We establish the necessary and sufficient conditions for covariance stationarity of ARCH($\infty$), for both the levels and the squares. The result applies to any form of the conditional variance coefficients. This includes GARCH($p,q$) and also specifications with hyperbolically decaying coefficients, such as the autoregressive coefficients of the autoregressive fractionally integrated moving average model. The covariance stationarity condition for the levels rules out long memory in the squares.

1. INTRODUCTION

Introduced by the seminal work of Engle (1982), the autoregressive conditional heteroskedastic (ARCH) models are certainly the most popular class of nonlinear time series models, in particular thanks to the development of the generalized autoregressive conditional heteroskedastic model of order $p,q$ (GARCH($p,q$)) by Bollerslev (1986) defined by

$$
\varepsilon_t = z_t \sigma_t, \quad t \in \mathbb{Z},
$$

$$
\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2, \quad \text{a.s.},
$$

where $\omega > 0$, $\beta_i \geq 0$, $\alpha_j \geq 0$ $(i = 1, \ldots, p, j = 1, \ldots, q)$ for integers $p \geq 0$ and $q > 0$ and a.s. means almost surely. The minimum conditions, which we assume hereafter, imposed on the rescaled innovation $z_t$ are i.i.d.-ness, $|z_t| < \infty$ a.s., and $z_t^2$ not degenerate. When $p = 0$, one gets the ARCH($q$).

Giraitis, Kokoska, and Leipus (2000) find sufficient conditions for strict stationarity of the ARCH($\infty$). This is the most general formulation of GARCH processes, generalizing (2) to

I thank Peter M. Robinson for useful comments on previous versions of the paper. Also, I am grateful to the co-editor (Bruce E. Hansen) and an anonymous referee whose suggestions greatly improved the paper. Address correspondence to: Paolo Zaffaroni, Servizio Studi, Banca d’Italia, Via Nazionale 91, 00184 Roma, Italy; e-mail: zaffaroni.paolo@insedia.interbusiness.it.
\[
\sigma_t^2 = \tau + \sum_{k=1}^{\infty} \psi_k \epsilon_{t-k}^2 \; \text{a.s.,} \quad \sum_{k=1}^{\infty} \psi_k < \infty,
\]

where \( \tau \geq 0 \) and \( \psi_k \geq 0 \). The ARCH(\( \infty \)) was introduced by Robinson (1991), who considered finite parameterizations of the \( \psi_k \) to build classes of alternatives in deriving score tests for no-ARCH. (See also the work of Hong, 1997, who proposes a one-sided test for no-ARCH that employs the ARCH(\( \infty \)) as a class of alternatives.)

The ARCH(\( \infty \)) representation of GARCH\((p,q)\) model (2) is obtained choosing exponentially decaying \( \psi_k \); e.g., the GARCH\((1,1)\) follows setting \( \tau = \omega/(1-\beta_1) \) and \( \psi_j = \alpha_1 \beta_1^{j-1} \); the ARCH\((q)\) model is obtained when \( \tau = \omega \) and \( \psi_j = 0 \) for \( j > q \). Although hyperbolic behavior of the ARCH coefficients \( \psi_j \) is allowed for, the strict stationarity condition of Giraitis et al. (2000),

\[
E(z_t^2) \sum_{k=1}^{\infty} \psi_k < 1,
\]

implies a bounded second moment of the \( \epsilon_t \). Indeed, for GARCH\((1,1)\) (4) works out to \( E(\beta_1 + \alpha_1 z_t^2) < 1 \), the well-known covariance stationarity condition for the \( \epsilon_t \) (see Bollerslev, 1986), which is more restrictive than the necessary and sufficient strict stationarity condition (see Nelson, 1990).

Exploiting the well-known linear ARMA\((m,p)\) (with \( m = \max[p,q] \)) representation of GARCH\((p,q)\), introduced by Bollerslev (1986), the autocovariance function (ACF) of the squares \( \epsilon_t^2 \) can be readily shown to be proportional to the ACF of an ARMA\((m,p)\) once the bounded fourth moment condition of the \( \epsilon_t \) is imposed. The critical step is precisely calculating this constant of proportionality, given by \( E[z_t^2 - E(z_t^2)]^2E(\sigma_t^4) \). For general GARCH\((p,q)\) the solution has been found independently by He and Teräsvirta (1999) and Karanasos (1999), extending the GARCH\((1,2)\) and GARCH\((2,1)\) cases analyzed by Bollerslev (1986). Noting that the previous results are either based on un-primitive assumptions or establish only necessity but not sufficiency, Ling (1999) and Ling and McAleer (2002) fill these gaps, providing the necessary and sufficient higher order moment conditions for GARCH\((p,q)\).

This paper shows how to evaluate \( E(\sigma_t^4) \) for ARCH(\( \infty \)) \( \sigma_t^2 \). It uses the linear representation of \( \epsilon_t^2 \) in a martingale difference sequence. The cases of both exponentially (such as GARCH\((p,q)\)) and hyperbolically decaying coefficients \( \psi_j \) are comprised. In particular, necessary and sufficient conditions for weak stationarity of the levels \( \epsilon_t \) and the squares \( \epsilon_t^2 \) are established. This is done in Section 2. Finally, we discuss in Section 3 the implications of the covariance stationarity conditions on the memory of the squares. It follows that covariance stationarity of the \( \epsilon_t \) precludes long memory in the \( \epsilon_t^2 \). This strengthens the result of Giraitis et al. (2000), who found that a sufficient condition for covariance stationarity of the \( \epsilon_t^2 \) rules out long memory. The proofs for the results are given in the Appendix.
2. STATIONARITY OF ARCH(∞)

The minimal condition for covariance stationarity of the \( \varepsilon_t \), is \( \kappa = E(\varepsilon_t^2) < \infty \), by (1). Following Robinson (1991), who considers the case \( \kappa = 1 \), setting \( \psi(L) := 1 - \kappa \sum_{j=1}^{\infty} \psi_j L^j \), we can rewrite (3) as

\[
\psi(L)\varepsilon_t^2 = \kappa \tau + \nu_t, \tag{5}
\]

setting \( \nu_t := \varepsilon_t^2 - \kappa \sigma_t^2 \). By (1) and the i.i.d.-ness of the \( z_t \), \( E(\nu_t|\mathcal{F}_{t-1}) = 0 \), where \( \mathcal{F}_t \) is the \( \sigma \)-field of events induced by the \( \varepsilon_s \) \( (s \leq t) \) (for the definition of conditional expectations when the corresponding unconditional expectations may not exist, see Loève, 1978, Sect. 27.2).

Assume that, for complex-valued \( z \), the following “invertibility” condition holds:

\[
\exists \delta(z) = \sum_{j=0}^{\infty} \delta_j z^j := \psi^{-1}(z), \quad \delta_0 = 1, \quad \text{s.t.} \sum_{j=0}^{\infty} \delta_j^2 < \infty. \tag{6}
\]

A sufficient condition for (6) is

\[
|\psi(z)| \neq 0, \quad |z| \leq 1,
\]

implying \( \psi(1) > 0 \), but we want to allow for the possibility that \( \psi(1) = 0 \) such as when the \( \psi_j \) are the AR(∞) coefficients of the autoregressive fractionally integrated moving average (ARFIMA) filter:

\[
\psi(L) = (1-L)^d \frac{a(L)}{b(L)}, \tag{7}
\]

\( L \) being the lag operator \( (L z_t = z_{t-1}) \) with \( d > 0 \) (by summability of the \( \psi_j \)) and where the finite-order polynomials \( a(L) \), \( b(L) \) have zeros outside the unit circle in the complex plane. Indeed, note that (6) does not imply absolute summability of the \( \delta_j \).

Given (6), (5) can be rewritten as

\[
\varepsilon_t^2 = \kappa \tau \delta(1) + \sum_{j=0}^{\infty} \delta_j \nu_{t-j}. \tag{8}
\]

These simple manipulations suggest that \( \tau, \kappa < \infty \), and \( \kappa \sum_{j=1}^{\infty} \psi_j < 1 \) imply covariance stationarity of the \( \varepsilon_t \). Indeed, it turns out that these are the necessary and sufficient conditions for \( E(\varepsilon_t^2) < \infty \). This is developed in Theorem 1, where the necessary and sufficient conditions for covariance stationarity of the squares \( \varepsilon_t^2 \) are also established.
THEOREM 1. Assume that condition (6) holds.

(i) The necessary and sufficient conditions for $E(\sigma_t^2) < \infty$ are

$$\tau < \infty,$$

$$\kappa < \infty,$$

$$\kappa \sum_{i=1}^{\infty} \psi_i < 1.$$  \hspace{1cm} (9, 10, 11)

Under these conditions

$$\phi := E(\varepsilon_t^2) = \frac{\kappa \tau}{1 - \kappa \sum_{i=1}^{\infty} \psi_i} < \infty.$$  \hspace{1cm} (12)

(ii) The necessary and sufficient conditions for $E(\sigma_t^4) < \infty$ are

$$\tau < \infty,$$

$$\theta := E(z_t^2 - \kappa)^2 < \infty,$$

$$\left( \theta \sum_{u=-\infty}^{\infty} \chi_\delta(u) \chi_\phi(u) \right) < 1,$$  \hspace{1cm} (13, 14)

setting $\psi_0 = 0$, $\psi_k = \psi_k (k \geq 1)$, and $\chi_c(u) := \sum_{k=0}^{\infty} c_k c_{k+u}$, $u = 0, \pm 1, \ldots$, for any square summable sequence $c_i$.

Under these conditions the $\varepsilon_t^2$ are covariance stationary with ACF

$\text{cov}(\varepsilon_t^2, \varepsilon_{t+u}^2) = E(\nu_t^2) \chi_\delta(u), \quad u = 0, \pm 1, \ldots,$

where

$$E(\nu_t^2) = \frac{\theta (\phi/\kappa)^2}{1 - \theta \sum_{u=-\infty}^{\infty} \chi_\delta(u) \chi_\phi(u)}^{-1} < \infty.$$  \hspace{1cm} (15)

Remark 1.1. The ACF of the squares $\varepsilon_t^2$ for GARCH($p,q$) can be obtained as follows. We first establish the ARCH($\infty$) representation of GARCH($p,q$). Exploiting the ARMA$(\max(p,q),p)$ representation of GARCH($p,q$), by standard arguments (see Nelson and Cao, 1992) one obtains

$$\frac{\alpha(L)}{1 - \beta(L)} = \alpha(L) \left( \frac{A_1}{1 - \rho_1 L} + \cdots + \frac{A_p}{1 - \rho_p L} \right),$$

where $\alpha(L) := \alpha_1 L + \cdots + \alpha_q L^q$, $\beta(L) := \beta_1 L + \cdots + \beta_p L^p$, $\rho_i (i = 1, \ldots, p)$ define the inverse of the roots of $|1 - \beta(z)| = 0$, assumed all outside the unit circle in the complex plane and distinct, and

$$A_i := \frac{1}{(1 - \rho_1 / \rho_i) \cdots (1 - \rho_{i-1} / \rho_i) (1 - \rho_{i+1} / \rho_i) \cdots (1 - \rho_p / \rho_i)}, \quad i = 1, \ldots, p,$$
where \( A_p = A_1 = 1 \) for \( p = 1 \). The \( \text{ARCH}(\infty) \) coefficients \( \psi_j \) of \( \text{GARCH}(p,q) \) are then

\[
\psi_j = \begin{cases} 
\alpha_1 \mu_{j-1} + \cdots + \alpha_j \mu_0, & j = 1, \ldots, q, \\
\alpha_1 \mu_{j-1} + \cdots + \alpha_q \mu_{j-q}, & j > q, 
\end{cases}
\]

\( \tau = \omega/(1 - \beta_1 - \cdots - \beta_p), \)

with \( \mu_j := A_1(\rho_1)^j + \cdots + A_p(\rho_p)^j, j \geq 0 \). Based on the previous definitions, Theorem 1 yields the ACF of \( \text{GARCH}(p,q) \text{~} \varepsilon_t^2 \text{~} \) with

\[
\phi = \kappa \omega/(1 - \kappa \alpha(1) - \beta(1)),
\]

deriving the \( \chi_\phi(u) \) and the \( \chi_\bar{\phi}(u) \), and thus \( E(\nu_t^2) \), based on (15) and deriving the \( \delta_i \) from (6).

Remark 1.2. Necessary and sufficient conditions for covariance stationarity of \( \text{GARCH}(p,q) \text{~} \varepsilon_t^2 \text{~} \) have been independently established by He and Teräsvirta (1999) and Karanasos (1999). When considering \( \text{GARCH}(p,q) \), these results and (12)−(14) are equivalent, but formal derivation of their equivalence is very cumbersome. He and Teräsvirta (1999) is closer to our result, because their work considers the slightly more general case of nonstandardized \( z_t \), though from a computational point of view our conditions (12)−(14) seem more appealing because they permit immediate use of known results for the theoretical ACF of \( \text{ARMA}(p,q) \).

As an example, for \( \text{GARCH}(1,1) \), setting \( \kappa = 1 \) and \( \pi_1 := \alpha_1 + \beta_1 \), i.e., the "persistence" parameter, we get

\[
\chi_\phi(0) = 1 + \frac{\alpha_1^2}{1 - \pi_1^2}, \quad \chi_\phi(u) = \alpha_1 \pi_1^{\lfloor u \rfloor - 1} \left( 1 + \frac{\alpha_1 \pi_1}{1 - \pi_1^2} \right),
\]

\[
\chi_\bar{\phi}(0) = \frac{\alpha_1^2}{1 - \beta_1^2}, \quad \chi_\bar{\phi}(u) = \frac{\alpha_1^2 \beta_1^{\lfloor u \rfloor}}{1 - \beta_1^2}, \quad u = \pm 1, \ldots.
\]

When the \( z_t \) have zero fourth-order cumulant (such as for Gaussian \( z_t \)) then \( \theta = 2 \). Under this condition, by simple manipulations, (14) yields the well-known covariance stationarity conditions for \( \varepsilon_t^2 \) (cf. Bollerslev, 1986, Sect. 3)

\[
3 \alpha_1^2 + \beta_1^2 + 2 \alpha_1 \beta_1 < 1,
\]

which, in turn, implies (11): \( E(\varepsilon_t^2) \alpha(1) + \beta(1) < 1 \).

For the \( \text{ARCH}(2) \), namely, \( \psi_k = 0, k \geq 3 \), setting \( (1 - \xi_1^{-1}L)(1 - \xi_2^{-1}L) := (1 - \psi_1 L - \psi_2 L^2) \) and assuming that the roots \( \xi_1, \xi_2 \) are greater than one in modulus (cf. (6)) and satisfy, for simplicity’s sake, \( \xi_1 \neq \xi_2 \), one gets
\[ \chi_\delta(u) = \frac{\xi_1^2 \xi_2^2}{(\xi_1 \xi_2 - 1)(\xi_2 - \xi_1)} \left( (\xi_1^2 - 1)^{-1} \xi_1^{1-\lfloor u \rfloor} - (\xi_2^2 - 1)^{-1} \xi_2^{1-\lfloor u \rfloor} \right), \]

\[ u = 0, \pm 1, \ldots, \quad \chi_\delta(0) = \psi_1^2 + \psi_2^2, \quad \chi_\delta(1) = \psi_1 \psi_2, \]

\[ \chi_\delta(u) = 0, \quad u = \pm 2, \pm 3, \ldots, \]

when \( \kappa = 1 \) (see Brockwell and Davis, 1987, Example 3.3.5). Using

\[ \psi_1 = \xi_1^{-1} + \xi_2^{-1}, \quad \psi_2 = -\xi_1^{-1} \xi_2^{-1}, \]

and by simple manipulations, from (14) one gets the covariance stationarity condition for ARCH(2) \( \epsilon_i^2 \) (cf. Milhøj, 1985, Theorem 3; He and Teräsvirta, 1999, equation (12)), setting \( \theta = 2 \),

\[ 3\psi_1^2 + 3\psi_2^2 + 3\psi_1^2 \psi_2 - 3\psi_2^3 + \psi_2 < 1. \]

Remark 1.3. When

\[ (E(z_i^4))^{1/2} \sum_{j=1}^{\infty} \psi_j < 1, \quad (17) \]

Giraitis et al. (2000, Theorem 2.1) show that \( E(\sigma_i^4) < \infty \). However, (17) is more restrictive than (14). For instance, for GARCH(1,1) with \( \theta = 2 \kappa = 2 \), their condition becomes

\[ 3^{1/2} \alpha_1 + \beta_1 < 1, \]

strictly implying (16), unless \( \beta_1 = 0 \), the ARCH(1).

Remark 1.4. Hyperbolically decaying specifications of the \( \psi_i \) and hence of the \( \delta_i \) are allowed for. Note, however, that from (11), \( \kappa < 1/(\sum_{i=1}^{\infty} \psi_i) \) and thus \( \kappa = 1 \) is ruled out when \( \sum_{i=1}^{\infty} \psi_i = 1 \). Thus, unless \( \kappa < 1 \), (7) is not compatible with covariance stationary \( \epsilon_j \).

Imposing \( \kappa = 1 \), a choice compatible with covariance stationary levels is obtained using the autoregressive coefficients of the ARFIMA filter as follows. Set (cf. (7))

\[ 1 - \sum_{j=1}^{\infty} \bar{\psi}_j L^j := (1 - L)^d \frac{a(L)}{b(L)}, \quad (18) \]

where \( d > 0 \) and \( a(L) \) and \( b(L) \) are finite-order polynomials, all of whose roots are outside the unit circle in the complex plane. Assuming that the nonnegativity constraints on the coefficients hold, i.e., \( \bar{\psi}_i \geq 0 \), set \( \psi_i = \bar{\psi}_i \) (\( i \geq 2 \)) and \( \psi_1 = \bar{\psi}_1 \epsilon \) for some given \( 0 < \epsilon < 1 \). Condition (14) is more involved, but again by suitable modification of the first \( n \) (say) coefficients \( \bar{\psi}_j \) (\( j = 1, \ldots, n \)), a feasible sequence of coefficients can be obtained from (18).
Remark 1.5. An alternative, frequency domain characterization of (14) is

$$\theta \int_{-\pi}^{\pi} f_{\delta}(\lambda) f_{\psi}(\lambda) \, d\lambda < 2\pi,$$

(19)

setting

$$f_{\delta}(\lambda) := |\delta(e^{i\lambda})|^2, \quad f_{\psi}(\lambda) := (1/\kappa^2)|1 - \psi(e^{i\lambda})|^2, \quad -\pi \leq \lambda < \pi,$$

where $\kappa^2 f_{\psi}(\lambda) = 1 + f_{\delta}^{-1}(\lambda) - 2\text{Re}(\delta^{-1}(e^{i\lambda})) (-\pi \leq \lambda < \pi)$ and $\text{Re}(\cdot)$ denotes the real part of its argument. This frequency domain specification seems easier to compute than (14). For instance, when considering parameterization (7) of the $\psi_{ij}$ (as for GARCH($p,q$)) this allows us to exploit the much greater computational simplicity of rational power spectra compared with their Fourier transforms. This might be relevant when imposing covariance stationarity of the $\epsilon^2_t$ in practical estimation, e.g., when estimating ARCH($\infty$) using the Whittle estimator (see Giraitis and Robinson, 2001).

The equivalent (19) representation of (14) could be used to show that (14) strictly implies (11). In fact, setting $\theta = 2\kappa = 2$ for simplicity’s sake, (19) can be rewritten as

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f_{\delta}(\lambda)(1 - 2\text{Re}(\delta^{-1}(e^{i\lambda}))) \, d\lambda < -\frac{1}{2}.$$ 

However, $f_{\delta}(\lambda) \geq 0 (-\pi \leq \lambda < \pi)$, and it is arbitrary, necessarily requiring

$$1 - 2\text{Re}(\delta^{-1}(e^{i\lambda})) = -1 + 2 \sum_{j=1}^{\infty} \psi_j \cos(j\lambda) < 0, \quad -\pi \leq \lambda < \pi.$$ 

Given the nonnegativity of the $\psi_j$ and $\cos(j\lambda) \leq 1$, with the equality achieved for $\lambda = 0 \pmod{2\pi}$, this is equivalent to

$$-1 + 2 \sum_{j=1}^{\infty} \psi_j < 0,$$

strictly implying $\psi(1) > 0$.

Finally, exploiting the relation between the $\psi_i$ ($i \geq 1$) and the $\delta_j$ ($j \geq 0$) (cf. (6)), (19) allows us to express the time domain characterization (14) more simply as

$$2\kappa \sum_{j=1}^{\infty} \psi_j \chi_{\delta}(j) < \frac{\kappa^2}{\theta} + \chi_{\delta}(0) - 1.$$
Remark 1.6. The original formulation of ARCH(\infty), analogous to Robinson (1991) but allowing \( \kappa \neq 1 \), is

\[
\sigma^2_t = \tilde{\tau} + \sum_{k=1}^{\infty} \psi_k (\epsilon^2_{t-k} - \kappa \tilde{\tau}), \quad \text{a.s.,}
\]

for some \( 0 < \tilde{\tau} < \infty \). The reparameterization (20) is clearly permitted only for covariance stationary \( \epsilon_t \), i.e., when \( \psi(1) > 0 \), given that \( \tilde{\tau} = \tau/\psi(1) \). In fact, assume that one starts directly from (20) rather than from (3). By the nonnegativity constraint \( \kappa \sum_{i=1}^{\infty} \psi_i \leq 1 \) because \( \psi(1) < 0 \) is not allowed. Imposing \( \psi(1) = 0 \) and assuming (6), the linear moving average representation (8) for the \( \epsilon^2_t \) would then be

\[
\epsilon^2_t = \zeta + \sum_{j=0}^{\infty} \delta_j \nu_{t-j},
\]

for any constant \( \zeta \), given that \( \psi(L) \epsilon^2_t = \psi(L) \epsilon^2_t - \psi(1) \zeta = \psi(L)(\epsilon^2_t - \zeta) \), which is meaningless.

Remark 1.7. (9)–(11) and (12)–(14) are the necessary and sufficient conditions for \( E(\epsilon^2_t) < \infty \) and for \( E(\epsilon^4_t) < \infty \), respectively.

3. MEMORY OF ARCH(\infty)

Giraitis et al. (2000, Proposition 3.1) show that (17) implies absolute summability of the ACF for the \( \epsilon^2_t \), ruling out long memory. However, considering that (17) is more restrictive than effectively required to obtain covariance stationary \( \epsilon^2_t \), it is important to assess the impact of the weaker condition (14) on the memory of the \( \epsilon^2_t \).

Insights can be obtained by looking at the linear representation (8) for \( \epsilon^2_t \). In fact, it follows that the memory of the \( \epsilon^2_t \) is expressed by the asymptotic behavior of \( \delta_j \) as \( j \to \infty \). Surprisingly, it turns out that even the much weaker covariance stationarity condition (11) for the levels \( \epsilon_t \) rules out long memory in the \( \epsilon^2_t \). In fact, from (6) and (11),

\[
\delta(1) = \sum_{j=0}^{\infty} \delta_j = 1 / \left( 1 - \kappa \sum_{j=1}^{\infty} \psi_j \right) < \infty.
\]

Expression (14) ensures that the uncorrelated \( \nu_t \) have finite variance but the rate of decay of the \( \delta_j \), imposed by (11), is already quick enough to imply their absolute summability. We summarize our results on the memory of the \( \epsilon^2_t \) as follows.
THEOREM 2. Assume that conditions (6) and (11) hold. Then
\[ \sum_{j=1}^{\infty} \delta_j < \infty, \tag{21} \]
where
\[ 0 \leq \delta_l = \kappa \psi_l + \sum_{s=2}^{\infty} \kappa^s \sum_{i_1=1}^{l-s+1} \cdots \sum_{i_{s-1}=1}^{l-i_1-\cdots-i_{s-2}-1} \psi_{i_1} \cdots \psi_{i_{s-1}} \psi_{l-i_1-\cdots-i_{s-1}}, \quad l \geq 1. \tag{22} \]

When the \( \psi_i \) decay toward zero more slowly than exponentially, namely, \( \psi_i / \xi^i \to \infty \) as \( i \to \infty \) for any \( 0 < \zeta < 1 \), (21) implies that, as \( u \to \infty \),
\[ \chi_\delta(u) \sim C \psi_u, \tag{23} \]
for some \( 0 < C < \infty \), with \( c(x) \sim d(x) \) as \( x \to x_0 \), meaning that \( c(x)/d(x) \to 1 \).

Remark 2.1. When \( \psi_i \sim ci^{-\delta} \) as \( i \to \infty \) for \( 0 < c, \delta - 1 < \infty \), as for the parameterization described in Remark 1.4, \( E(\epsilon_i^2) < \infty \) implies
\[ \chi_\delta(u) \sim C u^{-\delta}, \quad u \to \infty, \tag{24} \]
for some \( 0 < C < \infty \), ruling out long memory in the \( \epsilon_i^2 \). Under the same assumptions on the asymptotic behavior of the \( \psi_i \), the exact rate in (24) was also obtained in Giraitis et al. (2000, Proposition 3.2), although they impose (17), a sufficient condition for \( E(\epsilon_i^4) < \infty \).

Remark 2.2. Theorem 2 makes it clear that whereas the bounded second moment conditions (9)–(11) impart the degree of memory of the \( \epsilon_i^2 \), the stronger bounded fourth moment conditions (12)–(14) ensure that the martingale difference sequence \( \eta_i \) (the innovations in the linear representation (8) of the squares) are square integrable, without changing the memory implications of the model. This double role of the coefficients is simply a by-product of ARCH(\( \infty \)) nonlinearity.

Remark 2.3. In the case of exponentially decaying \( \psi_i \), i.e., when there does exist \( 0 < \zeta < 1 \) such that \( \psi_i / \xi^i \to c \) as \( i \to \infty \) for some \( 0 < c < \infty \), it clearly follows that
\[ \chi_\delta(u) \sim C \delta^u, \quad u \to \infty, \]
for some \( 0 < \delta < 1, 0 < C < \infty \), as in the GARCH(\( p,q \)) case.

REFERENCES


**APPENDIX**

**Proof of Theorem 1.** Part (i). Assume that \(E(\sigma_i^2) < \infty\). The necessary condition of (9)–(11) follows, taking expectation on both sides of (3). Sufficiency of (9)–(11) for \(E(\sigma_i^2) < \infty\) is proved in Giraitis et al. (2000, Theorem 2.1). Part (ii). Assume that \(E(\sigma_i^4) < \infty\). Then (12)–(14) follow, squaring both terms in (3) and taking expectations,

\[
E(\sigma_i^4) = (\phi/\kappa)^2 + \theta E(\sigma_i^4) \sum_{j_1,j_2=1}^{\infty} \psi_{j_1} \psi_{j_2} \left( \sum_{i=0}^{\infty} \delta_i \delta_{i+(j_1-j_2)} \right),
\]

using

\[
E[(\varepsilon_i^2 - \phi)(\varepsilon_{i+u}^2 - \phi)] = E(\nu_i^2) \sum_{j=0}^{\infty} \delta_j \delta_{j+u}, \quad u = 0, \pm 1, \ldots,
\]

as the \(\nu_i\) are square integrable martingale differences and \(\sum_{i=0}^{\infty} \delta_i^2 < \infty\) by (6). Given (12)–(14), \(E(\varepsilon_i^2) = \phi < \infty\) by part (i). Therefore \(\varepsilon_i\) is strictly stationary. Next, setting \(1_i := 1(\bigcap_{j=0}^{\infty} \{\sigma_{i-j}^2 < M\})\) for some constant \(M < \infty\) where \(1(A)\) equals one when the event \(A\) holds and zero otherwise, for any arbitrary sequence of constants \(s_j \geq 0, j = 1,2,\ldots,\)
\[ \sigma^2_{i}1_{i} \leq \tau 1_{i} + \sum_{j=1}^{\infty} \psi_{j} \epsilon_{i-j}^{2}1_{i-j}. \]

Writing \( \epsilon_{i-j}^{2} = \epsilon_{i-j}^{2} - \phi + \phi \), using (8), squaring and rearranging terms yields, as \( M \to \infty \),

\[ \sigma^4_{i}1_{i} \leq A_{i}^{2} + 2A_{i} \sum_{j=1}^{\infty} \psi_{j}(\epsilon_{i-j}^{2} - \phi)1_{i-j}. \]

By the law of iterated expectations, and collecting terms in \( \sim \) for some constant \( 0 \leq \zeta < \infty \), squaring and rearranging terms yields

\[ E(\sigma^4_{i}1_{i}) \leq E(A_{i}^{2})(1 - \theta \sum_{u=-\infty}^{\infty} \chi_{\delta}(u)\chi_{\phi}(u))^{-1} + o(1). \]

**Proof of Theorem 2.** From

\[ \left( 1 + \sum_{j=1}^{\infty} \delta_{j} L_{j} \right) \left( 1 - \kappa \sum_{i=1}^{\infty} \psi_{i} L_{i} \right) = 1, \]

by the fundamental theorem for polynomials and simple yet tedious calculations, (22) follows. We now establish that

\[ \delta_{j} = \kappa \psi_{j} + O \left( \left( \kappa \sum_{i=1}^{\infty} \psi_{i} \right)^{\xi} \right), \text{ as } j \to \infty, \]

for some constant \( 0 < \zeta < \infty \). By (11), for any \( M < \infty \)

\[ \kappa \sum_{j=1}^{M} \psi_{j} < 1, \quad \kappa \sum_{j=M+1}^{\infty} \psi_{j} < 1. \]
Consider a large enough $k$ such that $M < k$ for one of such $M$. Then, split the right-hand side of (22) as

$$
\delta_k = \kappa \psi_k + \sum_{l=2}^{[k/M]} \sum_{j_1=1}^{k-l+1} \cdots \sum_{j_{l-1}=1}^{k-j_{l-1}-j_{l-2}-1} B_{j_1,j_2,\ldots,j_{l-1}}(l,k)
$$

$$
+ \sum_{l=[k/M]+1}^{k} \sum_{j_1=1}^{k-l+1} \cdots \sum_{j_{l-1}=1}^{k-j_{l-1}-j_{l-2}-1} B_{j_1,j_2,\ldots,j_{l-1}}(l,k),
$$

(A.3)

where $[\cdot]$ is the integer part of its argument and

$$
B_{j_1,j_2,\ldots,j_{l-1}}(l,k) := \kappa^{l-1} \psi_{j_1} \psi_{j_2} \cdots \psi_{j_{l-1}} \psi_{k-j_1-\ldots-j_{l-2}}.
$$

Let us dispose of the first sum on the right-hand side of (A.3). As $l \leq [k/M]$, one obtains three terms:

$$
\sum_{l=2}^{[k/M]} \sum_{j_1=1}^{k-l+1} \cdots \sum_{j_{l-1}=1}^{k-j_{l-1}-j_{l-2}-1} B_{j_1,j_2,\ldots,j_{l-1}}(l,k)
$$

$$
= \sum_{l=1}^{[k/M]} \left( \sum_{j_1=1}^{M} \sum_{j_{l-1}=M+1}^{k-l+1} \cdots \sum_{j_{l-1}=M+1}^{k-j_{l-1}-j_{l-2}-1} B_{j_1,j_2,\ldots,j_{l-1}}(l,k) \right)
$$

$$
= A_1 + A_2 + A_3
$$

(A.5)

with

$$
A_1 := \sum_{l=2}^{[k/M]} \left\{ \sum_{j_1=1}^{M} \cdots \sum_{j_{l-1}=1}^{M} B_{j_1,j_2,\ldots,j_{l-1}}(l,k) \right\},
$$

(A.6)

$$
A_2 := \sum_{l=2}^{[k/M]} \left\{ \sum_{j_1=1}^{M} \cdots \sum_{j_{l-1}=1}^{M} B_{j_1,j_2,\ldots,j_{l-1}}(l,k) \right\},
$$

(A.7)

and the “mixed” term

$$
A_3 := \sum_{l=2}^{[k/M]} \sum_{s=1}^{l-1} \left\{ \sum_{j_1=M+1}^{k-l+1} \cdots \sum_{j_{s} = M+1}^{k-l+s-1} \sum_{j_{s+1} = M+1}^{k-l+s} \cdots \sum_{j_{l-1} = M+1}^{k-j_{l-1}-j_{l-2}} \sum_{j_{l} = 1}^{M} \cdots \sum_{j_{l-1} = 1}^{M} B_{j_1,j_2,\ldots,j_{l-1}}(l,k) \right\},
$$

(A.8)

where $\sum_{i_r} := \sum_{(r_1,\ldots,r_s) \subset \{1,\ldots,l-1\}}$, i.e., it selects all the groups of indexes of dimension $s$ (with $s = 1,\ldots,l-1$) drawn from a number $l$ of the and thus for any $s$ and any permutation $j_{i_1} + \cdots + j_{i_{l-s}} + j_{i_1} + \cdots + j_{r} = j_1 + \cdots + j_l$. Note that $\sum_{s=1}^{l-1} l = \binom{l}{s}$. 


For $A_1$, for some $0 < \zeta < 1$, writing $\sum_{l=2}^{[k/M]} = \sum_{l=2}^{[\zeta k/M]} + \sum_{l=[\zeta k/M]}^{[k/M]}$ yields, as $k \to \infty$,

$$A_1 = O\left(\psi_{\lfloor(1-\zeta)k\rfloor}X_1 + \left(\kappa \sum_{j=1}^{M} \psi_j\right)\left[\zeta k/M\right]\right)$$

$$= O\left(\psi_k + \left(\kappa \sum_{j=1}^{\infty} \psi_j\right)\left[\zeta k/M\right]\right)$$

since

$$\sum_{l=0}^{\infty} \left(\sum_{j_1=1}^{M} \cdots \sum_{j_{l-1}=1}^{M} \psi_{j_1} \psi_{j_2} \cdots \psi_{j_{l-1}} \kappa^l\right) \leq \frac{1}{1 - \kappa \sum_{j=1}^{\infty} \psi_j} := X_1 < \infty.$$  

Concerning $A_2$, along the same lines, as $k \to \infty$,

$$A_2 = O\left(\psi_{\lfloor(1-\zeta)k\rfloor} \sum_{l=0}^{\infty} \left(\sum_{j_1=M+1}^{\infty} \cdots \sum_{j_{l-1}=M+1}^{\infty} \kappa^l \psi_{j_1} \psi_{j_2} \cdots \psi_{j_{l-1}}\right) + \sum_{l=0}^{[k/M]} \left(\kappa \sum_{j=M+1}^{\infty} \psi_j\right)^l\right)$$

$$= O\left(\psi_k + \left(\kappa \sum_{j=1}^{\infty} \psi_j\right)\left[\zeta k/M\right]\right).$$

For the “mixed” term, by means of the same truncation, one obtains $A_3 = A_3' + A_3''$, where $A_3' = O(\psi_k X_2) = O(\psi_k)$ as $k \to \infty$, setting

$$X_2 := \sum_{l=0}^{\infty} \sum_{s=1}^{l-1} \sum_{l,s} \left(\sum_{j_1=M+1}^{\infty} \cdots \sum_{j_{l-1}=M+1}^{\infty} \sum_{j_1=1}^{M} \cdots \sum_{j_{l-1}=1}^{M} \psi_{j_1} \psi_{j_2} \cdots \psi_{j_{l-1}} \kappa^l\right)$$

$$= \sum_{l=0}^{\infty} \left(\kappa \sum_{j=1}^{\infty} \psi_j\right)^l < \infty.$$  

Likewise

$$A_3'' = O\left(\sum_{l=\lfloor\zeta k/M\rfloor}^{[k/M]} \sum_{s=1}^{l-1} \sum_{j_1=1}^{M} \sum_{j_{l-1}=1}^{M} \psi_{j_1} \psi_{j_2} \cdots \psi_{j_{l-1}} \sum_{j_1=1}^{M} \cdots \sum_{j_{l-1}=1}^{M} \sum_{j_1=1}^{M} \cdots \sum_{j_{l-1}=1}^{M} B_{j_1,j_2,\ldots,j_{l-1}}(l,k)\right)$$

$$= O\left(\left(\kappa \sum_{j=1}^{\infty} \psi_j\right)\left[\zeta k/M\right]\right) \text{ as } k \to \infty.$$  

For the second term on the right-hand side of (A.3), $l > [k/M]$ implies that for at least one summation, say, $\sum_{j_1=1}^{k-l+i-j_1-\cdots-j_{l-1}}$, then $k - l + i - j_1 - \cdots - j_{l-1} < M$. Thus, decomposition (A.5) is not feasible. However, the previous bounds developed for the case $[\zeta k/M] \leq l \leq [k/M]$ can still be applied. In fact
\[
\sum_{l=[k/M]+1}^{k} \sum_{j_1=1}^{k-l+1} \sum_{j_2=1}^{k-j_1-\cdots-j_{i-1}} B_{j_1,j_2,\ldots,j_{i-1}}(l, k) = O\left( \sum_{l=[k/M]+1}^{(M+1)k-l+1} \sum_{j_1=1}^{(M+1)k-j_1-\cdots-j_{i-1}} \kappa^j \psi_{j_1}\cdots\psi_{j_{i-1}} \right), \tag{A.9}
\]

given \( \max_{k \geq 1} \psi_k < \infty \) and \( \psi_k \geq 0 \) for any \( k \). Then, apply (A.5) to the expression on the right-hand side of (A.9) and proceed as indicated previously.

Condition (11) ensures summability of the \( \delta_j \). On the other hand when \( \kappa \sum_{i=1}^\infty \psi_i = 1 \), (A.2) is loose, suggesting a rate of decay of \( \delta_j \) slower than \( \psi_j \). For instance, for case (7) with \( \kappa = 1 \) and \( d > 0 \) it follows that \( \psi_j \sim c j^{-d-1} \) and \( \delta_j \sim c' j^{d-1} \) as \( j \to \infty \) for some \( 0 < c, c' < \infty \). Finally, note that when \( \psi_j \geq 0 \ (j \geq 1) \) then (22) implies \( \delta_j \geq 0 \ (j \geq 0) \).

Hence, for any \( u > 0 \) and some \( 0 < c < \infty \),
\[
\chi_{\delta}(u) = \sum_{j=0}^{u} \delta_j \delta_{j+u} + \sum_{j=u+1}^{\infty} \delta_j \delta_{j+u} \sim c \delta_u(1 + o(1)), \quad \text{as } u \to \infty,
\]
given that, for large enough \( u \),
\[
\sum_{j=u+1}^{\infty} \delta_j \delta_{j+u} \leq \delta_{u+n} \sum_{j=u+1}^{\infty} \delta_j = o(\delta_u),
\]
for some constant \( 1 \leq n < \infty \). Finally, comparing (A.2) with (22) and given \( \kappa > 0, \psi_i \geq 0 \ (i \geq 1) \), yields \( \delta_k \sim c \psi_k \) as \( k \to \infty \).