Strict stationarity testing and estimation of explosive and stationary GARCH models

Christian Francq and Jean-Michel Zakoïan

September 1, 2011
STRICT STATIONARITY TESTING AND ESTIMATION OF EXPLOSIVE
AND STATIONARY GARCH MODELS\textsuperscript{1}

CHRISTIAN FRANCQ\textsuperscript{a} AND JEAN-MICHEL ZAKOIAN\textsuperscript{b}

This paper studies the asymptotic properties of the quasi-maximum likeli-
hood estimator of 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 non-stationary situations, but is consistently estimated, with the same
estimator, in both cases. Tests of strict stationarity and non stationarity are
proposed. The tests developed for the classical GARCH(1,1) model are able to
detect non-stationarity in more general GARCH models. A numerical illustration
based on stock indices and individual stock returns is proposed.

\textbf{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 financ-
cial time series. A standard assumption is that the prices are non stationary
while the returns (or log-returns) are stationary. Numerous econometric

\textsuperscript{1}We are most thankful to the Editor James H. Stock and to three referees for their con-
structive comments and suggestions. We would like to thank Youri A. Davydov, Siegfried
Hörmann, Lajos Horváth and several seminar and conference audiences for helpful dis-
cussions. 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).
\textsuperscript{a}University Lille 3 (EQUIPPE), BP 60149, 59653 Villeneuve d’Ascq cedex, France and
CREST.\texttt{christian.francq@univ-lille3.fr}
\textsuperscript{b}CREST, 15 Boulevard Gabriel Peri, 92245 Malakoff Cedex, France and University
Lille 3 (EQUIPPE).\texttt{zakoian@ensae.fr}
tools, such as the unit root tests, have been introduced for testing the non-
stationarity of prices. For the log-returns, the most widely used models are
arguably the GARCH introduced by Engle (1982) and Bollerslev (1986). No
econometric tools are available for testing strict stationarity in the GARCH
framework. This is the main aim of this paper to develop such tools. The
problem is non standard because, contrary to stationarity in linear time
series models, which solely depends on the lag polynomials, the strict sta-
tionarity condition for GARCH models has a non explicit form, involving
the distribution of the underlying independent and identically distributed
(iid) sequence.

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

\begin{equation}
\begin{cases}
\epsilon_t = \sqrt{h_t} \eta_t, & t = 1, 2, \ldots \\
h_t = \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta h_{t-1}
\end{cases}
\end{equation}
with initial values $\epsilon_0$ and $h_0 \geq 0$, where $\omega_0 > 0$, $\alpha_0, \beta_0 \geq 0$, and $(\eta_t)$ is a sequence of independent and identically distributed (iid) 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)

(1.2) $\gamma_0 < 0$.

More precisely, if (1.2) holds we have

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

where

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

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

(1.5) $\gamma_0 > 0$,

$h_t \to \infty$ 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

\begin{equation}
\text{when } \gamma_0 = 0, \quad h_t \to \infty \text{ in probability}
\end{equation}

instead of almost surely in the case $\gamma_0 > 0$. The astonishing difficulties encountered in the case $\gamma_0 = 0$ are related to the fact that the sequence

$h_t = h_{t,t} + a_0(\eta_{t-1}) \ldots a_0(\eta_0) h_0$ does not increase with $t$.

### 1.2. The econometric problem

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

\begin{equation}
\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),
\end{equation}

where $\Theta$ is a compact subset of $(0, \infty)^3$ containing the true value $\theta_0 = (\omega_0, \alpha_0, \beta_0)'$, and $\sigma_t^2(\theta) = \omega + \alpha_t \sigma_{t-1}^2(\theta)$ for $t = 1, \ldots, n$ (with initial values for $\epsilon_0^2$ 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 will establish the asymptotic distribution of the statistic

\begin{equation}
\hat{\gamma}_n = \frac{1}{n} \sum_{t=1}^{n} \log(\hat{\alpha}_n \hat{\eta}_t^2 + \hat{\beta}_n).
\end{equation}

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 (1990) noted that the arguments given by Nelson (1990) for the a.s. convergence are in failure when $\gamma_0 = 0$. 
(2004a, 2004b) were the first to establish an asymptotic theory for estimators of non-stationary GARCH.\footnote{See Linton, Pan and Wang (2010) for extensions in the case of non iid errors.} However, they only considered a constrained QMLE of \((\alpha_0, \beta_0)\) (in the sense that the value of the intercept is fixed) which is consistent in the non stationary case, but is inconsistent in the stationary case. 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 \((\alpha_0, \beta_0)\), the only components which 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

\[
\sqrt{n} \left( \hat{\alpha}_n - \alpha_0, \hat{\beta}_n - \beta_0 \right)^\prime \overset{d}{\to} N \left\{ 0, (\kappa_\eta - 1) I^{-1} \right\}, \quad \text{as } n \to \infty
\]

where \(\overset{d}{\to}\) stands for the convergence in distribution, \(\kappa_\eta = E\eta_1^4\) and \(I\) is a matrix which has an explicit form and does not depend on \(\omega_0\). When \(\gamma_0 = 0\), the QMLE of \((\alpha_0, \beta_0)\) will be 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 \((\alpha_0, \beta_0)\) is consistent in every situation, we will show that the QMLE of \(\omega_0\) is only consistent in the stationary case. For this reason it is important to test the sign of \(\gamma_0\).

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 \((\alpha_0, \beta_0)\) 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 non stationarity. We also consider testing stationarity in more general
GARCH-type models. Numerical illustrations are provided in Section 5.
In particular, the stationarity of eleven major stock returns is analyzed.
Section 6 concludes. Proofs and complementary results are collected in the
Appendix.

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 $\gamma_0 = 0$ will be 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. 4

**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 $\Theta$ such that $\forall \theta \in \Theta$, $\beta < 1$,

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

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

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

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

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

4For instance for the process $(\epsilon_t)$ we have $E_\infty(\epsilon_t^2) = \omega_0/(1-\alpha_0-\beta_0)$ when $\alpha_0 + \beta_0 < 1.$
iv) When $\gamma_0 < 0$, $\kappa_\eta = E\eta_1^2 \in (1, \infty)$, $\theta_0$ belongs to the interior $\Theta$ of $\Theta$, and for $\Theta$ such that $\varnothing \in \Theta$, $\beta < 1$,

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

and

\[
(2.2) \quad J = E \left( \frac{1}{h_t^2} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta'}(\theta_0) \right).
\]

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

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

where

\[
I = \begin{pmatrix}
\frac{1}{\alpha_0^2} & \frac{\alpha_0 \beta_0 (1-\nu_1)}{\alpha_0^2 \beta_0 (1-\nu_2)} \\
\frac{\alpha_0 \beta_0 (1-\nu_1)}{\alpha_0^2 \beta_0 (1-\nu_2)} & \frac{\beta_0 (1-\nu_2)}{\alpha_0^2 (1-\nu_1)(1-\nu_2)}
\end{pmatrix}
\]

with $\nu_i = E \left( \frac{\beta_0}{\alpha_0 \eta_1^2 + \beta_0} \right)^i$.

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 $\eta_1^2$. Let $Z_t = \alpha_0 \eta_1^2 + \beta_0$. 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, $E(1 + Z_{t-1} + Z_{t-2} + \cdots + Z_{t-1} \cdots Z_1) \geq t$. We introduce the assumption

A: when $t$ tends to infinity,

\[
E \left( \frac{1}{1 + Z_1 + Z_1 Z_2 + \cdots + Z_1 \cdots Z_{t-1}} \right) = o \left( \frac{1}{\sqrt{t}} \right).
\]

Note that 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 A is of order $(\log t)/\sqrt{t}$ for any distribution such that $E|\log Z_t|^2 < \infty$ (details are available from the authors). This assumption will be used to prove that

\[
(2.4) \quad \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{1}{h_t} \to 0, \quad \text{as } n \to \infty
\]
in $L^1$ when $\gamma_0 = 0$.\footnote{In the case $\gamma_0 > 0$, (2.4) holds a.s. from Proposition A.1-i) and $A$ is not needed.}

**Theorem 2.2** Suppose that the assumptions of Theorem 2.1-iii) hold, in particular $\gamma_0 = 0$. Then, if $\theta_0 \in \breve{\Theta}$, $\kappa_\eta \in (1, \infty)$, $E|\log \eta_n^2| < \infty$ and $A$ 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)-(2.2), when $\gamma_0 < 0$ the asymptotic distribution of the QMLE $(\hat{\alpha}_n, \hat{\beta}_n)$ of $(\alpha_0, \beta_0)$ is given by

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

with

$$I_* = J_{\alpha\beta,\alpha\beta} - J_{\alpha\beta,\omega} J_{\omega,\alpha\beta}^{-1} J_{\omega,\omega},$$

and $J_{\omega,\omega} = \mathbb{E}_{\infty} \left( \frac{\partial \sigma_n^2}{\partial \omega} \frac{\partial \sigma_n^2}{\partial \omega} (\theta_0) \right)$, $J_{\alpha\beta,\alpha\beta} = \mathbb{E}_{\infty} \left( \frac{\partial \sigma_n^2}{\partial (\alpha, \beta)} \frac{\partial \sigma_n^2}{\partial (\alpha, \beta)} (\theta_0) \right)$ and $J_{\omega,\alpha\beta} = J'_{\alpha\beta,\omega} = \mathbb{E}_{\infty} \left( \frac{\partial \sigma_n^2}{\partial (\alpha, \beta)} \frac{\partial \sigma_n^2}{\partial (\alpha, \beta)} (\theta_0) \right)$. Letting

$$\hat{J}_{\alpha\beta,\alpha\beta} = \frac{1}{n} \sum_{t=1}^n \frac{\partial \sigma_n^2}{\sigma_n^2(\theta_n)} \frac{\partial \sigma_n^2}{\sigma_n^2(\theta_n)} (\hat{\theta}_n),$$

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

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

is a strongly consistent estimator of $I_*$ 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 $\gamma_0 \geq 0$.\footnote{In the case $\gamma_0 > 0$, (2.4) holds a.s. from Proposition A.1-i) and $A$ is not needed.}
Theorem 2.3  Let the assumptions i)-ii)-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 \( \hat{I}_* \to I_* \) a.s as \( n \to \infty \).

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

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

In any case, \( (\hat{\kappa}_\eta - 1) \hat{I}_*^{-1} \) is a consistent estimator of the asymptotic variance of the QMLE of \((\alpha_0, \beta_0)\).

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 \( \omega_0 \). From the proof of Theorem 2.1, it can be shown that the log-likelihood is completely flat in the direction where \((\alpha_0, \beta_0)\) is fixed and \( \omega_0 \) varies. More precisely, we have

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

for any sequence \( \lambda_n \) tending to zero as \( n \to \infty \).

Thus, in general, as noted by Jensen and Rahbek (2004b), there is no consistent estimator of \( \omega_0 \). 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 \( \Theta \) 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 \((\alpha_0, \beta_0)\)

The asymptotic behaviour of the QMLE \((\hat{\alpha}_n, \hat{\beta}_n)\) being independent of \(\omega_0\) when \(\gamma_0 > 0\), and the QMLE of \(\omega_0\) being generally inconsistent in view of Proposition 2.1, it seems natural to avoid estimating \(\omega_0\). To this aim a constrained QMLE of \((\alpha_0, \beta_0)\), in which the first component of \(\theta\) is fixed to an arbitrary value \(\omega\), can be introduced. The estimator

\[
(\hat{\alpha}_n, \hat{\beta}_n) = \min_{(\alpha, \beta) \in \Theta} \frac{1}{n} \sum_{t=1}^{n} \ell_t(\omega, \alpha, \beta),
\]

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 \((\hat{\alpha}^c_n(\omega), \hat{\beta}^c_n(\omega))\). In Appendix A.3 we prove that: under the assumptions of Theorem 2.2, in particular \(\gamma_0 = 0\),

\[
\sqrt{n} \left( \hat{\alpha}^c_n(\omega) - \alpha_0, \hat{\beta}^c_n(\omega) - \beta_0 \right) \xrightarrow{d} \mathcal{N} \left\{ 0, (\kappa_\eta - 1)I^{-1} \right\}, \quad \text{as } n \to \infty.
\]

However, the next result shows that the restricted QMLE of \((\alpha_0, \beta_0)\) is generally inconsistent in the stationary case.

**Proposition 2.2** Let \((\epsilon_t)\) be a stationary solution of the GARCH(1,1) model with parameters \(\omega_0, \alpha_0\) and \(\beta_0\), such that \(E\epsilon_t^4 < \infty\). Then, if \(\omega \neq \omega_0\)

\[
(\hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega)) \text{ does not converge in probability to } (\alpha_0, \beta_0).
\]

On the contrary, Theorem 2.1 and Theorem 2.2 show that

\[
(\hat{\alpha}_n, \hat{\beta}_n) \text{ is always CAN (under A when } \gamma_0 = 0).\]

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 \((\alpha_0, \beta_0)\) 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

\[ H_0: a\alpha_0 + b\beta_0 \leq c \quad \text{against} \quad H_1: a\alpha_0 + b\beta_0 > c, \]

where \( a, b, c \) are given numbers. A case of particular interest is \( a = b = c = 1 \), because \( Ee_t^2 < \infty \) if and only if \( \alpha_0 + \beta_0 < 1 \). Note however that we do not impose any constraint on \( a, b, c \) so that some values of \( \theta_0 \) satisfying \( H_0 \) may correspond to nonstationary GARCH models. A direct consequence of Theorem 2.1-2.3 is the following result, in which \( \Phi \) denotes the \( \mathcal{N}(0,1) \) cumulative distribution function. Let \( \alpha \in (0,1) \).

**Corollary 3.1** Assume that \( \theta_0 \in \Theta \) and the assumptions of Theorem 2.3 hold. For the testing problem (3.1), the test defined by the critical region

\[ C^{\alpha*} = \left\{ T_n^{\alpha*} := \sqrt{n}(a\hat{\alpha}_n + b\hat{\beta}_n - c) \bigg/ \sqrt{(\hat{\kappa}_\eta - 1)(a,b)\hat{I}_n^{-1}(a,b)} > \Phi^{-1}(1 - \alpha) \right\} \]

has the asymptotic significance level \( \alpha \) and is consistent.

The following result, showing that no consistent test exists for \( \omega_0 \), 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 \( \Theta \) such that \( E\log(\alpha_1\eta_t^2 + \beta_1) > 0 \) and \( \omega_1 \neq \omega_1^* \). When \( \omega_1 \) and \( \omega_1^* \) are sufficiently close, there exists no consistent test for testing \( H_0 : \theta_0 = \theta_1 \) against \( H_1 : \theta_0 = \theta_1^* \) at the asymptotic level \( \alpha \in (0,1/2) \).
Propositions 2.1 and 3.1 show that no asymptotically valid inference on \( \omega_0 \) can be done in the nonstationary case. It is thus of interest to test if a given series is stationary or not.

### 3.2. Strict stationarity testing

Consider the strict stationarity testing problems

\begin{equation}
H_0 : \gamma_0 < 0 \quad \text{against} \quad H_1 : \gamma_0 \geq 0,
\end{equation}

and

\begin{equation}
H_0 : \gamma_0 \geq 0 \quad \text{against} \quad H_1 : \gamma_0 < 0,
\end{equation}

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

**Theorem 3.1** Assume the following conditions are satisfied: iv) of Theorem 2.1 in the case \( \gamma_0 < 0 \), v) of Theorem 2.1 in the case \( \gamma_0 > 0 \) and the those of Theorem 2.2 (in particular 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

\begin{equation}
\sqrt{n}(\hat{\gamma}_n - \gamma_0) \overset{d}{\rightarrow} N(0, \sigma_\gamma^2)
\end{equation}

as \( n \to \infty \)

where

\[
\sigma^2_\gamma = \begin{cases} 
\sigma_u^2 + (\kappa_1 - 1)(a'J^{-1}a - (1 - \nu_1)^2)) & \text{when } \gamma_0 < 0, \\
\sigma_u^2 & \text{when } \gamma_0 \geq 0,
\end{cases}
\]

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 accolades in the first expression of $\sigma_γ^2$ is positive, showing that the asymptotic variance of $\hat{γ}_n$ is larger in the stationary case than in the non-stationary case. The next result provides an estimator of $\sigma_γ^2$ which is consistent in every situation (explosive and stationary). It allows to construct a confidence interval for the top Lyapunov exponent. Let

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

\[
\hat{\sigma}_γ^2 = \hat{\sigma}_u^2 + (\hat{κ}_n - 1)\{\hat{α}'\hat{J}^{-1}\hat{α} - (1 - \hat{ν}_1)^2\},
\]

\[
\hat{α} = \left(0, \frac{1 - \hat{ν}_1}{\hat{α}_n}, \frac{\hat{ν}_1}{\hat{β}_n}\right), \quad \hat{ν}_1 = \frac{1}{n} \sum_{t=1}^{n} \frac{\hat{β}_n}{\hat{α}_n \hat{η}_t^2 + \hat{β}_n}.
\]

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

(3.6) \quad \hat{σ}_γ^2 \to σ_γ^2 in probability (and a.s. when γ₀ ≠ 0) as n → ∞.

Therefore, at the asymptotic level α ∈ (0, 1), a confidence interval for γ₀ is

\[
[\hat{γ}_n - \frac{\hat{σ}_γ}{\sqrt{n}} \Phi^{-1}(1 - α/2), \hat{γ}_n + \frac{\hat{σ}_γ}{\sqrt{n}} \Phi^{-1}(1 - α/2)].
\]

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 critical region

(3.7) \quad C^{ST} = \left\{ T_n := \sqrt{n} \frac{\hat{γ}_n}{\sigma_u} > \Phi^{-1}(1 - α) \right\}

has its asymptotic significance level bounded by α, has the asymptotic probability of rejection α under γ₀ = 0, and is consistent for all γ₀ > 0.

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

(3.8) \quad C^{NS} = \{ T_n < \Phi^{-1}(α) \}

has its asymptotic significance level bounded by α, has the asymptotic probability of rejection α under γ₀ = 0, and is consistent for all γ₀ < 0.
4. TESTING NON STATIONARITY IN NON LINEAR GARCH

In this section we study the behaviour of the stationarity tests of Section 3.2 when the data are generated by the following GARCH-type model:

\begin{align}
\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{align}

with an initial value $h_0$, under the same assumptions on $(\eta_t)$ as in Model (1.1). In this model, $\omega : \mathbb{R} \to \mathbb{R}$ and $b_0 : \mathbb{R} \to [b_1, +\infty)$, for some $\omega, \infty, b_0 > 0$. 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 $b_0(x) = \alpha_0 x^2 + \beta_0$; the GJR model when $b_0(x) = \alpha_1 (\max\{x, 0\})^2 + \alpha_2 (\min\{x, 0\})^2 + \beta_0$. It can be shown that, if $E \max\{0, \log b_0(\eta_t)\} < \infty$,

\begin{align}
\Gamma := E \log b_0(\eta_t) < 0
\end{align}

is a necessary and sufficient condition for the strict stationarity of this model (see e.g. Francq and Zakoïan, 2006a). Our aim is to test strict stationarity, without estimating the non parametric 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 $0 < E|\log \eta_1^2| < \infty$ and $E|\log b_0(\eta_1)|^2 < \infty$. 

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

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

If $\Gamma < 0$ then, under regularity conditions implying 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 $\text{Var} \log \epsilon_t^2 < \infty$, we have, for some $\Gamma^*$,

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

**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

$$P(C_{NS}) \to 0 \quad \text{and} \quad P(C_{ST}) \to 1$$

where $C_{ST}$ and $C_{ST}$ are defined in Corollary 3.3.

If $\Gamma < 0$ then

$$P(C_{ST}) \to 0 \quad \text{and} \quad P(C_{NS}) \to 1.$$

5. NUMERICAL ILLUSTRATIONS

Before illustrating our asymptotic results for the tests, we study the behaviour 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 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 \(\theta_0\) corresponding to a second-order stationary, a strictly stationary without second-order moment, and a non stationary process. Since \(\eta_t \sim \mathcal{N}(0,1)\), the QMLE corresponds to the MLE. Similar results were obtained with other distributions for \(\eta_t\). Concerning the estimation of \((\alpha_0, \beta_0)\), the results are very similar for the 3 values of \(\theta_0\), confirming that stationarity is not necessary for the estimation of these parameters. By contrast, the column \(\omega\) in the non stationary case confirms the asymptotic results of Proposition 2.1 and illustrates the impossibility of estimating the parameter \(\omega_0\) with a reasonable accuracy under the nonstationarity condition (1.5). Note that the RMSE of estimation of \(\omega_0\) even worsen 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 \(n = 500\), \(n = 2,000\) and \(n = 4,000\) of a GARCH(1,1) model of the form (1.1), with different values of \(\theta_0\) and the standardized Student distribution with 7 degrees of freedom for \(\eta_t\). The standardized Student is often employed as distribution of GARCH errors in applied works.
TABLE I

Bias (Mean Errors) and MSE (Mean Squared Errors) for the QMLE of a GARCH(1,1), with $\eta_t \sim N(0,1)$ and $\theta_0 = (1,0.3,0.6), \theta_0 = (1,0.5,0.6)$ or $\theta_0 = (1,0.7,0.6)$, corresponding to a second-order stationary (2nd), strict stationary (ST) or non stationary (NS) model. Bias and MSE are computed over 1,000 independent simulations of length $n = 200$ or $n = 4,000$.

<table>
<thead>
<tr>
<th></th>
<th>2nd ($\gamma_0 = -0.180$)</th>
<th>ST ($\gamma_0 = -0.038$)</th>
<th>NS ($\gamma_0 = 0.078$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\omega$</td>
<td>$\alpha$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>$n = 200$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias</td>
<td>-0.21</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>MSE</td>
<td>0.58</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>$n = 4,000$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>MSE</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

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 $b = 1$, i.e. a test on the value of $\beta_0$, are presented in Tables II-III. It has to be noted that the test (3.2) behaves similarly when the tested value corresponds to a stationary solution (Table II) or to a non stationary 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 $\alpha_0$ 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-V show that, as expected, the frequency of rejection of the $C^{ST}$-test increases with $\gamma_0$ 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 first kind is
TABLE II

Relative frequency of rejection (in %) for the test (3.2) of the null hypothesis $H_0 : \beta_0 \leq 0.7$ against $H_1 : \beta_0 > 0.7$ (i.e. $a = 0$, $b = 1$ and $c = 0.7$ in (3.1)) at the nominal level $\alpha = 5\%$, when $\alpha_0 = 0.2$. The value $(\alpha_0, \beta_0) = (0.2, 0.7)$ corresponds to a stationary process.

<table>
<thead>
<tr>
<th>$\beta_0$</th>
<th>0.61</th>
<th>0.64</th>
<th>0.67</th>
<th>0.70</th>
<th>0.73</th>
<th>0.76</th>
<th>0.79</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 500$</td>
<td>3.5</td>
<td>4.3</td>
<td>5.2</td>
<td>8.9</td>
<td>12.6</td>
<td>26.8</td>
<td>49.6</td>
</tr>
<tr>
<td>$n = 2,000$</td>
<td>0.3</td>
<td>0.6</td>
<td>1.8</td>
<td>6.8</td>
<td>18.3</td>
<td>53.1</td>
<td>91.5</td>
</tr>
<tr>
<td>$n = 4,000$</td>
<td>0.2</td>
<td>0.3</td>
<td>1.0</td>
<td>5.5</td>
<td>27.7</td>
<td>76.9</td>
<td>99.0</td>
</tr>
</tbody>
</table>

TABLE III

As Table II, but for testing the same hypotheses, but when $\alpha_0 = 0.5$. The value $(\alpha_0, \beta_0) = (0.5, 0.7)$ corresponds to a non stationary process.

<table>
<thead>
<tr>
<th>$\beta_0$</th>
<th>0.61</th>
<th>0.64</th>
<th>0.67</th>
<th>0.70</th>
<th>0.73</th>
<th>0.76</th>
<th>0.79</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 500$</td>
<td>0.3</td>
<td>0.5</td>
<td>2.8</td>
<td>9.9</td>
<td>25.5</td>
<td>47.7</td>
<td>67.2</td>
</tr>
<tr>
<td>$n = 2,000$</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>6.2</td>
<td>41.6</td>
<td>81.8</td>
<td>97.0</td>
</tr>
<tr>
<td>$n = 4,000$</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>6.1</td>
<td>61.0</td>
<td>96.2</td>
<td>99.7</td>
</tr>
</tbody>
</table>

better controlled for tests of stationarity of ARCH(1) models, which is not surprising, the model being 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 standard-

TABLE IV

Relative frequency of rejection of the stationarity hypothesis $H_0 : \gamma_0 < 0$ of the test (3.7) at the nominal level $\alpha = 5\%$ for the GARCH(1,1) model with $\beta_0 = 0.8$. The parameter $\alpha_0 = 0.2575$ corresponds to $\gamma_0 = 0$.

<table>
<thead>
<tr>
<th>$\alpha_0$</th>
<th>0.18</th>
<th>0.20</th>
<th>0.22</th>
<th>0.2575</th>
<th>0.28</th>
<th>0.30</th>
<th>0.31</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 500$</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>7.5</td>
<td>27.8</td>
<td>61.4</td>
<td>75.2</td>
</tr>
<tr>
<td>$n = 2,000$</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>6.3</td>
<td>67.8</td>
<td>98.6</td>
<td>99.9</td>
</tr>
<tr>
<td>$n = 4,000$</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>5.3</td>
<td>92.4</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>
TABLE V
As Table IV, but for testing the nonstationarity hypothesis $H_0 : \gamma_0 \geq 0$
with the test (3.8).

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\alpha_0$</th>
<th>0.18</th>
<th>0.20</th>
<th>0.22</th>
<th>0.2575</th>
<th>0.28</th>
<th>0.30</th>
<th>0.31</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>98.3</td>
<td>91.7</td>
<td>69.3</td>
<td>19.8</td>
<td>4.1</td>
<td>0.7</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>2,000</td>
<td>100.0</td>
<td>100.0</td>
<td>98.3</td>
<td>11.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>4,000</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>9.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
</tbody>
</table>

TABLE VI
As Table IV, but for a GJR model. The parameter $\alpha_1 = 0.2575$ corresponds to $\Gamma = 0$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\alpha_1$</th>
<th>0.18</th>
<th>0.20</th>
<th>0.22</th>
<th>0.2575</th>
<th>0.28</th>
<th>0.30</th>
<th>0.31</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>0.1</td>
<td>0.1</td>
<td>1.1</td>
<td>7.8</td>
<td>15.8</td>
<td>32.7</td>
<td>35.2</td>
<td></td>
</tr>
<tr>
<td>2,000</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>6.6</td>
<td>31.7</td>
<td>65.8</td>
<td>77.4</td>
<td></td>
</tr>
<tr>
<td>4,000</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>5.6</td>
<td>45.1</td>
<td>87.7</td>
<td>96.1</td>
<td></td>
</tr>
</tbody>
</table>

ized Student distribution for $\eta_t$, we take $\beta_0 = 0.8$ and $\alpha_2 = 0.2575$, whereas $\alpha_1$ varies in 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 $\alpha_1 \in \{0.28, 0.30, 0.31\}$.

Tables 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 non-standard GARCH models, there is no guarantee that the asymptotic relative frequency of rejection 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,
TABLE VII

Test statistic $T_n$ of the strict stationarity tests (3.7) and (3.8). 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 $\gamma_0 < 0$, tends to $+\infty$ when $\gamma_0 > 0$.

<table>
<thead>
<tr>
<th>CAC</th>
<th>DAX</th>
<th>DJA</th>
<th>DJI</th>
<th>DJT</th>
<th>DJU</th>
<th>FTSE</th>
<th>Nasdaq</th>
<th>Nikkei</th>
<th>SMI</th>
<th>SP500</th>
</tr>
</thead>
<tbody>
<tr>
<td>-14.5</td>
<td>-15.8</td>
<td>-15.1</td>
<td>-13</td>
<td>-15.1</td>
<td>-14</td>
<td>-10.7</td>
<td>-8.5</td>
<td>-15.4</td>
<td>-23</td>
<td>-11.1</td>
</tr>
</tbody>
</table>

DJT, DJU, FTSE, Nasdaq, Nikkei, SMI and SP500, from January 2, 1990, to January 22, 2009, except for the indices for which such historical data do not exist. Table 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 conclusion can be opposite 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).

For four of these five stocks, the non stationarity assumption cannot be rejected, at any reasonable significance level. Interestingly, non stationarity 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 non stationarity of the series. As an example, Figure 1 displays the sample path of the MCBF

---

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

TABLE VIII

Test statistic $T_n$ and $p-$values of the nonstationarity test \((3.8)\) for stock returns.

<table>
<thead>
<tr>
<th></th>
<th>ICGN</th>
<th>MCBF</th>
<th>KVA</th>
<th>BTC</th>
<th>CCME</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>928</td>
<td>868</td>
<td>1,221</td>
<td>908</td>
<td>469</td>
</tr>
<tr>
<td>$\hat{\alpha}_n$</td>
<td>0.581</td>
<td>0.023</td>
<td>0.143</td>
<td>0.508</td>
<td>0.413</td>
</tr>
<tr>
<td>$\hat{\beta}_n$</td>
<td>0.696</td>
<td>0.979</td>
<td>0.927</td>
<td>0.765</td>
<td>0.750</td>
</tr>
<tr>
<td>$T_n$</td>
<td>-2.297</td>
<td>0.024</td>
<td>1.120</td>
<td>0.491</td>
<td>0.457</td>
</tr>
<tr>
<td>$p$-val</td>
<td>0.011</td>
<td>0.510</td>
<td>0.869</td>
<td>0.688</td>
<td>0.676</td>
</tr>
</tbody>
</table>

series. The positive estimated value of $\gamma_0$ 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 non stationary GARCH(1,1) processes. The practical implications of our results are the following.

i) If one is interested in inference on $(\alpha_0, \beta_0)$ in the GARCH(1,1) model, then stationarity testing is unnecessary. The standard QMLE of $(\alpha_0, \beta_0)$ is CAN in every (stationary or non stationary) situation. A key result, allowing 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 $\omega_0$ is used (e.g. estimating the variance of today’s or tomorrow’s conditional distribution for derivatives pricing or for computing forecast intervals), then stationarity testing is necessary because if $\gamma_0 > 0$, the QMLE of $\omega_0$ is not consistent.

iii) The constrained QMLE of $(\alpha_0, \beta_0)$, which was known to be CAN in the non stationary 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 non stationarity 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 of the present paper works for detecting stationarity in other volatility models (not restricted to augmented GARCH). Second, it could be worth developing specific stationarity tests for particular augmented GARCH models.

APPENDIX A: 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 \( \rho > e^{-\gamma_0} \),

\[
\rho^t h_t \to \infty \quad \text{and, if } E[\log(\eta^2_t)] < \infty, \quad \rho^t \epsilon_t^2 \to \infty \quad \text{a.s. as } t \to \infty.
\]

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

\[
h_t \to \infty \quad \text{and, if } E[\log(\eta^2_t)] < \infty, \quad \epsilon_t^2 \to \infty \quad \text{in probability.}
\]

iii) Let \( \psi \) be a decreasing bijection from \((0, \infty)\) to \((0, \infty)\) such that \( E\psi(\epsilon mass) < \infty \). When \( \gamma_0 = 0 \) and \( \alpha_0 > 0 \),

\[
(A.1) \quad \psi(\epsilon_t^2) \to 0 \quad \text{and } \psi(h_t) \to 0 \quad \text{in } L^1.
\]

**Proof.** To prove i) we note that, for \( t > 1 \),

\[
h_t = \omega_0 \left\{ 1 + \sum_{i=1}^{t-1} a_0(\eta_{t-1}) \ldots a_0(\eta_{t-i}) \right\} + a_0(\eta_{t-1}) \ldots a_0(\eta_0) h_0
\]

(A.2) \( \geq \omega_0 \prod_{i=1}^{t-1} a_0(\eta_i). \)

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 \( E[\log(\eta^2_t)] < \infty \) entails \( \log(\eta^2_t)/t \to 0 \) a.s as \( t \to \infty \). Therefore, \( \liminf_{t \to \infty} t^{-1} \log \rho^t \eta^2_t h_t \geq E \log \rho a_0(\eta_1) > 0 \), and \( \rho^t \epsilon_t^2 = \rho^t \eta^2_t h_t \to +\infty \) a.s. by already given arguments.

The proof of ii) follows from Klüppelberg et al. (2004). Their condition \( E[\log(\delta + \lambda \epsilon^2_t)] < \infty \) becomes in our notations \( E[\log a_0(\eta_1)] < \infty \), and this condition is satisfied because \( E \log^+ a_0(\eta_1) \leq \gamma_0 = 0 \) 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 \( h_t > \alpha_0 \epsilon^2_{t-1} \) with \( \alpha_0 > 0 \), we have \( \psi(h_t) < \psi^*(\epsilon^2_{t-1}) \)

\[
where \psi^*(x) = \psi(\alpha_0 x) \text{ satisfies the same assumptions as } \psi(x). \text{ Therefore, the second convergence in (A.1) will follow from the first convergence. It suffices to consider the case } \epsilon_0 = 0. \text{ Note that, even if } \epsilon_t^2 \text{ does not increase with probability one, } \epsilon^2_{t+1} \text{ is stochastically greater than } \epsilon_t^2 \text{ because}
\]

\[
\epsilon^2_{t+1} = \left\{ \omega_0 + \omega_0 a_0(\eta_1) \ldots + \omega_0 a_0(\eta_1) \ldots a_0(\eta_2) + \omega_0 a_0(\eta_1) \ldots a_0(\eta_1) \right\} \eta^2_{t+1}
\]

\[
\geq \left\{ \omega_0 + \omega_0 a_0(\eta_1) \ldots + \omega_0 a_0(\eta_1) \ldots a_0(\eta_2) \right\} \eta^2_{t+1}
\]

\[
= \epsilon_t^2.
\]
where stands for equality in distribution. The dominated convergence theorem and i)-ii) then entail
\[ E\psi(\epsilon_t^2) = \int_0^\infty P\{\epsilon_t^2 < \psi^{-1}(u)\} \, du \to \int_0^\infty \lim_{t \to \infty} \downarrow P\{\epsilon_t^2 < \psi^{-1}(u)\} \, du = 0, \]
which completes the proof. \(\square\)

### A.2. Asymptotic normality of the QMLE of \((\alpha_0, \beta_0)\)

Let \(\omega = \inf\{\omega | \theta \in \Theta\}, \underline{\alpha} = \inf\{\alpha | \theta \in \Theta\}, \underline{\beta} = \inf\{\beta | \theta \in \Theta\}, \underline{\tau} = \sup\{\omega | \theta \in \Theta\}, \bar{\beta} = \sup\{\beta | \theta \in \Theta\}.\) 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 \(\Theta_0 = \{\theta \in \Theta : \beta < e^{\omega}\}\) and \(\Theta_p = \{\theta \in [0, \infty)^3 : \beta < ||1/a_0(\eta_{t})||_p^{-1}\}\).

**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 \(\Theta_p^* \subset \Theta_0\),
\[ \sup_{\theta \in \Theta_p^*} \left| \frac{\sigma_t^2(\theta)}{h_t} - v_t(\alpha, \beta) \right| \to 0 \text{ a.s. as } t \to \infty. \]
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 \(\Theta_p^* \subset \Theta_p\),
\[ \sup_{\theta \in \Theta_p^*} \left| \frac{\sigma_t^2(\theta)}{h_t} - v_t(\alpha, \beta) \right| \to 0 \text{ in } L^p. \]

**Proof.** Without loss of generality, assume that \(\sigma_0^2(\theta) = 0\). We then have \(\sigma_t^2(\theta) = \sum_{j=1}^t \beta^j \omega + \alpha \epsilon_{t-j}^2\) and
\[
\frac{\sigma_t^2(\theta)}{h_t} = \sum_{j=1}^t \beta^j \left( \prod_{k=1}^{j-1} \frac{h_{t-k}}{h_{t-k+1}} \right) \frac{\omega + \alpha \epsilon_{t-j}^2}{h_t} = a_t + b_t,
\]
where
\[ a_t = \sum_{j=1}^t \beta^j \left( \prod_{k=1}^{j-1} \frac{h_{t-k}}{h_{t-k+1}} \right) \alpha \epsilon_{t-j}^2 \quad \text{and} \quad b_t = \sum_{j=1}^t \beta^j \frac{\omega}{h_t}. \]
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 $\overline{\gamma}_0 = \sup\{\beta \mid \theta \in \Theta_0^*\}$, $b_t \leq K \overline{\gamma}(t + \overline{\gamma}_0)/h_t \to 0$ a.s. by $i$) of Proposition A.1. It follows that $\sup_{\theta \in \Theta_0^*} b_t \to 0$ a.s. Now note that

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

(A.4)

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 $\sup_{\theta \in \Theta_0^*} \sum_{j=t_0+1}^{\infty} v_tj = \sum_{j=t_0+1}^{\infty} v_tj(\alpha_0, \overline{\gamma}_0)$, 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 $\rho > e^{-\gamma_0}$,

$$\sigma^2_{\theta} (\beta) \geq \sum_{j=1}^{t_0} a_tj = \sum_{j=1}^{t_0} v_tj + o(p^{t_0}) \quad \text{a.s.}$$

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 $\gamma_0 = 0 = E \log(\alpha_0 \eta_1^2 + \beta_0)$ 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 $iii)$ of Proposition A.1. 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. 

\textbf{Lemma A.2} If $\theta \in \Theta_0$, we have

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

\textbf{Proof.} Straightforward algebra shows that

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

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. iff $\beta v_t(\alpha, \beta) - \beta_0 + (\alpha - \alpha_0)\eta^2_{t-1} = 0$. By strict stationarity $v_{t-1}(\alpha, \beta) = 1$, a.s. and we have $\beta - \beta_0 + (\alpha - \alpha_0)\eta^2_{t-1} = 0$. Because the distribution of $\eta^2_{t-1}$ is non degenerate, the conclusion follows. \hfill \Box

**Proof of i)–iii) in Theorem 2.1.** 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 $Q_n(\theta) = n^{-1}\sum_{t=1}^{n} \{k_t(\theta) - \ell_t(\theta_0)\}$. 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 $\Theta_0^*$ is an arbitrary compact subset of $\Theta_0$. We have by stationarity and ergodicity of $v_t(\alpha, \beta)$, a.s.

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

because $\log x < x - 1$ for $x > 0$. The inequality is strict except when $v_t(\alpha, \beta) = 1$ 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 $(\alpha, \beta)'s$ such that $(\omega, \alpha, \beta)' \in \Theta$.

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

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

**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,

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

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

$$\frac{1}{v_t(\alpha, \beta)} \leq \left( \frac{\alpha_0}{\alpha} + \frac{\beta_0}{\alpha n_{t-1}^2} \right) \frac{a_0(\varepsilon)}{\beta} \leq K(\varepsilon) \frac{a_0(\varepsilon)}{\beta}.$$
The conclusion follows by the same arguments as before. 

Thus, for any integer $k$

\[
\sup_{(\alpha, \beta) \in \Theta_{\alpha, \beta}} \left( \frac{1}{v_t(\alpha, \beta)} \right)^k \leq \{K(\varepsilon)\}^k \sum_{i=1}^{\infty} I_{|\eta_{i-1}| \leq \varepsilon} \cdots I_{|\eta_{i-1}| \leq \varepsilon} I_{|\eta_{i-1}| > \varepsilon} \left( \frac{a_0(\varepsilon)}{\beta} \right)^{k(i-1)}.
\]

It follows that,

\[
E \sup_{(\alpha, \beta) \in \Theta_{\alpha, \beta}} \left( \frac{1}{v_t(\alpha, \beta)} \right)^k \leq \{K(\varepsilon)\}^k \{1 - p(\varepsilon)\} \sum_{i=1}^{\infty} p(\varepsilon)^{i-1} \left( \frac{a_0(\varepsilon)}{\beta} \right)^{k(i-1)}.
\]

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

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

\[
\frac{h_t}{\sigma_t^2(\theta)} \leq \frac{\omega_0 + a_0(\eta_{i-1})h_{t-1}}{\omega + a h_{t-1}\eta_{t-1}^2 + \beta \sigma_{t-1}^2(\theta)} \leq \frac{\omega_0}{\omega} + \frac{a_0(\varepsilon)}{\beta} + \beta_0 \varepsilon := H(\varepsilon),
\]

and for $|\eta_{i-1}| \leq \varepsilon$ and $|\eta_{i-2}| > \varepsilon$,

\[
\frac{h_t}{\sigma_t^2(\theta)} \leq \frac{\omega_0}{\omega} + \frac{a_0(\varepsilon)}{\beta} H(\varepsilon).
\]

More generally,

\[
(A.6) \quad \sup_{\theta \in \Theta_0} \frac{h_t}{\sigma_t^2(\theta)} \leq V_t
\]

where

\[
V_t = \sum_{i=1}^{\infty} I_{|\eta_{i-1}| \leq \varepsilon} \cdots I_{|\eta_{i-1}| \leq \varepsilon} I_{|\eta_{i-1}| > \varepsilon} \left( \frac{\omega_0}{\omega} \sum_{j=0}^{i-2} \left( \frac{a_0(\varepsilon)}{\beta} \right)^j \varepsilon \right)^{i-1} \left( \frac{a_0(\varepsilon)}{\beta} \right)^{i-1} H(\varepsilon).
\]

The conclusion follows by the same arguments as before. 

Now we show that

\[
(A.7) \quad \lim_{n \to \infty} \sup_{\theta \in \Theta_0} |R_n(\theta)| = 0 \quad \text{a.s.}
\]

where $\Theta_0$ 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),

\[
(A.8) \quad \sup_{\theta \in \Theta_0} \left| \frac{1}{n} \sum_{t=1}^{n} \eta_t^2 \sigma_t^2(\theta) \left( v_t(\alpha, \beta) - \sigma_t^2(\theta) \right) \frac{h_t}{v_t(\alpha, \beta)} \right| \leq \frac{K \varepsilon}{n} \sum_{t=1}^{n} \frac{V_t}{v_t(\alpha, \beta)}
\]

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 $\Theta$. 


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

$$\frac{K}{n} \sum_{i=1}^{n} \left\| \frac{V_i}{\psi_t(\alpha, \beta)} \right\| \sup_{\theta \in \Theta_p} \left\| \nu_i(\alpha, \beta) - \frac{\sigma^2_t(\theta)}{h_t} \right\|_p,$$

which tends to zero by ii) of Lemma A.1. \qed

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

$$d_t^0(\alpha, \beta) = \sum_{j=1}^{\infty} \frac{\eta_{t-j}^2}{a(\eta_{t-j})} \prod_{k=1}^{j-1} \frac{\beta}{a(\eta_{t-k})}, \quad d_t^2(\alpha, \beta) = \sum_{j=2}^{\infty} \frac{(j-1)a_{\eta_{t-j}}^2 - j+1}{\beta a(\eta_{t-j})} \prod_{k=1}^{j-1} \frac{\beta}{a(\eta_{t-k})}.$$

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

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left( \frac{\partial}{\partial \alpha} f_t(\theta_0), \frac{\partial}{\partial \beta} f_t(\theta_0) \right) \xrightarrow{d} \mathcal{N} \{0, (\kappa_\eta - 1)I \} \quad \text{as } n \to \infty.$$

**Proof.** Using the Wold-Cramér device, it suffices to show that for all $\lambda = (\lambda_1, \lambda_2)'$, the sequence

$$\nabla_t = \left( \frac{\partial}{\partial \alpha} f_t(\theta_0), \frac{\partial}{\partial \beta} f_t(\theta_0) \right) \lambda = \frac{1 - \eta_t^2}{h_t} \lambda' \left( \frac{\partial \sigma^2_t(\theta_0)}{\partial \alpha}, \frac{\partial \sigma^2_t(\theta_0)}{\partial \beta} \right)$$

satisfies

(A.9) \quad \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \nabla_t \xrightarrow{d} \mathcal{N} \{0, (\kappa_\eta - 1)\lambda' \lambda \}.

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

$$\frac{\partial \sigma^2_t(\theta)}{\partial \alpha} = \sum_{j=1}^{t} \beta^j - 2 \epsilon_t, \quad \frac{\partial \sigma^2_t(\theta)}{\partial \beta} = \sum_{j=2}^{t} (j-1) \beta^{j-2} (\omega + \alpha \epsilon_t^2).$$

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

$$\frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \alpha} = \sum_{j=1}^{t} \beta^j \left\{ \prod_{k=1}^{j} \frac{\sigma_{t-k}^2(\theta)}{\sigma_{t-k+1}^2(\theta)} \right\} \epsilon_t \leq d_t^0(\alpha, \beta)$$

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

Moreover, we have

$$0 \leq d_t^0(\alpha, \beta) = \frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \alpha} \leq s_{t_0} + r_{t_0}.$$
where
\[ s_{t_0} = \sum_{j=1}^{t_0} \frac{\eta_j^2}{\alpha_j} \prod_{k=1}^{j-1} \frac{\beta_k}{\alpha_k} - \frac{\epsilon_j^2}{\sigma_{t-j}^2(\theta)} \prod_{k=1}^{j} \frac{\beta_k}{\sigma_{t-k+1}^2(\theta)}. \]
\[ r_{t_0} = \sum_{j=t_0+1}^{\infty} \frac{\eta_j^2}{\alpha_j} \prod_{k=1}^{j-1} \frac{\beta_k}{\alpha_k}. \]

For all \( p \geq 1, \| r_{t_0} \|_p \to 0 \) as \( t_0 \to \infty \) because \( \| \beta / \alpha_0 \|_p < 1 \). Since \( \| \eta_j^2 / \alpha_j \|_p < 1/\alpha_0, \| \beta \sigma_{t-1}(\theta) / \sigma_t(\theta) \|_p < 1, \)

\begin{align*}
\frac{\beta}{\alpha_0} - \frac{\beta \sigma_{t-1}(\theta)}{\sigma_t(\theta)} \to 0
\end{align*}

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 \( d_t^0(\alpha, \beta) \) is replaced by \( d_t^3(\alpha, \beta) \). It follows that \( d_t^0(\alpha, \beta) \) and \( d_t^3(\beta) \) have moments of any order, and

\[ (A.10) \quad \frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \alpha} - d_t^0(\alpha, \beta) \to 0 \quad \text{and} \quad \frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \beta} - d_t^3(\alpha, \beta) \to 0 \]
in \( L^p \) for any \( p \geq 1 \). Standard computations show that \( \mathcal{I} = \mathcal{E} d_1 \) where \( d_t = (d_t^0(\alpha_0, \beta_0), d_t^3(\alpha_0, \beta_0)) \). 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 \( n \to \infty \),

\[ \text{Var} \left( \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \nabla_t \right) = \frac{\kappa_n - 1}{n} \sum_{t=1}^{n} \mathbb{E} |\lambda d_t|^2 + o(1) \to (\kappa_n - 1) \lambda^2. \]

Moreover, for all \( \varepsilon > 0 \)
\begin{align*}
\frac{1}{n} \sum_{t=1}^{n} \mathbb{E} \left[ \left| \nabla_t \right|^2 \right] & \leq (\kappa_n - 1) E \left\{ |\lambda_1| d_t^0(\beta_0) + |\lambda_2| d_t^3(\alpha_0, \beta_0) \right\} \to 0.
\end{align*}

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

**Lemma A.5** Let \( \varpi \) be an arbitrary compact subset of \( [0, \infty) \). Assume that \( \mathbb{E} \log \eta_1^2 < \infty \). When \( \gamma_0 > 0 \) we have

\[ (A.11) \quad \sum_{i=1}^{\infty} \sup_{\omega \in \Theta_0} \left| \frac{\partial}{\partial \omega} f_i(\omega) \right| < \infty \quad \text{a.s.,} \]
\[ (A.12) \quad \sum_{i=1}^{\infty} \sup_{\omega \in \Theta_0} \left| \frac{\partial^2}{\partial \omega \partial \theta} f_i(\omega) \right| < \infty \quad \text{a.s.,} \]
\[ (A.13) \quad \sup_{\omega \in \varpi} \left| \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2 f_i(\omega, \alpha_0, \beta_0)}{\partial \alpha_{i+1} \partial \beta_{j+1}} - \mathcal{I}(i, j) \right| = o(1) \quad \text{a.s. for all } i, j \in \{1, 2\}, \]
\[ (A.14) \quad \frac{1}{n} \sum_{i=1}^{n} \sup_{\omega \in \Theta_0} \left| \frac{\partial^3}{\partial \alpha_i \partial \beta_j \partial \beta_k} f_i(\omega) \right| = O(1) \quad \text{a.s. for all } i, j, k \in \{2, 3\}. \]
When $\gamma_0 = 0$,

\begin{align}
\sup_{\omega \in \Omega} \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}} - I(i, j) \right| = \alpha_P(1) \quad \text{for all } i, j \in \{1, 2\},
\end{align}

\begin{align}
\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_P(1) \quad \text{for all } i, j, k \in \{2, 3\}.
\end{align}

**Proof.** Let first suppose $\gamma_0 > 0$. Using Proposition A.1-i) and Lemma A.1-i), $\sigma_t^2(\theta) > \alpha e^{\gamma_0}$, and 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

\begin{align}
\left| \frac{\partial}{\partial \omega} \ell_t(\theta) \right| = \left| \frac{-h_t \eta_t^2}{\sigma_t^2(\theta)} + \frac{1}{\sigma_t^2(\theta)} \sum_{j=1}^{t} \beta^{j-1} \right| \leq K \rho^t \left( \frac{\eta_t^2}{\nu_{t-1}(\alpha, \beta)} + 1 \right) \quad \text{a.s.}
\end{align}

Since, in view of Lemma A.3, $\sum_{t=1}^{n} K \rho^t (\eta_t^2 / \nu_t(\alpha, \beta) + 1)$ 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 $Z(1, 2) = \lim_{n \to \infty} n^{-1} \sum_{t=1}^{n} d_t^2(\alpha_0, \beta_0) d_t^2(\alpha_0, \beta_0)$ a.s. Moreover, we have

\begin{align}
\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) \left( \frac{1}{\sigma_t^2} \sum_{j=2}^{t} (j-1) \beta_0^{j-2} (\omega + \alpha_0 \epsilon_{t-j}) \right)
\end{align}

\begin{align}
+ \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{align}

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

\begin{align}
\sup_{\omega \in \Omega} \left| \sum_{j=1}^{t} \beta_0^{j-1} \epsilon_{t-j}^2 - d_t(\alpha_0, \beta_0) \right| &\to 0,
\end{align}

\begin{align}
\sup_{\omega \in \Omega} \left| \frac{\sum_{j=2}^{t} (j-1) \beta_0^{j-2} (\omega + \alpha_0 \epsilon_{t-j})}{h_t} - d_t(\alpha_0, \beta_0) \right| &\to 0,
\end{align}

\begin{align}
\sup_{\omega \in \Omega} \left| \frac{\sum_{j=2}^{t} (j-1) \beta_0^{j-2} \epsilon_{t-j}^2}{h_t} - \frac{1}{\alpha_0} d_t^2(\alpha_0, \beta_0) \right| &\to 0,
\end{align}

\begin{align}
\sup_{\omega \in \Omega} \left| \frac{h_t}{\sigma_t^2} - 1 \right| &\to 0.
\end{align}

In view of $\sigma_t^2 - h_t = \sum_{j=1}^{t} \beta_0^{j-1} (\omega - \omega_0)$ and $1/h_t = o(p^t)$ for some $p \in (e^{-\gamma_0}, \beta_0^{-1})$, 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, with now $\sigma^2_i = \sigma^2_i(\theta)$
\[
\frac{\partial^3 \ell_i(\theta)}{\partial \alpha \partial \beta} = \left(2 - 6\eta^2 \frac{h_i}{\sigma^2_i}\right) \left(\frac{1}{\sigma^2_t} \sum_{j=1}^i \beta^{j-1} \epsilon^2_{i-j}\right)^2 \left(\frac{1}{\sigma^2_t} \sum_{j=2}^i (j-1) \beta^{j-2} (\omega + \alpha \epsilon^2_{i-j})\right)
+ \left(2\eta^2 \frac{h_i}{\sigma^2_t} - 1\right) \left(\frac{1}{\sigma^2_t} \sum_{j=1}^i \beta^{j-1} \epsilon^2_{i-j}\right) \left(\frac{2}{\sigma^2_t} \sum_{j=2}^i (j-1) \beta^{j-2} \epsilon^2_{i-j}\right)
\]
\[
= \left(2 - 3\eta^2 \frac{1}{v_i(\alpha, \beta)}\right) (d^3_0(\alpha, \beta))^2 d^3_0(\alpha, \beta)
+ \left(2\eta^2 \frac{1}{v_i(\alpha, \beta)} - 1\right) d^3_0(\alpha, \beta) \left(\frac{2}{\alpha}\right) a(\alpha, \beta) + o(1) \text{ a.s.}
\]
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 $d^3_0(\alpha, \beta)$ and $d^3_0(\alpha, \beta)$ 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)-(A.14) can be replaced by $L^p$ convergences, by the arguments used to show (A.10). We then obtain (A.15) and (A.16) from Proposition A.1 iii). \hfill \square

**Proof of iv)-v) in Theorem 2.1.** 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 $\Theta$, even asymptotically. However, the partial derivative with respect to $(\alpha, \beta)$ is asymptotically equal to zero at the minimum since $(\hat{\alpha}_n, \hat{\beta}_n) \rightarrow (\alpha_0, \beta_0)$ and $\theta_0$ belongs to the interior of $\Theta$. Hence, an expansion of the criterion derivative gives

\[
(A.22) \quad \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \theta} \ell_t(\hat{\theta}_n) \right) \sim \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \theta} \ell_t(\theta_0) + J_n (\hat{\theta}_n - \theta_0)
\]
where $J_n$ is a $3 \times 3$ matrix whose elements have the form

\[
J_n(i, j) = \frac{1}{n} \sum_{t=1}^n \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell_t(\theta_0)
\]
where $\theta^*_n = (\omega^*_n, \alpha^*_n, \beta^*_n)'$ is between $\hat{\theta}_n$ and $\theta_0$. Since $\theta_0 \in \Theta_0$, we have $\theta^*_n \in \Theta_0$ for $n$ large enough. By (A.12) in Lemma A.5 and the compactness of $\Theta$, we have, for $i = 2, 3,$

\[
(A.23) \quad J_n(i, 1) \sqrt{n}(\hat{\omega}_n - \omega_0) \leq \sum_{t=1}^\infty \sup_{\theta \in \Theta_0} \left\| \frac{\partial^2}{\partial \omega \partial \theta} \ell_t(\theta) \right\| \frac{1}{\sqrt{n}} (\hat{\omega}_n - \omega_0) \rightarrow 0 \text{ a.s.}
\]
An expansion of the function

\[(\alpha, \beta) \mapsto \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial \alpha^2} \ell_t(\omega^*_t, \alpha, \beta)\]

gives

\[J_n(2, 2) = \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial \alpha^2} \ell_t(\omega^*_t, \alpha_0, \beta_0) + \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^3 \ell_t(\omega^*_t, \alpha^*_t, \beta^*_t)}{\partial (\alpha, \beta) \partial \alpha^2} \begin{pmatrix} \alpha^*_t - \alpha_0 \\ \beta^*_t - \beta_0 \end{pmatrix}\]

where \((\alpha^*, \beta^*)\) is between \((\alpha^*_t, \beta^*_t)\) and \((\alpha_0, \beta_0)\). Using (A.13), (A.14) and the consistency of \((\alpha^*, \beta^*)\), we get \(J_n(2, 2) \to \mathcal{I}(1, 1)\) \(\text{a.s.}\), and similarly, for \(i, j = 1, 2\)

\[(A.24) \quad J_n(i + 1, j + 1) \to \mathcal{I}(i, j) \quad \text{a.s.}\]

The conclusion follows, by considering the last two components in (A.22) and from Lemma A.4, (A.23) and (A.24). \(\square\)

**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

\[\frac{1}{\sigma^2_t} \sum_{j=1}^{t} \beta^{j-1} \epsilon_{t-j}^2 \leq \frac{1}{\alpha},\]

and \(\beta^*_2 < 1\) for \(n\) large enough, we obtain

\[|J_n(2, 1)\sqrt{n}(\hat{\omega}_n - \omega_0)| \leq \frac{K}{\sqrt{n}} \sum_{t=1}^{n} \left\{ \frac{2h_t \eta_t^2}{\sigma^2_t(\beta^*_2)} + 1 \right\} \left\{ \sum_{j=1}^{t} (\beta^*_2)^{j-1} \epsilon_{t-j}^2 \right\} \left\{ \sum_{j=1}^{t}(\beta^*_2)^{j-1} \right\} \frac{1}{\sigma^2_t(\beta^*_2) h_t} \]

\[\text{(A.25)} \quad \leq \frac{K}{\sqrt{n}} \sum_{t=1}^{n} \left\{ \frac{2h_t \eta_t^2}{\sigma^2_t(\beta^*_2)} + 1 \right\} \frac{h_t}{\sigma^2_t(\beta^*_2) h_t} \]

Hence, by Lemma A.3

\[E[J_n(2, 1)\sqrt{n}(\hat{\omega}_n - \omega_0)] \leq \frac{K}{\sqrt{n}} \sum_{t=1}^{n} E \frac{1}{h_t} \]

The same bound is obtained when \(J_n(2, 1)\) is replaced by \(J_n(3, 1)\). 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 \(i = 2, 3\),

\[E[J_n(i, 1)\sqrt{n}(\hat{\omega}_n - \omega_0)] \to 0 \text{ as } n \to \infty.\]

\[\text{(A.26)} \quad \]
Finally, using $iii)$ in Theorem 2.1, (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 $\gamma_0 > 0$. 

\begin{proof}
\textbf{Proof of Theorem 2.3.} The convergence results in $i)$ can be shown in a standard way, using Taylor expansions of the functions $\kappa_n = \kappa_n(\hat{\theta}_n)$ and $\frac{1}{n} \sum_{t=1}^{n} \frac{\partial \sigma_i^2(\theta_n)}{\partial \theta} \frac{\partial \sigma_i^2(\theta_n)}{\partial \theta} (\hat{\theta}_n - \theta_0)$ around $\theta_0$, and the ergodic theorem together with the consistency of $\hat{\theta}_n$.

Now consider the case $ii)$. For some $\theta^* = (\omega^*, \alpha^*, \beta^*)^T$ between $\hat{\theta}_n$ and $\theta_0$ we have

\begin{equation}
\kappa_n = \frac{1}{n} \sum_{t=1}^{n} \eta_t^4 - 2 \frac{1}{n} \sum_{t=1}^{n} \frac{\partial \sigma_i^2(\theta^*)}{\partial \theta} \frac{\partial \sigma_i^2(\theta^*)}{\partial \theta} (\hat{\theta}_n - \theta_0) = \frac{1}{n} \sum_{t=1}^{n} \eta_t^4 + R_n.
\end{equation}

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

$$|R_n| \leq K \sum_{t=1}^{n} \eta_t^4 \left( \frac{h_t}{\sigma_i^2(\theta_0)} \right)^2 \left( \rho^t |\hat{\omega}_n - \omega_0| + d_i^2(\alpha^*, \beta^*)|\alpha_n - \alpha_0| + d_i^2(\alpha^*, \beta^*)|\beta_n - \beta_0| \right),$$

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 $\Theta$, and the existence moments at any order for $d_i^2(\alpha^*, \beta^*)$, $d_i(\alpha^*, \beta^*)$ and $h_t/\sigma_i^2(\theta^*)$. Hence the first part of $ii)$ is proven. Now, similarly to (A.25), we have

\begin{equation}
n\hat{\mathcal{J}}_{\omega, \alpha} \leq \sum_{t=1}^{n} \frac{h_t}{\sigma_i^2(\theta_n)} \sum_{j=1}^{t} \hat{\beta}_{n,j}^{-1} \leq K \sum_{t=1}^{n} \frac{h_t}{\sigma_i^2(\theta_n)} \rho^t,
\end{equation}

for $\rho \in (0,1)$ when $n$ is large enough, by Proposition A.1 $i)$. It follows that $n\hat{\mathcal{J}}_{\omega, \alpha} = O(1)$ a.s. Similarly $n\hat{\mathcal{J}}_{\omega, \beta} = O(1)$ a.s. Moreover, we have $n\hat{\mathcal{J}}_{\omega, \omega} \geq 1/\sigma_i^2(\hat{\theta}_n) > 0$. Thus we have shown that

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

Now we turn to $\hat{\mathcal{J}}_{\alpha, \beta, \omega}$. Considering the top-left term, a Taylor expansion around $\theta_0$ gives

\begin{equation}
\mathcal{J}_{\alpha, \alpha} = \frac{1}{n} \sum_{t=1}^{n} \left( \frac{1}{\sigma_i^2(\theta_n)} \sum_{j=1}^{t} \hat{\beta}_{n,j}^{-1} \epsilon_{t-j}^2 \right)^2
\end{equation}

\begin{equation}
\mathcal{J}_{\alpha, \beta} = \frac{1}{n} \sum_{t=1}^{n} \{d_i^2(\alpha_0, \beta_0)\}^2 + \frac{1}{n} \sum_{t=1}^{n} \left[ \left\{ \frac{1}{\sigma_i^2(\theta_n)} \frac{\partial \sigma_i^2(\theta)}{\partial \alpha} \right\} - \{d_i^2(\alpha_0, \beta_0)\} \right]^2 + S_n
\end{equation}

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

$$K \sum_{t=1}^{n} \left( \frac{\sum_{j=1}^{t} \hat{\beta}_{n,j}^{-1} \epsilon_{t-j}^2}{\sigma_i^2(\theta^*)} \right) \left( \frac{\sum_{j=1}^{t} (j-1) \hat{\beta}_{n,j}^{-1} \epsilon_{t-j}^2}{\sigma_i^2(\theta^*)} \right)^2 \rho^t |\hat{\omega}_n - \omega_0| + d_i^2(\alpha^*, \beta^*)|\alpha_n - \alpha_0| + d_i^2(\alpha^*, \beta^*)|\beta_n - \beta_0|$$

$$+ K \sum_{t=1}^{n} \left( \frac{\sum_{j=1}^{t} \hat{\beta}_{n,j}^{-1} \epsilon_{t-j}^2}{\sigma_i^2(\theta^*)} \right) \left( \frac{\sum_{j=1}^{t} (j-1) \hat{\beta}_{n,j}^{-1} \epsilon_{t-j}^2}{\sigma_i^2(\theta^*)} \right) |\beta_n - \beta_0| = o(1), \quad \text{a.s.}$$
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 $1/\alpha_0^2 = I(1,1)$ since $d^\alpha_t(\alpha_0, \beta_0) = v_t(\alpha_0, \beta_0)/\alpha_0 = 1/\alpha_0$ by Lemma A.2. We thus have shown that $\hat{J}_{\alpha,\alpha}$ a.s. converges to $I(1,1)$. The other two terms in $\hat{J}_{\alpha,\beta,\alpha,\beta}$ can be handled similarly, which completes the proof of (ii).

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

$$
\frac{K}{n} \sum_{t=1}^n \eta^4_t \left( \frac{h_t}{\sigma^2_t(\theta^*_t)} \right)^2 \frac{1}{\omega^* + \alpha^* \beta^* \epsilon^2_{t-2}} + \frac{K}{n} \sum_{t=1}^n \eta^4_t \left( \frac{h_t}{\sigma^2_t(\theta^*_t)} \right)^2 \left( d^\alpha_t(\alpha^*, \beta^*)|\hat{\alpha}_n - \alpha_0| + d^\beta_t(\alpha^*, \beta^*)|\hat{\beta}_n - \beta_0| \right).
$$

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 $d^\alpha_t(\alpha^*, \beta^*)$, $d^\beta_t(\alpha^*, \beta^*)$ and $h_t/\sigma^2_t(\theta^*_t)$. Hence $\hat{\kappa}_n \to \kappa$ in probability. Now, in view of the first inequality of (A.28), we have $n\hat{J}_{\omega,\alpha} \leq K \sum_{t=1}^n \frac{h_t}{\sigma^2_t(\theta^*_t)} \frac{1}{h_t}$, and thus $n\hat{J}_{\omega,\alpha} = O_P(1)$ by Lemma A.3 and the arguments used to show (A.26). Proceeding as in the proof of (i), we deduce that $\hat{J}_{\alpha,\beta,\omega} \hat{J}_{\omega,\alpha,\beta} = o_P(1)$. By similar arguments, the right-hand side of (A.29) converges to $I(1,1)$ in probability. □

**Proof of Proposition 2.1** If a consistent estimator $\hat{\theta}_n$ existed, then the test of critical region $C = \{ |\hat{\theta}_n - \theta| > |\hat{\theta}_n - \theta^*| \}$ would have null asymptotic errors of first and second kind, in contradiction with Proposition 3.1 below. □

### A.3. Constrained QMLE of $(\alpha_0, \beta_0)$

**Proof of (2.8).** By the arguments used to prove Theorem 2.1-(iii), we have

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

A Taylor expansion of the criterion derivative gives

$$
0 = \frac{1}{\sqrt{n}} \sum_{t=1}^n \partial (\omega, \hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega))
$$

$$
= \frac{1}{\sqrt{n}} \sum_{t=1}^n \partial (\omega, \alpha_0, \beta_0)
$$

$$
(\text{A.31}) \quad + \left( \frac{1}{\sqrt{n}} \sum_{t=1}^n \partial^2 (\omega, \alpha_0, \beta_0) \sum_{t=1}^n \partial (\omega, \alpha_0, \beta_0) \right) \sqrt{n}(\hat{\alpha}_n^c(\omega) - \alpha_0, \hat{\beta}_n^c(\omega) - \beta_0)'
$$

ectaart.cls ver. 2006/04/11 file: ECMA_9405final.tex date: September 1, 2011
where \((\alpha^*, \beta^*)\) is between \((\hat{\alpha}_n^*(\omega), \hat{\beta}_n^*(\omega))\) and \((\alpha_0, \beta_0)\). Another Taylor expansion yields, for \((\alpha^{**}, \beta^{**})\) between \((\hat{\alpha}_n^*(\omega), \hat{\beta}_n^*(\omega))\) and \((\alpha_0, \beta_0)\)
\[
\left| \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial(\alpha, \beta)\partial(\alpha, \beta)} L_t(\omega, \alpha_0, \beta_0) - \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial(\alpha, \beta)\partial(\alpha, \beta)} L_t(\omega, \alpha^*, \beta^*) \right|
\leq |\alpha^* - \alpha_0| \frac{1}{n} \sum_{t=1}^{n} \left| \frac{\partial^3}{\partial(\alpha, \beta)\partial(\alpha, \beta)} L_t(\omega, \alpha^{**}, \beta^{**}) \right|
\quad + |\beta^* - \beta_0| \frac{1}{n} \sum_{t=1}^{n} \left| \frac{\partial^3}{\partial(\alpha, \beta)\partial(\alpha, \beta)} L_t(\omega, \alpha, \beta) \right|
\leq |\alpha^* - \alpha_0| \frac{1}{n} \sum_{t=1}^{n} \sup_{\theta \in \Theta} \left| \frac{\partial^3}{\partial(\alpha, \beta)\partial(\alpha, \beta)} L_t(\omega, \alpha, \beta) \right|
\quad + |\beta^* - \beta_0| \frac{1}{n} \sum_{t=1}^{n} \sup_{\theta \in \Theta} \left| \frac{\partial^3}{\partial(\alpha, \beta)\partial(\alpha, \beta)} L_t(\omega, \alpha, \beta) \right|
= o_p(1),
\]

using (A.30) and (A.16). Therefore, using (A.15), the term in parentheses in (A.31) converges to \(I\). To conclude, it remains to prove that

\[
(A.32) \quad \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial(\alpha, \beta)} L_t(\omega, \alpha_0, \beta_0) \xrightarrow{d} \mathcal{N}(0, (k_n - 1)I).
\]

We have
\[
\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial(\alpha, \beta)} L_t(\omega, \alpha_0, \beta_0) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial(\alpha, \beta)} L_t(\omega_0, \alpha_0, \beta_0)
\quad + \frac{\omega - \omega_0}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial^2}{\partial(\alpha, \beta)\partial(\alpha, \beta)} L_t(\omega^*, \alpha_0, \beta_0)
\]

where \(\omega^*\) is between \(\omega_0\) and \(\omega\). The last term tends to zero in probability, using A, similarly to (A.26). The first term converges in distribution to the normal law of (A.32) by Lemma A.4. \(\square\)

**Proof of Proposition 2.2.** The ergodic theorem entails that, for \(\beta < 1\), almost surely,
\[
L_n(\alpha, \beta) := \frac{1}{n} \sum_{t=1}^{n} \frac{e^2_t}{\sigma_t^2(\omega, \alpha, \beta)} + \log \sigma_t^2(\omega, \alpha, \beta)
\rightarrow L(\alpha, \beta) = E_\infty \left\{ \frac{h_t}{\sigma_t^2(\omega, \alpha, \beta)} + \log \sigma_t^2(\omega, \alpha, \beta) \right\}
\]
as \(n \to \infty\). The dominated convergence theorem implies that
\[
\frac{\partial L}{\partial \alpha}(\alpha_0, \beta_0) = E_\infty \left\{ \left(1 - \frac{h_t}{\sigma_t^2(\omega, \alpha_0, \beta_0)} \right) \frac{1}{\sigma_t^2(\omega, \alpha_0, \beta_0)} \frac{\partial \sigma_t^2(\omega, \alpha_0, \beta_0)}{\partial \alpha} \right\}
\quad = (\omega - \omega_0) \left( \sum_{i \geq 0} \beta_i^0 \right) E_\infty \left\{ \frac{1}{\sigma_t^2(\omega, \alpha_0, \beta_0)} \frac{\partial \sigma_t^2(\omega, \alpha_0, \beta_0)}{\partial \alpha} \right\} \neq 0.
\]

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
\[
L_n \{ \hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega) \} = L_n(\alpha_0, \beta_0) + \frac{\partial L_n}{\partial \alpha}(\hat{\alpha}_n, \hat{\beta}_n)\{\hat{\alpha}_n^c(\omega) - \alpha_0\} + \frac{\partial L_n}{\partial \beta}(\hat{\alpha}_n, \hat{\beta}_n)\{\hat{\beta}_n^c(\omega) - \beta_0\},
\]
where $(\hat{\alpha}_n, \hat{\beta}_n)$ is between $(\alpha_n^c(\omega), \beta_n^c(\omega))$ and $(\alpha_0, \beta_0)$. Note that since $E\epsilon_1^4 < \infty$, almost surely, for $\beta < 1$
\[
\lim_{n \to \infty} \sup_{\alpha} \sup_{\beta < 1} \left| \frac{\partial L_n}{\partial \alpha}(\alpha, \beta_0) \right| \leq \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left( 1 + \frac{\epsilon_i^2}{\omega} \right) \sum_{i=1}^{n} \beta_i^2 \epsilon_{i-1-i}^2 < \infty,
\]
\[
\lim_{n \to \infty} \sup_{\alpha} \sup_{\beta < 1} \left| \frac{\partial L_n}{\partial \beta}(\alpha, \beta_0) \right| \leq \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left( 1 + \frac{\epsilon_i^2}{\omega} \right) \sum_{i=1}^{n} \beta_i^2 \sigma_{i-1-i}^2 < \infty.
\]
Now suppose that
\[
(A.34) \quad (\hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega)) \to (\alpha_0, \beta_0), \quad \text{in probability as } n \to \infty.
\]
Then, it follows from (A.33) that
\[
L_n \{ \hat{\alpha}_n^c(\omega), \hat{\beta}_n^c(\omega) \} \to L(\alpha_0, \beta_0), \quad \text{in probability as } n \to \infty.
\]
Then, taking the limit in probability in the following 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. \hfill \Box

\section*{A.4. Stationarity test}

\textbf{Proof of Theorem 3.1.} Let $\gamma_n(\theta) = n^{-1} \sum_{i=1}^{n} \log \{ \alpha \eta_i^2(\theta) + \beta \}$ and $\eta_i(\theta) = \epsilon_i/\sigma_i(\theta)$.
First consider the case $\gamma_0 < 0$. A Taylor expansion gives
\[
\hat{\gamma}_n = \gamma_n(\theta_0) + \frac{\partial \gamma_n(\theta_0)}{\partial \theta} \hat{\theta} + o_P(n^{-1/2})
\]
with
\[
\left( A.35 \right) \quad \frac{\partial \gamma_n(\theta_0)}{\partial \theta} = -\frac{1}{n} \sum_{i=1}^{n} \frac{1}{a_0(\eta_i)} \left\{ \alpha_0 \eta_i^2 \frac{1}{h_2} \frac{\partial \sigma_i^2(\theta_0)}{\partial \theta} - \left( \frac{0}{\eta_i^2} \right) \right\} = -\Psi + o_P(1),
\]
where, in view of $\alpha_0 \eta_i^2/a_0(\eta_i) = 1 - \beta_0/a_0(\eta_1)$
\[
\Psi = (1 - \nu_1)\Omega - a, \quad \Omega = E\infty \frac{1}{h_2} \frac{\partial \sigma_i^2(\theta_0)}{\partial \theta}.
\]
Moreover the QMLE satisfies
\[(A.37) \quad \sqrt{n}(\hat{\theta}_n - \theta_0) = -\mathcal{J}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (1 - \eta_t^2) \frac{1}{h_t} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} + o_p(1).\]

In view of (A.35), (A.36) and (A.37), we have
\[(A.38) \quad \sqrt{n}(\hat{\gamma}_n - \gamma_0) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} u_t + \Psi' \mathcal{J}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (1 - \eta_t^2) \frac{1}{h_t} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} + o_p(1).\]

Note that
\[
\text{Cov}\left( \frac{1}{\sqrt{n}} \sum_{t=1}^{n} u_t, \Psi' \mathcal{J}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (1 - \eta_t^2) \frac{1}{h_t} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \right) = c \Omega' \mathcal{J}^{-1} \Psi,
\]
where \(c = \text{Cov}(u_t, 1 - \eta_t^2).\) The Slutsky lemma and the central limit theorem for martingale differences thus entail
\[\sqrt{n}(\hat{\gamma}_n - \gamma_0) \overset{d}{\to} N\left(0, \sigma_u^2 + 2c \Omega' \mathcal{J}^{-1} \Psi + (\kappa_0 - 1) \Psi' \mathcal{J}^{-1} \Psi \right).\]

Now let \(\bar{\theta}_0 = (\omega_0, \alpha_0, 0)'\). Noting that \(\partial_{\bar{\theta}_0} \partial \sigma_t^2(\theta_0)/\partial \theta = h_t\) almost surely, we have
\[E\left\{ \frac{1}{h_t} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \left( 1 - \frac{1}{h_t} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \bar{\theta}_0 \right) \right\} = 0,
\]
which entails \(\mathcal{J}(\bar{\theta}_0) = \Omega\) and \(\Omega' \mathcal{J}^{-1} \Omega = 1\). It follows that \(\Omega' \mathcal{J}^{-1} \Psi = 0\). Noting that \(\Psi' \mathcal{J}^{-1} \Psi = a' \mathcal{J}^{-1} a - (1 - \nu_0)^2\) the asymptotic distribution in (3.5) follows in the case \(\gamma_0 < 0\).

Now consider the case \(\gamma_0 \geq 0\). Let \(\theta'_0\) be a sequence such that \(\|\theta'_0 - \theta_0\| \leq \|\hat{\theta}_n - \theta_0\|\). By Proposition A.1 (using Assumption A when \(\gamma_0 = 0\)), we have
\[\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{1}{\sigma_t^2(\theta'_0)} \frac{\partial \sigma_t^2(\theta'_0)}{\partial \omega} = o(1), \text{ a.s. (resp. in probability) as } n \to \infty
\]
when \(\gamma_0 > 0\) (resp. when \(\gamma_0 = 0\)). It can be deduced that, under the same conditions, \(\sqrt{n} \frac{\partial^2 \gamma_0(\theta'_0)}{\partial \omega \partial \theta} = o(1)\), and
\[\sqrt{n}(\hat{\theta} - \theta_0) \frac{\partial^2 \gamma_0(\theta'_0)}{\partial \theta \partial \theta'}(\hat{\theta} - \theta_0) = o(1),
\]
which entails that (A.35) still holds. By the same arguments, (A.36) holds with
\[
\Omega = E\begin{pmatrix} 0 & 0 \\ 0 & 1/\alpha_0 & \nu_1 / \{\beta_0(1 - \nu_1)\} \end{pmatrix}
\]
and \(\Psi = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\).

The conclusion follows. \(\square\)

**Proof of Corollary 3.2.** The proof relies on arguments already used and is omitted. \(\square\)

**Proof of Corollary 3.3.** By arguments used in the proof of Theorem 2.3, \(\hat{\sigma}_u^2\) converges almost surely to \(\sigma_u^2\) when \(\gamma < 0\) or \(\gamma \geq 0\). Therefore \(T_n = \sqrt{n}(\hat{\gamma}_n - \gamma_0)/\hat{\sigma}_u + \sqrt{n} \gamma_0/\hat{\sigma}_u\)
where

\[(A.40)\]

\[S_n = \sum_{t=1}^{n} \frac{\epsilon_t^2}{\sigma_t^2(\theta_1)} + \log \sigma_t^2(\theta_1) - \frac{\epsilon_t^2}{\sigma_t^2(\theta_1^*)} - \log \sigma_t^2(\theta_1^*)\]

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

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

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

\[S_n \to S_0 = \varphi(\eta_1^2, \eta_2^2, \ldots; \theta_1, \theta_1^*) \text{ a.s. under } H_0\]

and

\[S_n \to S_1 = -\varphi(\eta_1^2, \eta_2^2, \ldots; \theta_1^*, \theta_1) \text{ a.s. under } H_1.\]

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

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

where

\[\varsigma_{t-1} = \frac{\omega_1 - \omega_1}{\omega_1 + \omega_1 \sum_{k=1}^{t} \alpha_k \eta_{t-k}^2}, \quad \varsigma_{t-1}^* = \frac{\omega_1 - \omega_1}{\omega_1 + \omega_1 \sum_{k=1}^{t} \alpha_k \eta_{t-k}^2} \cdot \eta_{t-k}^2.\]

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

\[\varsigma_{t-1} = \frac{\sum_{j=1}^{t} \beta_1^j (\omega_1 - \omega_1)}{\sigma_t^2(\theta_1)}, \quad \varsigma_{t-1}^* = \frac{\sum_{j=1}^{t} \beta_1^j (\omega_1 - \omega_1)}{\sigma_t^2(\theta_1)} \cdot \eta_{t-k}^2.\]

Using (A.39), it follows that

\[(A.41) \quad |S_0 - S_1| \leq K |\omega_1 - \omega_1| \sum_{t=0}^{\infty} \rho(t \eta_t^2 + 1).\]
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
\[
\lim_{n \to \infty} P_{H_1}(S_n > c_n) = P(S_1 > c), \text{ where } c \text{ is such that } P(S_0 > c) = \alpha.
\]
For any $\varepsilon > 0$, we have
\[
P(S_1 > c) \leq P(S_0 + |S_1 - S_0| > c) \leq P(S_0 > c - \varepsilon) + P(|S_1 - S_0| > \varepsilon).
\]
In the right-hand side of the last inequality, by continuity, the first probability is close to $0$ when $\varepsilon$ 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. \hfill \Box

\section{A.6. Stationarity test in nonlinear GARCH models}

\textbf{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 $\varepsilon_t^2 \to \infty$ a.s. at an exponential rate as $t \to \infty$. Moreover, it can be seen that the analogous of Lemma A.1-i) still holds.

Indeed, for all $\theta$ such as $\beta < e^{\Gamma}$, when $t \to \infty$, $|\sigma_t^2(\theta)/h_t - w_t(\alpha, \beta)| \to 0$ where, similar to \((A.5)\), \((w_t(\alpha, \beta))\) is defined as the stationary solution of
\[
(A.42) \quad b_0 n(\theta_{t-1}) w_t(\alpha, \beta) = \beta w_{t-1}(\alpha, \beta) + \alpha n_{t-1}^2.
\]
Moreover, for all $\theta$ 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 solution of
\[
(\alpha^*, \beta^*)^* = \arg \min_{(\alpha, \beta) \in \Theta_n, \beta} \mathbb{E} \{w_t^{-1}(\alpha, \beta) - 1 + \log w_t(\alpha, \beta)\}.
\]

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

It follows that
\[
\hat{\gamma}_n \overset{a.1}{=} \frac{1}{n} \sum_{t=1}^{n} \log \left( \hat{\alpha}_n \frac{\eta_t^2}{w_t(\hat{\alpha}_n, \hat{\beta}_n)} + \hat{\beta}_n \right) \overset{\mathbb{E}}{=} \frac{1}{n} \sum_{t=1}^{n} \log \left( \frac{\eta_0(n) w_{t+1}(\hat{\alpha}_n, \hat{\beta}_n)}{w_t(\hat{\alpha}_n, \hat{\beta}_n)} \right) \to \Gamma \text{ a.s. as } n \to \infty.
\]

For the previous convergence, we note that for some neighborhood $V^*$ of $(\alpha^*, \beta^*)$ we have $E \sup_{(\alpha, \beta) \in V^*} \mathbb{E} \log w_t(\alpha, \beta) < \infty$, which entails $\sup_{(\alpha, \beta) \in V^*} \mathbb{E} \log w_{n+1}(\alpha, \beta)/n \to 0$ 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 \textit{et al.} 2003).
The convergence of $\hat{\sigma}_u^2$ to

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

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

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

$$\Gamma^* = E\infty \log \left( \alpha^* \frac{\sigma^2}{\sigma^2(\theta^*)} + \beta^* \right) = E\infty \log \frac{\sigma^2_{t+1}(\theta^*) - \omega^*}{\sigma^2(\theta^*)} < 0.$$

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


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.
