1, for $\theta\in\Theta_0$, there exist a random variable K and $\rho\in(\beta e^{-\gamma_0},1)$ such that for t large enough,

$$\text{(A.17)}\quad \left|\frac{\partial}{\partial\omega}\ell_t(\theta)\right|=\left|\frac{-h_t\eta_t^2}{\sigma_t^4(\theta)}+\frac{1}{\sigma_t^2(\theta)}\right|\sum_{j=1}^{t}\beta^{j-1}\le K\rho^t\left(\frac{\eta_t^2}{v_t(\alpha,\beta)}+1\right)\quad \text{a.s.}$$

Since, in view of Lemma A.3, $\sum_{t=1}^{\infty}K\rho^t(\eta_t^2/v_t(\alpha,\beta)+1)$ has a finite expectation, it is a.s. finite, uniformly in $\theta\in\Theta_0$. Thus (A.11) is proved and (A.12) can be obtained by the same arguments. For brevity, we only prove (A.13) in the case $i=2$ and $j=3$. First note that $\mathcal{I}(1,2)=\lim_{n\to\infty}n^{-1}\sum_{t=1}^{n}d_t^{\alpha}(\alpha_0,\beta_0)d_t^{\beta}(\alpha_0,\beta_0)$ a.s. Moreover, we have

$$\begin{aligned}\frac{\partial^2\ell_t(\omega,\alpha_0,\beta_0)}{\partial\alpha\partial\beta}&=\left(2\eta_t^2\frac{h_t}{\sigma_t^2}-1\right)\left(\frac{1}{\sigma_t^2}\sum_{j=1}^{t}\beta_0^{j-1}\epsilon_{t-j}^2\right)\\&\quad\times\left(\frac{1}{\sigma_t^2}\sum_{j=2}^{t}(j-1)\beta_0^{j-2}(\omega+\alpha_0\epsilon_{t-j}^2)\right)\\&\quad+\left(1-\eta_t^2\frac{h_t}{\sigma_t^2}\right)\left(\frac{1}{\sigma_t^2}\sum_{j=2}^{t}(j-1)\beta_0^{j-2}\epsilon_{t-j}^2\right),\end{aligned}$$

where $\sigma_t^2=\sigma_t^2(\omega,\alpha_0,\beta_0)$. We obtain the result by showing that, a.s. as $t\to\infty$,

$$\text{(A.18)}\quad \left|\frac{\sum_{j=1}^{t}\beta_0^{j-1}\epsilon_{t-j}^2}{h_t}-d_t^{\alpha}(\alpha_0,\beta_0)\right|\to 0,$$

$$\text{(A.19)}\qquad \sup_{\omega\in\varpi}\left|\frac{\sum_{j=2}^{t}(j-1)\beta_0^{j-2}(\omega+\alpha_0\epsilon_{t-j}^2)}{h_t}-d_t^{\beta}(\alpha_0,\beta_0)\right|\to 0,$$

$$\text{(A.20)}\qquad \left|\frac{\sum_{j=2}^{t}(j-1)\beta_0^{j-2}\epsilon_{t-j}^2}{h_t}-\frac{1}{\alpha_0}d_t^{\beta}(\alpha_0,\beta_0)\right|\to 0,$$

$$\text{(A.21)}\qquad \sup_{\omega\in\varpi}\left|\frac{h_t}{\sigma_t^2}-1\right|\to 0.$$

In view of $\sigma_t^2 - h_t = \sum_{j=1}^{t}\beta_0^{j-1}(\omega-\omega_0)$ and $1/h_t = o(\rho^t)$ for some $\rho \in (e^{-\gamma_0}, \beta_0^{-1})$, the convergence (A.21) holds. The convergences in (A.18)–(A.20) are obtained by the arguments used to establish the first convergence in Lemma A.1. Next we turn to (A.14). For instance consider the case $i=j=2$ and $k=3$. We have, now with $\sigma_t^2=\sigma_t^2(\theta)$,

$$\begin{aligned}
\frac{\partial^3\ell_t(\theta)}{\partial^2\alpha\,\partial\beta} &= \left(2-6\eta_t^2\frac{h_t}{\sigma_t^2}\right)\left(\frac{1}{\sigma_t^2}\sum_{j=1}^{t}\beta^{j-1}\epsilon_{t-j}^2\right)^2 \\
&\quad\times\left(\frac{1}{\sigma_t^2}\sum_{j=2}^{t}(j-1)\beta^{j-2}(\omega+\alpha\epsilon_{t-j}^2)\right) \\
&\quad+\left(2\eta_t^2\frac{h_t}{\sigma_t^2}-1\right)\left(\frac{1}{\sigma_t^2}\sum_{j=1}^{t}\beta^{j-1}\epsilon_{t-j}^2\right)\left(\frac{2}{\sigma_t^2}\sum_{j=2}^{t}(j-1)\beta^{j-2}\epsilon_{t-j}^2\right) \\
&= \left(2-3\eta_t^2\frac{1}{v_t(\alpha,\beta)}\right)(d_t^{\alpha}(\alpha,\beta))^2 d_t^{\beta}(\alpha,\beta) \\
&\quad+\left(2\eta_t^2\frac{1}{v_t(\alpha,\beta)}-1\right)d_t^{\alpha}(\alpha,\beta)\frac{2}{\alpha}d_t^{\beta}(\alpha,\beta)+o(1)\quad\text{a.s.,}
\end{aligned}$$

where the term $o(1)$ is again obtained by arguments similar to those used to show the first convergence in Lemma A.1. Noting that $d_t^{\alpha}(\alpha,\beta)$ and $d_t^{\beta}(\alpha,\beta)$ admit moments of any order, (A.14) then follows from the ergodic theorem, Lemma A.3, and the Cauchy–Schwarz inequality.

In the case $\gamma=0$, the a.s. convergences of the proof of (A.13) and (A.14) can be replaced by L^p convergences via the arguments used to show (A.10). We then obtain (A.15) and (A.16) from Proposition A.1(iii). *Q.E.D.*

PROOF OF THEOREM 2.1(iv) and (v): Part (iv) has already been proven (see Berkes, Horváth, and Kokoszka (2003) and Francq and Zakoïan (2004)).

It remains to prove the asymptotic normality of $(\hat{\alpha}_n, \hat{\beta}_n)$ when $\gamma_0 > 0$. Notice that we cannot use the fact that the derivative of the criterion cancels at $\hat{\theta}_n = (\hat{\omega}_n, \hat{\alpha}_n, \hat{\beta}_n)$, since we have no consistency result for $\hat{\omega}_n$. Thus the minimum could lie on the boundary of Θ, even asymptotically. However, the partial derivative with respect to (α, β) is asymptotically equal to zero at the minimum, since $(\hat{\alpha}_n, \hat{\beta}_n) \to (\alpha_0, \beta_0)$ and θ_0 belongs to the interior of Θ. Hence, an expansion of the criterion derivative gives

$$\text{(A.22)} \quad \begin{pmatrix} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial \omega} \ell_t(\hat{\theta}_n) \\ 0 \end{pmatrix} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial \theta} \ell_t(\theta_0) + \mathcal{J}_n \sqrt{n}(\hat{\theta}_n - \theta_0),$$

where $\mathcal{J}_n$ is a 3×3 matrix whose elements have the form

$$\mathcal{J}_n(i, j) = \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell_t(\theta_i^*),$$

where $\theta_i^* = (\omega_i^*, \alpha_i^*, \beta_i^*)'$ is between $\hat{\theta}_n$ and θ_0. Since $\theta_0 \in \Theta_0$, we have $\theta_i^* \in \Theta_0$ for n large enough. By (A.12) in Lemma A.5 and the compactness of Θ, we have, for $i = 2, 3$,

$$\text{(A.23)} \quad \mathcal{J}_n(i, 1)\sqrt{n}(\hat{\omega}_n - \omega_0) \le \sum_{t=1}^{\infty} \sup_{\theta \in \Theta_0} \left\| \frac{\partial^2}{\partial \omega \, \partial \theta} \ell_t(\theta) \right\| \frac{1}{\sqrt{n}}(\hat{\omega}_n - \omega_0) \to 0 \quad \text{a.s.}$$

An expansion of the function

$$(\alpha, \beta) \mapsto \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial \alpha^2} \ell_t(\omega_2^*, \alpha, \beta)$$

gives

$$\mathcal{J}_n(2, 2) = \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial \alpha^2} \ell_t(\omega_2^*, \alpha_0, \beta_0) + \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^3 \ell_t(\omega_2^*, \alpha^*, \beta^*)}{\partial(\alpha, \beta) \, \partial \alpha^2} \begin{pmatrix} \alpha_2^* - \alpha_0 \\ \beta_2^* - \beta_0 \end{pmatrix},$$

where (α^*, β^*) is between (α_2^*, β_2^*) and (α_0, β_0). Using (A.13), (A.14), and the consistency of (α^*, β^*), we get $\mathcal{J}_n(2, 2) \to \mathcal{I}(1, 1)$ a.s., and similarly for $i, j = 1, 2$,

$$\text{(A.24)} \quad \mathcal{J}_n(i + 1, j + 1) \to \mathcal{I}(i, j) \quad \text{a.s.}$$

The conclusion follows by considering the last two components in (A.22), and from Lemma A.4, (A.23), and (A.24). *Q.E.D.*

PROOF OF THEOREM 2.2: The proof of the asymptotic normality still relies on the Taylor expansion (A.22). The asymptotic distribution of the first term in the right-hand side of (A.22) is still given by Lemma A.4. To deal with the second term, we cannot use (A.23) because (A.12) requires $\gamma_0 > 0$. Instead, noting that

$$\frac{1}{\sigma_t^2(\theta)} \sum_{j=1}^{t} \beta^{j-1} \epsilon_{t-j}^2 \leq \frac{1}{\alpha}$$

and $\beta_2^* < 1$ for n large enough, we obtain

$$\text{(A.25)} \quad |\mathcal{J}_n(2,1)\sqrt{n}(\hat{\omega}_n - \omega_0)|$$
$$\leq \frac{K}{\sqrt{n}} \sum_{t=1}^{n} \left(\frac{2h_t\eta_t^2}{\sigma_t^2(\theta_2^*)} + 1 \right) \frac{\left\{ \sum_{j=1}^{t} (\beta_2^*)^{j-1} \epsilon_{t-j}^2 \right\} \left\{ \sum_{j=1}^{t} (\beta_2^*)^{j-1} \right\}}{\sigma_t^4(\theta_2^*)}$$
$$\leq \frac{K}{\sqrt{n}} \sum_{t=1}^{n} \left(\frac{2h_t\eta_t^2}{\sigma_t^2(\theta_2^*)} + 1 \right) \frac{h_t}{\sigma_t^2(\theta_2^*)} \frac{1}{h_t}.$$

Hence, by Lemma A.3,

$$E|\mathcal{J}_n(2,1)\sqrt{n}(\hat{\omega}_n - \omega_0)| \leq \frac{K}{\sqrt{n}} \sum_{t=1}^{n} E \frac{1}{h_t}.$$

The same bound is obtained when $\mathcal{J}_n(2,1)$ is replaced by $\mathcal{J}_n(3,1)$. Moreover,

$$h_t = \omega_0(1 + Z_{t-1} + Z_{t-1}Z_{t-2} + \cdots + Z_{t-1}\cdots Z_1) + Z_{t-1}\cdots Z_0\sigma_0^2.$$

By Assumption A, it follows that, for $i = 2, 3$,

$$\text{(A.26)} \quad E|\mathcal{J}_n(i,1)\sqrt{n}(\hat{\omega}_n - \omega_0)| \to 0 \quad \text{as} \quad n \to \infty.$$

Finally, using Theorem 2.1(iii), (A.15), and (A.16), the a.s. convergence (A.24) can be replaced by the same convergence in probability. The conclusion follows as in the case $\gamma_0 > 0$. *Q.E.D.*

PROOF OF THEOREM 2.3: The convergence results in (i) can be shown in a standard way, using Taylor expansions of the functions $\hat{\kappa}_\eta = \kappa_\eta(\hat{\theta}_n)$ and

$$\frac{1}{n} \sum_{t=1}^{n} \frac{1}{\sigma_t^2(\hat{\theta}_n)} \frac{\partial \sigma_t^2}{\partial \theta_i} \frac{\partial \sigma_t^2}{\partial \theta_j}(\hat{\theta}_n)$$

around θ_0, and the ergodic theorem together with the consistency of $\hat{\theta}_n$.

Now consider the case (ii). For some $\theta^* = (\omega^*, \alpha^*, \beta^*)'$ between $\hat{\theta}_n$ and θ_0, we have

$$\text{(A.27)} \quad \hat{\kappa}_\eta = \frac{1}{n}\sum_{t=1}^{n}\eta_t^4 - \frac{2}{n}\sum_{t=1}^{n}\frac{\epsilon_t^4}{\sigma_t^4(\theta^*)}\frac{1}{\sigma_t^2(\theta^*)}\frac{\partial\sigma_t^2(\theta^*)}{\partial\theta'}(\hat{\theta}_n - \theta_0)$$

$$:= \frac{1}{n}\sum_{t=1}^{n}\eta_t^4 + R_n.$$

By Proposition A.1 and already given arguments, for some $\rho \in (0,1)$,

$$|R_n| \le \frac{K}{n}\sum_{t=1}^{n}\eta_t^4\left(\frac{h_t}{\sigma_t^2(\theta_2^*)}\right)^2 \times (\rho^t|\hat{\omega}_n - \omega_0| + d_t^\alpha(\alpha^*, \beta^*)|\hat{\alpha}_n - \alpha_0| + d_t^\beta(\alpha^*, \beta^*)|\hat{\beta}_n - \beta_0|) = o(1) \quad \text{a.s.},$$

where the last equality follows from the strong consistency of $\hat{\alpha}_n$ and $\hat{\beta}_n$, the fact that $|\hat{\omega}_n - \omega_0|$ is bounded by compactness of Θ, and the existence moments at any order for $d_t^\alpha(\alpha^*, \beta^*)$, $d_t^\beta(\alpha^*, \beta^*)$, and $h_t/\sigma_t^2(\theta^*)$. Hence the first part of (ii) is proven. Now, similarly to (A.25), we have

$$\text{(A.28)} \quad n\widehat{\mathcal{J}}_{\omega,\alpha} \le \sum_{t=1}^{n}\frac{1}{\hat{\alpha}_n}\frac{h_t}{\sigma_t^2(\hat{\theta}_n)}\frac{\sum_{j=1}^{t}\hat{\beta}_n^{j-1}}{h_t} \le K\sum_{t=1}^{n}\frac{h_t}{\sigma_t^2(\hat{\theta}_n)}\rho^t$$

for $\rho \in (0,1)$ when n is large enough, by Proposition A.1(i). It follows that $n\widehat{\mathcal{J}}_{\omega,\alpha} = O(1)$ a.s. Similarly $n\widehat{\mathcal{J}}_{\omega,\beta} = O(1)$ a.s. Moreover, we have $n\widehat{\mathcal{J}}_{\omega,\omega} \ge 1/\sigma_1^4(\hat{\theta}_n) > 0$. Thus we have shown that

$$\widehat{\mathcal{J}}_{\alpha\beta,\omega}\widehat{\mathcal{J}}_{\omega,\omega}^{-1}\widehat{\mathcal{J}}_{\omega,\alpha\beta} = o(1) \quad \text{a.s.}$$

Now we turn to $\widehat{\mathcal{J}}_{\alpha\beta,\alpha\beta}$. Considering the top left term, a Taylor expansion around θ_0 gives

$$\text{(A.29)} \quad \widehat{\mathcal{J}}_{\alpha,\alpha} = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{1}{\sigma_t^2(\hat{\theta}_n)}\sum_{j=1}^{t}\hat{\beta}_n^{j-1}\epsilon_{t-j}^2\right)^2$$

$$= \frac{1}{n}\sum_{t=1}^{n}\{d_t^\alpha(\alpha_0, \beta_0)\}^2 + \frac{1}{n}\sum_{t=1}^{n}\left[\left\{\frac{1}{\sigma_t^2(\theta)}\frac{\partial\sigma_t^2}{\partial\alpha}(\theta_0)\right\}^2 - \{d_t^\alpha(\alpha_0, \beta_0)\}^2\right] + S_n,$$

where, for θ^* such that $\|\theta_0 - \theta^*\| \le \|\theta_0 - \hat{\theta}_n)\|$, $|S_n|$ is bounded by

$$\begin{aligned}
&\frac{K}{n}\sum_{t=1}^{n}\left(\frac{\sum_{j=1}^{t}(\beta^*)^{j-1}\epsilon_{t-j}^2}{\sigma_t^2(\theta^*)}\right)^2 \\
&\qquad \times\{\rho^t|\hat{\omega}_n-\omega_0| + d_t^{\alpha}(\alpha^*,\beta^*)|\hat{\alpha}_n-\alpha_0| + d_t^{\beta}(\alpha^*,\beta^*)|\hat{\beta}_n-\beta_0|\} \\
&\qquad + \frac{K}{n}\sum_{t=1}^{n}\left(\frac{\sum_{j=1}^{t}(\beta^*)^{j-1}\epsilon_{t-j}^2}{\sigma_t^2(\theta^*)}\right)\left(\frac{\sum_{j=1}^{t}(j-1)(\beta^*)^{j-2}\epsilon_{t-j}^2}{\sigma_t^2(\theta^*)}\right)|\hat{\beta}_n-\beta_0| \\
&\quad = o(1) \quad \text{a.s.}
\end{aligned}$$

by already used arguments. Moreover, the second term in the right-hand side of (A.29) converges to 0 a.s. by (A.10), while the first term is equal to $1/\alpha_0^2 = \mathcal{I}(1,1)$ since $d_t^{\alpha}(\alpha_0,\beta_0) = v_t(\alpha_0,\beta_0)/\alpha_0 = 1/\alpha_0$ by Lemma A.2. We thus have shown that $\widehat{\mathcal{J}}_{\alpha,\alpha}$ a.s. converges to $\mathcal{I}(1,1)$. The other two terms in $\widehat{\mathcal{J}}_{\alpha\beta,\alpha\beta}$ can be handled similarly, which completes the proof of (ii).

Turning to (iii), we note that $\partial\sigma_t^2(\theta^*)/\partial\omega \le K$ for n large enough, since $\beta_0 < 1$. Moreover, $\sigma_t^2(\theta^*) \ge \omega^* + \alpha^*\beta^*\epsilon_{t-2}^2$. Therefore, (A.27) continues to hold with $|R_n|$ bounded by

$$\begin{aligned}
&\frac{K}{n}\sum_{t=1}^{n}\eta_t^4\left(\frac{h_t}{\sigma_t^2(\theta_2^*)}\right)^2\frac{1}{\omega^*+\alpha^*\beta^*\epsilon_{t-2}^2} \\
&\quad + \frac{K}{n}\sum_{t=1}^{n}\eta_t^4\left(\frac{h_t}{\sigma_t^2(\theta_2^*)}\right)^2 \\
&\quad \times (d_t^{\alpha}(\alpha^*,\beta^*)|\hat{\alpha}_n-\alpha_0| + d_t^{\beta}(\alpha^*,\beta^*)|\hat{\beta}_n-\beta_0|).
\end{aligned}$$

Therefore, $|R_n| = o_P(1)$ by Proposition A.1(iii), the weak consistency of $\hat{\alpha}_n$ and $\hat{\beta}_n$, and the existence of moments for $d_t^{\alpha}(\alpha^*,\beta^*)$, $d_t^{\beta}(\alpha^*,\beta^*)$, and $h_t/\sigma_t^2(\theta_2^*)$. Hence $\hat{\kappa}_\eta \to \kappa$ in probability. Now, in view of the first inequality of (A.28), we have $n\widehat{\mathcal{J}}_{\omega,\alpha} \le K\sum_{t=1}^{n}\frac{h_t}{\sigma_t^2(\hat{\theta}_n)}\frac{1}{h_t}$, and thus $n\widehat{\mathcal{J}}_{\omega,\alpha} = O_P(1)$ by Lemma A.3 and the arguments used to show (A.26). Proceeding as in the proof of (i), we deduce that $\widehat{\mathcal{J}}_{\alpha\beta,\omega}\widehat{\mathcal{J}}_{\omega,\omega}^{-1}\widehat{\mathcal{J}}_{\omega,\alpha\beta} = o_P(1)$. By similar arguments, the right-hand side of (A.29) converges to $\mathcal{I}(1,1)$ in probability. *Q.E.D.*

PROOF OF PROPOSITION 2.1: If a consistent estimator $\hat{\theta}_n$ existed, then the test of critical region $C = \{\|\hat{\theta}_n - \theta_1\| > \|\hat{\theta}_n - \theta_1^*\|\}$ would have null asymptotic errors of the first and second kind, in contradiction with Proposition 3.1. *Q.E.D.*

A.3. *Constrained QMLE of* (α_0, β_0)

PROOF OF (2.8): By the arguments used to prove Theorem 2.1(iii), we have

$$(\hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega)) \to (\alpha_0, \beta_0) \quad \text{in probability} \quad \text{as} \quad n \to \infty. \tag{A.30}$$

A Taylor expansion of the criterion derivative gives

$$\begin{aligned} 0 &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial(\alpha, \beta)} \ell_t(\omega, \hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega)) \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial(\alpha, \beta)} \ell_t(\omega, \alpha_0, \beta_0) \\ &\quad + \left(\frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial(\alpha, \beta)' \, \partial(\alpha, \beta)} \ell_t(\omega, \alpha^*, \beta^*) \right) \\ &\quad \times \sqrt{n}(\hat{\alpha}_n^c(\omega) - \alpha_0, \hat{\beta}_n^c(\omega) - \beta_0)', \end{aligned} \tag{A.31}$$

where (α^*, β^*) is between $(\hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega))$ and (α_0, β_0). Another Taylor expansion yields, for $(\alpha^{**}, \beta^{**})$ between $(\hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega))$ and (α_0, β_0),

$$\begin{aligned} &\left\| \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial(\alpha, \beta)' \, \partial(\alpha, \beta)} \ell_t(\omega, \alpha_0, \beta_0) \right. \\ &\qquad \left. - \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial(\alpha, \beta)' \, \partial(\alpha, \beta)} \ell_t(\omega, \alpha^*, \beta^*) \right\| \\ &\quad \le |\alpha^* - \alpha_0| \frac{1}{n} \sum_{t=1}^{n} \left\| \frac{\partial^3}{\partial\alpha \, \partial(\alpha, \beta)' \, \partial(\alpha, \beta)} \ell_t(\omega, \alpha^{**}, \beta^{**}) \right\| \\ &\qquad + |\beta^* - \beta_0| \frac{1}{n} \sum_{t=1}^{n} \left\| \frac{\partial^3}{\partial\beta \, \partial(\alpha, \beta)' \, \partial(\alpha, \beta)} \ell_t(\omega, \alpha^{**}, \beta^{**}) \right\| \\ &\quad \le |\alpha^* - \alpha_0| \frac{1}{n} \sum_{t=1}^{n} \sup_{\theta \in \Theta} \left\| \frac{\partial^3}{\partial\alpha \, \partial(\alpha, \beta)' \, \partial(\alpha, \beta)} \ell_t(\omega, \alpha, \beta) \right\| \\ &\qquad + |\beta^* - \beta_0| \frac{1}{n} \sum_{t=1}^{n} \sup_{\theta \in \Theta} \left\| \frac{\partial^3}{\partial\beta \, \partial(\alpha, \beta)' \, \partial(\alpha, \beta)} \ell_t(\omega, \alpha, \beta) \right\| \\ &\quad = o_P(1), \end{aligned}$$

using (A.30) and (A.16). Therefore, using (A.15), the term in parentheses in (A.32) converges to $\mathcal{I}$. To conclude, it remains to prove that

$$\text{(A.32)} \qquad \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial}{\partial(\alpha,\beta)'}\ell_t(\omega,\alpha_0,\beta_0)\xrightarrow{d}\mathcal{N}(0,(\kappa_\eta-1)\mathcal{I}).$$

We have

$$\begin{aligned}\frac{1}{\sqrt{n}}&\sum_{t=1}^{n}\frac{\partial}{\partial(\alpha,\beta)'}\ell_t(\omega,\alpha_0,\beta_0)\\ &=\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial}{\partial(\alpha,\beta)'}\ell_t(\omega_0,\alpha_0,\beta_0)\\ &\quad+\frac{\omega-\omega_0}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial^2}{\partial\omega\,\partial(\alpha,\beta)'}\ell_t(\omega^*,\alpha_0,\beta_0),\end{aligned}$$

where ω^* is between ω_0 and ω. The last term tends to zero in probability, using Assumption A, similarly to (A.26). The first term converges in distribution to the normal law of (A.32) by Lemma A.4. *Q.E.D.*

PROOF OF PROPOSITION 2.2: The ergodic theorem entails that, for $\beta < 1$, almost surely

$$\begin{aligned}L_n(\alpha,\beta) &:= \frac{1}{n}\sum_{t=1}^{n}\frac{\epsilon_t^2}{\sigma_t^2(\omega,\alpha,\beta)}+\log\sigma_t^2(\omega,\alpha,\beta)\\ &\to \quad L(\alpha,\beta)=E_\infty\left\{\frac{h_t}{\sigma_t^2(\omega,\alpha,\beta)}+\log\sigma_t^2(\omega,\alpha,\beta)\right\}\end{aligned}$$

as $n\to\infty$. The dominated convergence theorem implies that

$$\begin{aligned}\frac{\partial L}{\partial\alpha}(\alpha_0,\beta_0)&=E_\infty\left\{\left(1-\frac{h_t}{\sigma_t^2(\omega,\alpha_0,\beta_0)}\right)\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\alpha}(\omega,\alpha_0,\beta_0)\right\}\\ &=(\omega-\omega_0)\left(\sum_{i\geq 0}\beta_0^i\right)E_\infty\left(\frac{1}{\sigma_t^4}\frac{\partial\sigma_t^2}{\partial\alpha}(\omega,\alpha_0,\beta_0)\right)\neq 0.\end{aligned}$$

It follows that the minimum of the function $L(\alpha,\beta)$ is reached at $(\alpha^*,\beta^*)\neq(\alpha_0,\beta_0)$.

A Taylor expansion of $L_n(\cdot)$ yields

$$\begin{aligned}\text{(A.33)} \qquad L_n\{\hat{\alpha}_n^c(\omega),\hat{\alpha}_n^c(\omega)\}&=L_n(\alpha_0,\beta_0)+\frac{\partial L_n}{\partial\alpha}(\tilde{\alpha}_n,\tilde{\beta}_n)\{\hat{\alpha}_n^c(\omega)-\alpha_0\}\\ &\quad+\frac{\partial L_n}{\partial\beta}(\tilde{\alpha}_n,\tilde{\beta}_n)\{\hat{\beta}_n^c(\omega)-\beta_0\},\end{aligned}$$

where $(\tilde{\alpha}_n, \tilde{\beta}_n)$ is between $(\hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega))$ and (α_0, β_0). Note that since $E\epsilon_t^4 < \infty$, almost surely, for $\overline{\beta} < 1$,

$$\limsup_{n\to\infty} \sup_{\alpha} \sup_{\beta<\overline{\beta}} \left| \frac{\partial L_n}{\partial \alpha}(\alpha, \beta_0) \right| \leq \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \left(1 + \frac{\epsilon_t^2}{\omega}\right) \frac{\sum_{i\geq 0} \overline{\beta}^i \epsilon_{t-1-i}^2}{\omega} < \infty,$$

$$\limsup_{n\to\infty} \sup_{\alpha} \sup_{\beta<\overline{\beta}} \left| \frac{\partial L_n}{\partial \beta}(\alpha, \beta_0) \right| \leq \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \left(1 + \frac{\epsilon_t^2}{\omega}\right) \frac{\sum_{i\geq 0} \overline{\beta}^i \sigma_{t-1-i}^2}{\omega} < \infty.$$

Now suppose that

$$(\hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega)) \to (\alpha_0, \beta_0) \quad \text{in probability} \quad \text{as} \quad n\to\infty. \tag{A.34}$$

Then it follows from (A.33) that

$$L_n\{\hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega)\} \to L(\alpha_0, \beta_0) \quad \text{in probability} \quad \text{as} \quad n\to\infty.$$

Then taking the limit in probability in the inequality

$$L_n\{\hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega)\} \leq L_n(\alpha^*, \beta^*),$$

we find that $L(\alpha_0, \beta_0) \leq L(\alpha^*, \beta^*)$, which is in contradiction with the definition of $(\alpha^*, \beta^*) \neq (\alpha_0, \beta_0)$. Thus (A.34) cannot be true. *Q.E.D.*

A.4. *Stationarity Test*

PROOF OF THEOREM 3.1: Let $\gamma_n(\theta) = n^{-1}\sum_{t=1}^{n} \log\{\alpha\eta_t^2(\theta) + \beta\}$ and $\eta_t(\theta) = \epsilon_t/\sigma_t(\theta)$. First consider the case $\gamma_0 < 0$. A Taylor expansion gives

$$\hat{\gamma}_n = \gamma_n(\theta_0) + \frac{\partial \gamma_n(\theta_0)}{\partial \theta'}(\hat{\theta}_n - \theta_0) + o_P\left(n^{-1/2}\right) \tag{A.35}$$

with

$$\frac{\partial \gamma_n(\theta_0)}{\partial \theta} = \frac{-1}{n} \sum_{t=1}^{n} \frac{1}{a_0(\eta_t)} \left\{ \alpha_0 \eta_t^2 \frac{1}{h_t} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} - \begin{pmatrix} 0 \\ \eta_t^2 \\ 1 \end{pmatrix} \right\} = -\Psi + o_P(1), \tag{A.36}$$

where, in view of $\alpha_0\eta_1^2/a_0(\eta_1) = 1 - \beta_0/a_0(\eta_1)$,

$$\Psi = (1-\nu_1)\Omega - a, \quad \Omega = E_\infty \frac{1}{h_t}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta}.$$

Moreover, the QMLE satisfies

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = -\mathcal{J}^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(1-\eta_t^2)\frac{1}{h_t}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta} + o_P(1). \tag{A.37}$$

In view of (A.35), (A.36), and (A.37), we have

$$\begin{aligned}&\sqrt{n}(\hat{\gamma}_n - \gamma_0) \\ &\quad = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} u_t + \Psi'\mathcal{J}^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(1-\eta_t^2)\frac{1}{h_t}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta} + o_P(1).\end{aligned} \tag{A.38}$$

Note that

$$\mathrm{Cov}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} u_t, \Psi'\mathcal{J}^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(1-\eta_t^2)\frac{1}{h_t}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta}\right) = c\Omega'\mathcal{J}^{-1}\Psi,$$

where $c = \mathrm{Cov}(u_t, 1-\eta_t^2)$. The Slutsky lemma and the central limit theorem for martingale differences thus entail

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \xrightarrow{d} \mathcal{N}(0, \sigma_u^2 + 2c\Omega'\mathcal{J}^{-1}\Psi + (\kappa_\eta - 1)\Psi'\mathcal{J}^{-1}\Psi).$$

Now let $\overline{\theta}_0 = (\omega_0, \alpha_0, 0)'$. Noting that $\overline{\theta}_0'\partial\sigma_t^2(\theta_0)/\partial\theta = h_t$ almost surely, we have

$$E\left\{\frac{1}{h_t}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta}\left(1 - \frac{1}{h_t}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta'}\overline{\theta}_0\right)\right\} = 0,$$

which entails $\mathcal{J}\overline{\theta}_0 = \Omega$ and $\Omega'\mathcal{J}^{-1}\Omega = 1$. It follows that $\Omega'\mathcal{J}^{-1}\Psi = 0$. Noting that $\Psi'\mathcal{J}^{-1}\Psi = a'\mathcal{J}^{-1}a - (1-\nu_1)^2$, the asymptotic distribution in (3.5) follows in the case $\gamma_0 < 0$.

Now consider the case $\gamma_0 \geq 0$. Let θ_n^* be a sequence such that $\|\theta_n^* - \theta_0\| \leq \|\hat{\theta}_n - \theta_0\|$. By Proposition A.1 (using Assumption A when $\gamma_0 = 0$), we have

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{1}{\sigma_t^2(\theta_n^*)}\frac{\partial\sigma_t^2(\theta_n^*)}{\partial\omega} \\ &\quad = o(1) \quad \text{a.s.} \quad \text{(resp., in probability)} \quad \text{as} \quad n \to \infty\end{aligned}$$

when $\gamma_0 > 0$ (resp., when $\gamma_0 = 0$). It can be deduced that under the same conditions, $\sqrt{n}\frac{\partial^2 \gamma_n(\theta_n^*)}{\partial\omega\,\partial\theta} = o(1)$ and

$$\sqrt{n}(\hat{\theta} - \theta_0)' \frac{\partial^2 \gamma_n(\theta_n^*)}{\partial\theta\,\partial\theta'} (\hat{\theta} - \theta_0) = o(1),$$

which entails that (A.35) still holds. By the same arguments, (A.36) holds with

$$\Omega = E \begin{pmatrix} 0 \\ d_t^{\alpha}(\theta_0) \\ d_t^{\beta}(\theta_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 1/\alpha_0 \\ \nu_1/\{\beta_0(1-\nu_1)\} \end{pmatrix} \quad \text{and} \quad \Psi = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The conclusion follows. *Q.E.D.*

The proof of Corollary 3.2 relies on arguments already used and is omitted.

PROOF OF COROLLARY 3.3: By arguments used in the proof of Theorem 2.3, $\hat{\sigma}_u^2$ converges almost surely to σ_u^2 when $\gamma < 0$ or $\gamma \geq 0$. Therefore, $T_n = \sqrt{n}(\hat{\gamma}_n - \gamma_0)/\hat{\sigma}_u + \sqrt{n}\gamma_0/\hat{\sigma}_u$ converges in probability to $-\infty$ when $\gamma < 0$, to $+\infty$ when $\gamma > 0$, and in distribution to $\mathcal{N}(0,1)$ when $\gamma_0 = 0$. *Q.E.D.*

A.5. *Inconsistency of Tests for ω*

PROOF OF PROPOSITION 3.1: The most powerful test is the Neyman–Pearson test of rejection region $\mathrm{C} = \{\mathcal{S}_n > c_n\}$, where

$$\mathcal{S}_n = \sum_{t=1}^{n} \frac{\epsilon_t^2}{\sigma_t^2(\theta_1)} + \log\sigma_t^2(\theta_1) - \frac{\epsilon_t^2}{\sigma_t^2(\theta_1^*)} - \log\sigma_t^2(\theta_1^*)$$

and c_n is a positive constant corresponding to the $\underline{\alpha}$-quantile of the (continuous) distribution of $\mathcal{S}_n$ under H_0 (see, e.g., Lehmann and Romano, (2005, Theorem 3.2.1)). By Proposition A.1(i), noting that $\sigma_t^2(\theta_1^*) - \sigma_t^2(\theta_1) = \sum_{j=1}^{t} \beta_1^{j-1}(\omega_1^* - \omega_1)$, there exists $\rho \in (0,1)$ such that

$$\left|\frac{\sigma_t^2(\theta_1^*) - \sigma_t^2(\theta_1)}{\sigma_t^2(\theta_1^*)}\right| \leq K\rho^t \quad \text{and} \quad \left|\frac{\sigma_t^2(\theta_1^*) - \sigma_t^2(\theta_1)}{\sigma_t^2(\theta_1)}\right| \leq K\rho^t \tag{A.39}$$

under both H_0 and H_1. Therefore, for some measurable function $\varphi(\cdot)$, as $n \to \infty$,

$$\mathcal{S}_n \to S_0 = \varphi(\eta_1^2, \eta_2^2, \ldots; \theta_1, \theta_1^*) \quad \text{a.s.} \quad \text{under } H_0$$

and

$$\mathcal{S}_n \to S_1 = -\varphi(\eta_1^2, \eta_2^2, \ldots; \theta_1^*, \theta_1) \quad \text{a.s.} \quad \text{under } H_1.$$

More precisely, in the ARCH(1) case, we have

$$S_0 = \sum_{t=1}^{\infty} \eta_t^2 s_{t-1} + \log(1 - s_{t-1}), \quad S_1 = \sum_{t=1}^{\infty} \eta_t^2 s_{t-1}^* - \log(1 + s_{t-1}^*), \tag{A.40}$$

where

$$s_{t-1} = \frac{\omega_1^* - \omega_1}{\omega_1^* + \omega_1 \sum_{k=1}^{t-1} \alpha_1^k \eta_{t-1}^2 \cdots \eta_{t-k}^2},$$

$$s_{t-1}^* = \frac{\omega_1^* - \omega_1}{\omega_1 + \omega_1^* \sum_{k=1}^{t-1} \alpha_1^k \eta_{t-1}^2 \cdots \eta_{t-k}^2}.$$

In the GARCH(1, 1) case, (A.40) still holds with

$$s_{t-1} = \frac{\sum_{j=1}^{t} \beta_1^{j-1}(\omega_1^* - \omega_1)}{\sigma_t^2(\theta_1^*)}, \quad s_{t-1}^* = \frac{\sum_{j=1}^{t} \beta_1^{j-1}(\omega_1^* - \omega_1)}{\sigma_t^2(\theta_1)}.$$

Using (A.39), it follows that

$$|S_0 - S_1| \le K|\omega_1 - \omega_1^*| \sum_{t=t_0}^{\infty} \rho^t(\eta_t^2 + 1). \tag{A.41}$$

Since the laws of S_0 and S_1 are continuous when $\omega_1 \neq \omega_1^*$, the power of the Neyman–Pearson test tends to

$$\lim_{n\to\infty} P_{H_1}(\mathcal{S}_n > c_n) = P(S_1 > c), \quad \text{where } c \text{ is such that} \quad P(S_0 > c) = \underline{\alpha}.$$

For any $\varepsilon > 0$, we have

$$P(S_1 > c) \le P(S_0 + |S_1 - S_0| > c) \le P(S_0 > c - \varepsilon) + P(|S_1 - S_0| > \varepsilon).$$

In the right-hand side of the last inequality, by continuity, the first probability is close to $\underline{\alpha}$ when ε is close to zero, and in view of (A.41), the second probability is close to zero when $|\omega_1 - \omega_1^*|$ is small. It follows that when $|\omega_1 - \omega_1^*|$ is small, $P(S_1 > c) < 1$, which shows the inconsistency of the Neyman–Pearson test and thus that of any test. *Q.E.D.*

A.6. *Stationarity Test in Nonlinear GARCH Models*

PROOF OF PROPOSITION 4.1: We start by considering the case $\Gamma > 0$. By the arguments given in the proof of Proposition A.1(i), $h_t \to \infty$ and $\epsilon_t^2 \to \infty$ a.s. at an exponential rate as $t \to \infty$. Moreover, it can be seen that the analog of Lemma A.1(i) still holds. Indeed, for all θ such as $\beta < e^{\Gamma}$, when $t \to \infty$, then $|\sigma_t^2(\theta)/h_t - w_t(\alpha, \beta)| \to 0$ where, similar to (A.5), $(w_t(\alpha, \beta))$ is defined as the stationary solution of

$$b_0(\eta_{t-1})w_t(\alpha, \beta) = \beta w_{t-1}(\alpha, \beta) + \alpha\eta_{t-1}^2. \tag{A.42}$$

Moreover, for all θ such as $\beta \geq e^{\Gamma}$, $\sigma_t^2(\theta)/h_t \to \infty$ a.s. as $t \to \infty$.

The pseudo-true value is thus the solution of

$$(\alpha^*, \beta^*)' = \arg \min_{(\alpha,\beta)\in\Theta_{\alpha,\beta}} E\{w_t^{-1}(\alpha, \beta) - 1 + \log w_t(\alpha, \beta)\}.$$

The consistency of $(\hat{\alpha}_n, \hat{\beta}_n)$ to (α^*, β^*) can then be shown by the arguments used for the proof of Theorem 2.1(ii).

It follows that

$$\hat{\gamma}_n \overset{o(1)}{=} \frac{1}{n}\sum_{t=1}^{n} \log\left(\hat{\alpha}_n \frac{\eta_t^2}{w_t(\hat{\alpha}_n, \hat{\beta}_n)} + \hat{\beta}_n\right) = \frac{1}{n}\sum_{t=1}^{n} \log \frac{a_0(\eta_t)w_{t+1}(\hat{\alpha}_n, \hat{\beta}_n)}{w_t(\hat{\alpha}_n, \hat{\beta}_n)}$$
$$\to \Gamma \quad \text{a.s.} \quad \text{as} \quad n \to \infty.$$

For the previous convergence, we note that for some neighborhood V^* of (α^*, β^*), we have $E \sup_{(\alpha,\beta)\in V^*} |\log w_1(\alpha, \beta)| < \infty$, which entails $\sup_{(\alpha,\beta)\in V(\theta^*)} \log w_{n+1}(\alpha, \beta)/n \to 0$ a.s. The latter moment condition comes from the fact that the strict stationarity condition entails $\beta^* < e^{\Gamma}$, which entails the existence of a moment of order $s > 0$ for $w_{n+1}(\alpha^*, \beta^*)$ (by the arguments given in Lemma 2.3 of Berkes, Horváth, and Kokoszka (2003)).

The convergence of $\hat{\sigma}_u^2$ to

$$\sigma_*^2 = \operatorname{Var} \log\left\{\frac{a(\eta_0)v_1(\alpha^*, \beta^*)}{v_0(\alpha^*, \beta^*)}\right\}$$

is obtained by similar arguments, noting that $E|\log w_1(\alpha^*, \beta^*)|^2 < \infty$ because $Ew_1^s(\alpha^*, \beta^*) < \infty$ for some $s > 0$ and because $\log w_1(\alpha^*, \beta^*) \geq \alpha^* \log \eta_0^2 - \log b_0(\eta_0)$. By (A.42), the a.s. limit of $\hat{\sigma}_u^2$ also can be written as $\operatorname{Var}\log\{\alpha^*\eta_1^2/w_1(\alpha^*, \beta^*) + \beta^*\}$. This limit is positive, unless η_1^2 is measurable with respect to the sigma-field generated by $\{\eta - u, u < 1\}$, which is impossible since the assumption $E|\log \eta_1^2|^2 > 0$ entails that η_1^2 has a nondegenerate distribution.

Now consider the case $\Gamma < 0$. Using standard arguments,

$$\Gamma^* = E_\infty \log\left(\alpha^* \frac{\epsilon_t^2}{\sigma_t^2(\theta^*)} + \beta^*\right) = E_\infty \log \frac{\sigma_{t+1}^2(\theta^*) - \omega^*}{\sigma_t^2(\theta^*)} < 0.$$

The convergence of $\hat{\sigma}_u^2$ to a positive limit is obtained by arguments already used. *Q.E.D.*

REFERENCES

AUE, A., AND L. HORVÁTH (2011): "Quasi-Likelihood Estimation in Stationary and Nonstationary Autoregressive Models With Random Coefficients," *Statistica Sinica* (forthcoming). [824]

BERKES, I., L. HORVÁTH, AND P. KOKOSZKA (2003): "GARCH Processes: Structure and Estimation," *Bernoulli*, 9, 201–227. [822,848,859]

——— (2005): "Near-Integrated GARCH Sequences," *The Annals of Applied Probability*, 15, 890–913. [822]

BILLINGSLEY, P. (1995): *Probability and Measure*. New York: Wiley. [846]

BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327. [821]

BOUGEROL, P., AND N. PICARD (1992): "Stationarity of GARCH Processes and of Some Nonnegative Time Series," *Journal of Econometrics*, 52, 115–127. [822]

ENGLE, R. F. (1982): "Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of the United Kingdom Inflation," *Econometrica*, 50, 987–1007. [821]

FRANCQ, C., AND J. M. ZAKOÏAN (2004): "Maximum Likelihood Estimation of Pure GARCH and ARMA-GARCH Processes," *Bernoulli*, 10, 605–637. [822,848]

——— (2006): "Mixing Properties of a General Class of GARCH(1, 1) Models Without Moment Assumptions on the Observed Process," *Econometric Theory*, 22, 815–834. [831]

——— (2010): *GARCH Models: Structure, Statistical Inference and Financial Applications*. Chichester: Wiley. [828]

——— (2012): "Supplement to 'Strict Stationarity Testing and Estimation of Explosive and Stationary Generalized Autoregressive Conditional Heteroscedasticity Models'," *Econometrica Supplementary Material*, 80, http://www.econometricsociety.org/ecta/Supmat/9405_data and programs.zip. [824]

GLOSTEN, L. R., R. JAGANATHAN, AND D. RUNKLE (1993): "On the Relation Between the Expected Values and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48, 1779–1801. [831]

GOLDIE, C. M., AND R. A. MALLER (2000): "Stability of Perpetuities," *The Annals of Probability*, 28, 1195–1218. [823]

HÖRMANN, S. (2008): "Augmented GARCH Sequences: Dependence Structure and Asymptotics," *Bernoulli*, 14, 543–561. [831]

JENSEN, S. T., AND A. RAHBEK (2004a): "Asymptotic Normality of the QMLE Estimator of ARCH in the Nonstationary Case," *Econometrica*, 72, 641–646. [823]

——— (2004b): "Asymptotic Inference for Nonstationary GARCH," *Econometric Theory*, 20, 1203–1226. [823,827,828,841,846]

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, 601–622. [823,838,839]

LEHMANN, E. L., AND J. P. ROMANO (2005): *Testing Statistical Hypotheses* (Third Ed.). New York: Springer-Verlag. [857]

LINTON, O., J. PAN, AND H. WANG (2010): "Estimation for a Non-Stationary Semi-Strong GARCH(1, 1) Model With Heavy-Tailed Errors," *Econometric Theory*, 26, 1–28. [823]

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, 575–596. [821]

NELSON, D. B. (1990): "Stationarity and Persistence in the GARCH(1, 1) Model," *Econometric Theory*, 6, 318–334. [822,823]

VAN DER VAART, A. W. (1998): *Asymptotic Statistics*. Cambridge, U.K.: Cambridge University Press.

University Lille 3 (EQUIPPE), BP 60149, 59653 Villeneuve d'Ascq cedex, France and CREST; christian.francq@univ-lille3.fr

and

CREST, 15 Boulevard Gabriel Peri, 92245 Malakoff Cedex, France and University Lille 3 (EQUIPPE); zakoian@ensae.fr.

Manuscript received July, 2010; final revision received August, 2011.