Title
The Variance of Sample Autocorrelations: Does Barlett's Formula Work With ARCH Data?

Authors
Kokoszka, Piotr S.
Politis, D N

Publication Date
2008-10-09

The variance of sample autocorrelations: does Bartlett's formula work with ARCH data?

Piotr S. Kokoszka* Dimitris N. Politis†

October 9, 2008

Abstract

We review the notion of linearity of time series, and show that ARCH or stochastic volatility (SV) processes are not only non-linear: they are not even weakly linear, i.e., they do not even possess a martingale representation. Consequently, the use of Bartlett's formula is unwarranted in the context of data typically modelled as ARCH or SV processes such as financial returns. More surprisingly, we show that even the squares of an ARCH or SV process are not weakly linear. Finally, we present an alternative to Bartlett's formula that is applicable (and consistent) in the context of financial returns data.

1 Introduction

In the theory and practice of time series analysis, an often used assumption is that a time series $\{X_t, t \in \mathbf{Z}\}$ of interest is *linear* [18], i.e., that

$$X_t = a + \sum_{i=-\infty}^{\infty} \alpha_i \xi_{t-i}, \quad \text{where } \xi_t \sim \text{ i.i.d. } (0,1), \tag{1}$$

*Department of Mathematics and Statistics, Utah State University, Logan, UT 84322-3900; email: `Piotr.Kokoszka@usu.edu`. Research partially supported by NSF grants DMS-0413653 and DMS-0804165.

†Department of Mathematics, University of California at San Diego, La Jolla, CA 92093-0112; email: `dpolitis@ucsd.edu`. Research partially supported by NSF grant DMS-07-06732

i.e., the ξ_ts are independent and identically distributed with mean zero and variance one; of course, in the above, the coefficients α_i must decay to zero fast enough as $|i| \to \infty$ so that the infinite sum converges in some fashion.

Recall that a linear time series $\{X_t\}$ is called *causal* if $\alpha_k = 0$ for $k < 0$, i.e., if

$$X_t = a + \sum_{i=0}^{\infty} \alpha_i \xi_{t-i}, \quad \text{where } \xi_t \sim \text{ i.i.d. } (0,1). \tag{2}$$

Eq. (2) should not be confused with the Wold decomposition that *all* purely nondeterministic time series possess [6]. In the Wold decomposition the 'error' series $\{\xi_t\}$ is only assumed to be a white noise, i.e., uncorrelated, and not i.i.d. A slightly weaker form of (2) amounts to relaxing the i.i.d. assumption on the errors to the assumption of a martingale difference, i.e., to assume that

$$X_t = a + \sum_{i=0}^{\infty} \alpha_i \nu_{t-i}, \qquad \sum_{i=0}^{\infty} \alpha_i^2 < \infty, \tag{3}$$

where $\{\nu_t\}$ is a stationary martingale difference adapted to $\mathcal{F}_t$, the σ-field generated by $\{X_s, s \le t\}$, i.e., that

$$E[\nu_t | \mathcal{F}_{t-1}] = 0 \text{ and } E[\nu_t^2 | \mathcal{F}_{t-1}] = 1 \text{ for all } t. \tag{4}$$

For conciseness, we will use here the term *weakly linear* for a time series $\{X_t, t \in \mathbf{Z}\}$ that satisfies (3) and (4).

Many asymptotic theorems in the literature have been proven under the assumption of linearity or weak linearity; see e.g. [6] [13] [18]. A central such result is the celebrated Bartlett's formula [2] for the asymptotic variance of the sample autocorrelations that has been shown to hold under weak linearity; see e.g. Chapter 8 of [1].

In the last thirty years, however, there has been a surge of research activity on *non*linear time series models; an early such example is the family of bilinear models [17]. Many of these nonlinear models were motivated from financial returns data as, e.g., the ARCH and GARCH models that were introduced in the 1980s [4] [11].

As far back as 1978, Granger and Andersen [17] warned against the use of Bartlett's formula in the context of bilinear time series. Despite additional warnings [23] [25], even to this day practitioners often give undue credence to the Bartlett $\pm 1.96/\sqrt{n}$ bands—that many software programs automatically overlay on the correlogram—even when a nonlinear time series model is entertained for the data!

As will be apparent from the main developements of this paper, ARCH processes are not only non-linear: they are not even weakly linear. Consequently, the use of Bartlett's formula is unwarranted in the context of data that are typically modelled as ARCH processes such as financial returns data. Perhaps more surprisingly, we show that even the squares of an ARCH process is not a weakly linear time series.

Example 1. To give a preview of our general results, consider a simple ARCH(1) process, i.e., suppose

$$X_t = \varepsilon_t \sqrt{\beta_0 + \beta_1 X_{t-1}^2} \tag{5}$$

where $\varepsilon_t \sim$ i.i.d. $(0,1)$. If $\beta_1 < 1$, then $EX_t^2 = \beta_0/(1-\beta_1)$. Denote $Y_t = X_t^2$ and let $\mathcal{F}_{t-1}$ be the σ–field generated by $Y_{t-1}, Y_{t-2}, \ldots$. Let $\sigma_t^2 = \beta_0 + \beta_1 L Y_t$ be the volatility function where L denotes the lag-operator, i.e., $LY_t = Y_{t-1}$. Since $\beta_1 L Y_t = \sigma_t^2 - \beta_0$, it follows that

$$(1-\beta_1 L)(Y_t - EY_t) = Y_t - \beta_1 L Y_t - (1-\beta_1) EY_t$$

$$= \sigma_t^2 \varepsilon_t^2 - (\sigma_t^2 - \beta_0) - (1-\beta_1)\frac{\beta_0}{1-\beta_1} = \sigma_t^2 \varepsilon_t^2 - \sigma_t^2.$$

Setting $\nu_t = \sigma_t^2(\varepsilon_t^2 - 1)$, the above calculation shows that

$$(1-\beta_1 L)(Y_t - EY_t) = \nu_t \tag{6}$$

and hence

$$Y_t = \frac{\beta_0}{1-\beta_1} + \sum_{i=0}^{\infty} \beta_1^i \nu_{t-i}. \tag{7}$$

In view of the fact that the innovations ν_t consitute a *white noise*—see e.g. [16], eq. (6) is simply the recursive equation of an AR(1) model with nonzero mean. In this light, eq. (7) is the usual MA representation of an AR(1) process thereby giving an allusion towards linearity.

Nevertheless, this allusion is false: linearity does not hold true here—not even in its weak form; this is a consequence of the fact that the innovations ν_t do not have a constant *conditional* variance as required in the martingale representation (3) and (4). To see this, just note that

$$E[\nu_t^2 | \mathcal{F}_{t-1}] = E[\sigma_t^4 (\varepsilon_t^2 - 1)^2 | \mathcal{F}_{t-1}] = \sigma_t^4 E[(\varepsilon_t^2 - 1)^2] = \sigma_t^4 \mathrm{Var}[\varepsilon_t^2].$$

The above simple example shows that the common intuition that the squares of an ARCH process are weakly linear is inaccurate. We will show in Section 2 that, under weak assumptions, neither the general ARCH(p) or ARCH(∞) nor stochastic volatility models are weakly linear, and that this negative result also extends to their squares. As a consequence, using Bartlett's formula on a correlogram of financial returns *or* their squares is unjustified[1] as made clear in Section 3. Last but not least, in Section 4, we present an alternative to Bartlett's formula that is indeed applicable (and consistent) in the context of ARCH models.

2 ARCH processes are not weakly linear

Consider a time series $\{X_t, t \in \mathbf{Z}\}$ that obeys the following model:

$$X_t = \sigma_t \varepsilon_t \quad \text{where } \varepsilon_t \sim \text{ i.i.d. } (0,1). \tag{8}$$

The above is a popular assumption with financial returns data as it captures the phenomenon of volatility clustering; two general models of interest can be put in this type of framework [26]:

- *ARCH*(∞) models where

  $$\sigma_t^2 = \beta_0 + \sum_{j=1}^{\infty} \beta_j X_{t-j}^2; \tag{9}$$

  this class includes all ARCH(p) and (invertible) GARCH(p,q) models.

- *Stochastic volatility* models where $L_t = \log \sigma_t$ satisfies the independent AR(p) equation

  $$L_t = \phi_0 + \sum_{j=1}^{p} \phi_j L_{t-j} + u_t \tag{10}$$

  where $u_t \sim$ i.i.d. $(0, \tau^2)$ and $\{u_t, t \in \mathbf{Z}\}$ is independent to $\{\varepsilon_t, t \in \mathbf{Z}\}$.

[1] In the context of bilinear series, Granger and Andersen [17] recommended using Bartlett's formula on the correlogram of the squared data; at the time this sounded like an insightful recommendation. Note, however, that a bilinear model of order one is tantamount to an ARCH(1) with $b = 0$ and $\beta_j = 0$ for $j > 1$ in eq. (9). Hence, in view of our general results, even with bilinear series using Bartlett's formula on the correlogram of the squared data is unjustified.

We introduce the following conditions:

(ia) For each t, σ_t is $\mathcal{F}_{t-1}$ measurable and square integrable.

(ib) The sequences $\{\sigma_t\}$ and $\{\varepsilon_t\}$ are independent, and σ_t is square integrable for each t.

(iia) There is t for which σ_t^2 is not equal to a constant.

(iib) There is t for which $E[\sigma_t^2|\mathcal{F}_{t-1}]$ is not equal to a constant.

(iii) For each t, ε_t is independent of $\mathcal{F}_{t-1}$, and it is square integrable with $E\varepsilon_t = 0, E\varepsilon_t^2 > 0$.

Definition 1 *We say that $\{X_t, t \in \mathbf{Z}\}$ is an ARCH-type process if (8) holds together with conditions (ia), (iia) and (iii).*

Definition 2 *We say that $\{X_t, t \in \mathbf{Z}\}$ is an SV-type process if (8) holds together with conditions (ib), (iib) and (iii).*

Proposition 1 *If $\{X_t\}$ is either an ARCH-type or an SV-type process, then it is not weakly linear.*

Proof. Suppose *ad absurdum* that $\{X_t\}$ is weakly linear. Since by (8) $EX_t = 0$, the constant term a in (3) must be zero, and the representation would have to be

$$X_t = \sum_{i=0}^{\infty} \alpha_i \nu_{t-i}. \tag{11}$$

The square summability of the α_j implies that

$$X_t = \lim_{m\to\infty} \sum_{i=1}^{m} \alpha_i \nu_{t-i}, \quad \text{in } L^2.$$

Since the conditional expectation of an L^2 random variable wrt a σ–field $\mathcal{F}$ coincides with the orthogonal projection on $L^2(\mathcal{F})$, and this projection is a continuous operator in L^2, we conclude that

$$E[X_t|\mathcal{F}_{t-1}] = \sum_{i=0}^{\infty} \alpha_i E[\nu_{t-i}|\mathcal{F}_{t-1}].$$

Since $E[\nu_t|\mathcal{F}_{t-1}] = 0$, we further obtain that

$$E[X_t|\mathcal{F}_{t-1}] = \sum_{i=1}^{\infty} \alpha_i \nu_{t-i}. \tag{12}$$

We will now show that for both ARCH–type and SV–type processes

$$E[X_t|\mathcal{F}_{t-1}] = 0. \tag{13}$$

If $\{X_t\}$ is ARCH–type, then

$$E[X_t|\mathcal{F}_{t-1}] = E[\sigma_t\varepsilon_t|\mathcal{F}_{t-1}] = \sigma_t E\varepsilon_t = 0.$$

Similarly, if $\{X_t\}$ is SV–type, then If $\{X_t\}$ is ARCH–type, then

$$E[X_t|\mathcal{F}_{t-1}] = E\{[\sigma_t\varepsilon_t|\sigma(\sigma_t, \mathcal{F}_{t-1})]|\mathcal{F}_{t-1}\} = E\varepsilon_t E[\sigma_t|\mathcal{F}_{t-1}] = 0.$$

By (11), (12) and (13), $X_t = \alpha_0\nu_t$, so by (4),

$$E[X_t^2|\mathcal{F}_{t-1}] = E[\alpha_0^2\nu_t^2|\mathcal{F}_{t-1}] = \alpha_0^2. \tag{14}$$

For an ARCH-type process, we obtain, on the other hand,

$$E[X_t^2|\mathcal{F}_{t-1}] = E[\sigma_t^2\varepsilon_t^2|\mathcal{F}_{t-1}] = \sigma_t^2 E\varepsilon_t^2. \tag{15}$$

Equations (14) and (15) imply that, for each t, $\sigma_t^2 E\varepsilon_t^2 = \alpha_0^2$. Since $E\varepsilon_t^2 > 0$, this contradicts assumption (iia) of Definition 1.

Similarly, for an SV-type process, $E[X_t^2|\mathcal{F}_{t-1}] = E\varepsilon_t^2 E[\sigma_t^2|\mathcal{F}_{t-1}]$, and we obtain a contradiction with condition (iib). $\diamond$

Proposition 1 covers both ARCH(∞) and stochastic volatility models mentioned above, and implies that these popular models for financial returns are not weakly linear. As a consequence, using Bartlett's formula is not justified under its auspices.

We next turn to the squares $Y_t = X_t^2$ of ARCH and SV processes. We denote

$$\mathcal{F}_{t-1}^X = \sigma\{X_{t-1}, X_{t-2}, \ldots\}, \quad \mathcal{F}_{t-1}^Y = \sigma\{Y_{t-1}, Y_{t-2}, \ldots\}.$$

Giraitis, Kokoszka and Leipus [14] showed that if (8) and (9) hold, and $(E\varepsilon_0^4)^{1/2}\sum_{j=1}^{\infty}\beta_j < \infty$, then the series Y_t admits the representation

$$Y_t = a + \sum_{i=0}^{\infty}\alpha_i\nu_{t-i} \tag{16}$$

in which the ν_t are martingale differences in the sense that $E[\nu_t|\mathcal{F}_{t-1}^Y] = 0$ and $E\nu_t^2 =: v^2$ is a finite constant. Nevertheless, the conditional variance

$$E[\nu_t^2|\mathcal{F}_{t-1}^Y] = \sigma_t^4 \mathrm{Var}[\varepsilon_0^2]$$

is *not* constant. Proposition 2 below shows that in general the squares of ARCH and SV processes are not weakly linear because they do not admit representation (16) whose innovations ν_t have nonzero *constant* conditional variance. As a consequence, Bartlett's formula cannot justifiably be used on the correlogram of squared returns.

For the purpose of Proposition 2, we now consider two cases that pose some required restrictions on the ARCH and SV processes considered.

ARCH case: Each σ_t^2 is $\mathcal{F}_{t-1}^Y$ measurable, ε_t^2 is independent of $\mathcal{F}_{t-1}^Y$, and $\{\sigma_t^2\}$ is not a.s. equal to a deterministic constant sequence.

SV case: The sequences $\{\sigma_t^2\}$ and $\{\varepsilon_t^2\}$ are independent, and the following two conditions hold:
(i) Y_t is not $\mathcal{F}_{t-1}^Y$ measurable;
(ii) the sequence

$$v_t^2 := E[\sigma_t^4|\mathcal{F}_{t-1}^Y] - (E[\sigma_t^2|\mathcal{F}_{t-1}^Y])^2$$

is not equal to an a.s. constant positive sequence (i.e. to a sequence such that $v_t^2 = v^2 > 0$ a.s. for each t).

Conditions (i) and (ii) automatically hold in the ARCH case. Indeed, if Y_t were $\mathcal{F}_{t-1}^Y$ measurable, then $\varepsilon_t^2 = \sigma_t^{-2} Y_t$ would be $\mathcal{F}_{t-1}^Y$ measurable, and so $E[\varepsilon_t^2|\mathcal{F}_{t-1}^Y] = \varepsilon_t^2$. Since, in the ARCH case, ε_t^2 is independent of $\mathcal{F}_{t-1}^Y$, we also have $E[\varepsilon_t^2|\mathcal{F}_{t-1}^Y] = E\varepsilon_t^2$, implying $\varepsilon_t^2 = E\varepsilon_t^2$. Thus, unless ε_t^2 is a.s. constant, candition (i) holds in the ARCH case. Condition (ii) holds in the ARCH case because the $\mathcal{F}_{t-1}^Y$ measurability of σ_t^2 implies that $v_t^2 = \sigma_t^4 - \sigma_t^4 = 0$.

Condition (i) practically always holds in the SV case because $Y_t = \sigma_t^2 \varepsilon_t^2$ need not be a function of $\sigma_{t-1}^2\varepsilon_{t-1}^2, \sigma_{t-2}^2\varepsilon_{t-2}^2, \ldots$. Because of condition (i), v_t^2 is in general a random variable, not a constant, so (ii) also practically always holds in the SV case.

Proposition 2 *Suppose $Y_t = X_t^2$, where X_t follows equation (8), and either ARCH or SV case holds. Then Y_t is not weakly linear.*

Proof. To lighten the notation denote $\mathcal{F}_{t-1} = \mathcal{F}^Y_{t-1}$ and suppose

$$Y_t = a + \sum_{i=0}^{\infty} \alpha_i \nu_{t-i}, \qquad \sum_{i=0}^{\infty} \alpha_i^2 < \infty \tag{17}$$

and (4) holds. Conditioning on $\mathcal{F}_{t-1}$, we obtain

$$E[Y_t|\mathcal{F}_{t-1}] = a + \sum_{i=1}^{\infty} \alpha_i \nu_{t-i}. \tag{18}$$

Subtracting (18) from (16), we thus obtain

$$Y_t - E[Y_t|\mathcal{F}_{t-1}] = \alpha_0 \nu_t.$$

If $\alpha_0 = 0$, then $Y_t = E[Y_t|\mathcal{F}_{t-1}]$ would be $\mathcal{F}_{t-1}$ measurable, which would contradict assumption (i). Thus for each t

$$\nu_t = \alpha_0^{-1}\{Y_t - E[Y_t|\mathcal{F}_{t-1}]\}. \tag{19}$$

The proof will thus be complete, if we show that the sequence $\{Y_t - E[Y_t|\mathcal{F}_{t-1}]\}$ does not have a constant positive conditional variance. For this purpose we introduce the following notation:

$$\xi_t = \varepsilon_t^2, \quad \lambda = E\xi_t, \qquad \rho_t = \sigma_t^2.$$

In the ARCH case,

$$E[Y_t|\mathcal{F}_{t-1}] = E[\xi_t\rho_t|\mathcal{F}_{t-1}] = \lambda\rho_t,$$

and so

$$E\left\{(Y_t - E[Y_t|\mathcal{F}_{t-1}])^2|\mathcal{F}_{t-1}\right\} = E[(\xi_t\rho_t - \lambda\rho_t)^2|\mathcal{F}_{t-1}] = \rho_t^2 \mathrm{Var}[\xi_o],$$

which is not a constant sequence.

In the SV case,

$$E[Y_t|\mathcal{F}_{t-1}] = E[\xi_t\rho_t|\mathcal{F}_{t-1}] = E[E[\xi_t\rho_t|\sigma\{\rho_t, \mathcal{F}_{t-1}\}]|\mathcal{F}_{t-1}]$$

$$= \lambda E[\rho_t|\mathcal{F}_{t-1}].$$

Therefore

$$E\left\{(Y_t - E[Y_t|\mathcal{F}_{t-1}])^2|\mathcal{F}_{t-1}\right\} = E\left\{(\xi_t\rho_t - \lambda E[\rho_t|\mathcal{F}_{t-1}])^2|\mathcal{F}_{t-1}\right\}$$

$$= E\{\xi_t^2\rho_t^2|\mathcal{F}_{t-1}\} - 2\lambda E\{\xi_t\rho_t E[\rho_t|\mathcal{F}_{t-1}]|\mathcal{F}_{t-1}\} + \lambda^2 \left\{(E[\rho_t|\mathcal{F}_{t-1}])^2|\mathcal{F}_{t-1}\right\}$$

$$= \lambda^2 E\{\rho_t^2|\mathcal{F}_{t-1}\} - 2\lambda^2(E[\rho_t|\mathcal{F}_{t-1}])^2 + \lambda^2(E[\rho_t|\mathcal{F}_{t-1}])^2,$$

and so we obtain

$$E[\nu_t^2|\mathcal{F}_{t-1}] = \alpha_0^{-2}\lambda^2 \left\{E\{\rho_t^2|\mathcal{F}_{t-1}\} - (E[\rho_t|\mathcal{F}_{t-1}])^2\right\}. \tag{20}$$

By condition (ii), the ν_t are not martingale differences with constant conditional variance, and the proof is complete. $\diamond$

3 Bartlett's formula does not work with ARCH data

For a second order stationary sequence $\{X_t\}$, define the population and sample autocovariances at lag k by

$$R_k = \mathrm{Cov}(X_1, X_{1+k}), \qquad \hat{R}_k = n^{-1}\sum_{i=1}^{n-k}(X_i - \bar{X}_n)(X_{i+k} - \bar{X}_n),$$

where $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$, and the corresponding autocorrelations

$$\rho_k = R_0^{-1}R_k, \qquad \hat{\rho}_k = \hat{R}_0^{-1}\hat{R}_k.$$

Define also the p–dimensional vectors

$$\boldsymbol{\rho} = [\rho_1, \rho_2, \ldots, \rho_p]^T, \qquad \hat{\boldsymbol{\rho}} = [\hat{\rho}_1, \hat{\rho}_2, \ldots, \hat{\rho}_p]^T.$$

If $\{X_t\}$ is a linear process (1) with iid innovations ξ_t having finite fourth moment, then

$$\sqrt{n}(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}) \xrightarrow{d} N(\mathbf{0}, \mathbf{W}) \tag{21}$$

as $n \to \infty$. The entries of the $p \times p$ matrix $\mathbf{W}$ are given by the celebrated Bartlett's formula:

$$w_{ij} = \sum_{k=-\infty}^{\infty}\left[\rho_{k+i}\rho_{k+j} + \rho_{k-i}\rho_{k+j} + 2\rho_i\rho_j\rho_k^2 - 2\rho_i\rho_k\rho_{k+j} - 2\rho_j\rho_k\rho_{k+i}\right]$$

see e.g. formula (47) in Section 8.4.5. of [1].

However, as we have seen in Section 2, the ARCH and SV processes typically used to model financial returns are not even weakly linear, so Bartlett's formula cannot be expected to hold with such data. To illustrate, note that for all white noise data, i.e., when $\rho_k = 0$ for all $k \geq 1$, Bartlett's formula implies that $\mathrm{var}(\sqrt{n}\hat{\rho}_1) \to 1$. Now ARCH processes are uncorrelated so they do satisfy the white noise condition but the limiting variance of $\sqrt{n}\hat{\rho}_1$ is generally greater than unity as the following simple example shows.

Example 1 (continued). Suppose $\{X_t\}$ is the ARCH(1) process (5), and assume for simplicity that the ε_t are standard normal. Then, if $3\beta_1 < 1$, $\{X_t\}$ is a strictly stationary sequence with finite $(4+\delta)$th moment, and

$$EX_1^2 = \frac{\beta_0}{1-\beta_1}, \qquad EX_1^4 = \frac{3\beta_0^2(1+\beta_1)}{(1-\beta_1)(1-3\beta_1^2)}, \tag{22}$$

see Section 3 of [3]. By Theorem 2.1 of [25], $\sqrt{n}\hat{\rho}_1 \xrightarrow{d} N(0,\tau^2)$, where

$$\tau^2 = R_0^{-2}\left[\mathrm{Var}(X_1X_2) + 2\sum_{i=1}^{\infty}\mathrm{Cov}(X_1X_2,\ X_{1+i}X_{2+i})\right]. \tag{23}$$

The mixing condition in Theorem 2.1 of [25] holds because ARCH(1) processes are even β–mixing with exponential rate, see e.g. [7].

By (22) we have $R_0 = \beta_0/(1-\beta_1)$. The calculation of $\mathrm{Var}(X_1X_2)$ is also straightforward:

$$\begin{aligned}
\mathrm{Var}(X_1X_2) &= E[X_1^2X_2^2] = E[X_1^2(\beta_0 + \beta_1\sigma_1^2\varepsilon_1^2)\varepsilon_2^2] \\
&= E[X_1^2(\beta_0 + \beta_1\sigma_1^2\varepsilon_1^2)] = \beta_0 E[X_1^2] + \beta_1 E[X_1^4] \\
&= \beta_0\frac{\beta_0}{1-\beta_1} + \beta_1\frac{3\beta_0^2(1+\beta_1)}{(1-\beta_1)(1-3\beta_1^2)} = \frac{\beta_0^2(1+3\beta_1)}{(1-\beta_1)(1-3\beta_1^2)}.
\end{aligned}$$

Since

$$\mathrm{Cov}(X_1X_2,\ X_{1+i}X_{2+i}) = 0, \quad i \geq 1, \tag{24}$$

we obtain

$$\tau^2 = \frac{(1-\beta_1)(1+3\beta_1)}{1-3\beta_1^2}. \tag{25}$$

One can check that τ^2 increases monotonically from 1 to ∞, as β_1 increases from 0 to $1/\sqrt{3}$; in particular, $\tau^2 \to \infty$ as $EX_1^4 \to \infty$.

For general ARCH(∞) processes (9), expressing τ^2 in terms of the coefficients β_j in closed form appears difficult and is not necessarily useful. The main properties established in the example above do however carry over to the general case, as the following proposition shows.

Proposition 3 *Suppose $\{X_t\}$ is the ARCH process of Definition 1 with exponentially decaying α–mixing coefficients, and finite $(4+\delta)$th moment. Then, convergence (21) holds with*

$$w_{ij} = \delta_{ij} \frac{E[X_1^2 X_{1+i}^2]}{(EX_1^2)^2}, \tag{26}$$

where δ_{ij} is the Kronecker delta.

Moreover, if $\{X_t\}$ admits representation (9), then

(i) $w_{ii} \geq 1$, and $w_{ii} > 1$ if $\mathrm{var}[\varepsilon_1^2] > 1$,

(ii) if $\beta_i > 0$, then $w_{ii} \to \infty$ as $EX_1^4 \to \infty$.

Proof. Since (24) holds for any ARCH–type process, formula (26) follows directly from Theorems 3.1 and 3.2 of [25].

We now prove the statements for the $\{X_t\}$ admitting representation (9).

(i) Observe that

$$E[X_1^2 X_{1+i}^2] - (EX_1^2)^2 = \mathrm{Cov}(X_1^2, X_{1+i}^2).$$

By Lemma 2.1 of [14], $\mathrm{Cov}(X_1^2, X_{1+i}^2) \geq 0$, so $w_{ii} \geq 1$. By formula (2.11) of [14], $w_{ii} > 1$ if $\mathrm{var}[\varepsilon_1^2] > 1$.

(ii) Set $\lambda = E\varepsilon_1^2$, and note that

$$E[X_1^2 X_{1+i}^2] = \lambda E[X_1^2(\beta_0 + \sum_{j=1}^{\infty} \beta_j X_{1+i-j}^2)]$$

$$= \lambda\beta_0 EX_1^2 + \lambda \sum_{j=1}^{\infty} \beta_j E[X_1^2 X_{1+i-j}^2] \geq \lambda\beta_i E[X_1^4].$$

Thus,

$$w_{ii} \geq \lambda\beta_i \frac{EX_1^4}{(EX_1^2)^2}$$

and the proof is complete. $\diamond$

Note that all ARCH and GARCH models used in practice have exponentially decaying α–mixing coefficients, see Sections 3 and 4 of [7]. Convergence (21) for GARCH processes with finite fourth moment also follows from Theorem 1 of [28] because for them the so called physical dependence measure decays exponentially fast, see Section 5 of [27]. The results of [25] are used to find the exact form the asymptotic covariance matrix.

Proposition 4 *Suppose* $\{X_t\}$ *is the SV process of Definition 2 with exponentially decaying* α*–mixing coefficients, and finite* $(4+\delta)$*th moment. Then, convergence (21) and formula (26) also holds.*

Proof. Follows by direct application of Theorems 3.1 and 3.2 of [25]. $\diamond$

Note that if an SV process is of the form $X_t = \exp(L_t)\varepsilon_t$ with L_t defined by (10), and the errors u_t have a density, then $\{L_t\}$ is α–mixing with exponential rate, see Section 6 of [5]. Since multiplying by an iid sequence ε_t does not affect mixing, $\{X_t\}$ is then also α–mixing with exponential rate.

4 An alternative to Bartlett's formula for financial returns data

Propositions 3 and 4 suggest a simple method-of-moments estimator of the asymptotic variance of $\sqrt{n}\hat{\rho}_i$ for ARCH and/or SV processes, namely

$$\hat{w}_{ii} = \frac{(n-i)^{-1}\sum_{d=1}^{n-i} X_d^2 X_{d+i}^2}{n^{-1}(\sum_{d=1}^{n} X_d^2)^2}. \tag{27}$$

Equation (27) is our proposed alternative to Bartlett's formula in the case of data that are either ARCH or SV processes with finite fourth moments. For ease of reference, we state the consistency properties of the proposed estimator as our final proposition.

Proposition 5 *(i) If* $\{X_t\}$ *is strictly stationary and ergodic with* $EX_t^4 < \infty$*, then, for any* $i \geq 1$*,* $\hat{w}_{ii} \xrightarrow{a.s.} w_{ii}$ *where* w_{ii} *was defined in eq. (26).*
(ii) Under the conditions of Proposition 3 or those of Proposition 4, $\hat{w}_{ii}$ *is an a.s.–consistent estimator of the variance of the asymptotic distribution of* $\sqrt{n}\hat{\rho}_i$*.*

Proof. Part (i) follows from the ergodic theorem for stationary sequences; see e.g. Theorem 9.6 of [19]. Ergodicity is a very weak property that is implied by any form of mixing [5]; hence part (ii) is immediate. $\diamond$

Note that if $EX_t^4 = \infty$, formulas (26) and (27) are no longer useful. Nevertheless, if $EX_t^2 < \infty$, the sample autocorrelations are still consistent; see e.g. the work of Davis and Mikosch [8] [9] [10] from which it follows that an ARCH process (with normal errors) is typically in the domain of attraction of an α stable law. If $\alpha \in (2, 4)$, then

$$n^{1-2/\alpha} L(n)(\hat{\rho}_k - \rho_k) \xrightarrow{d} S_k \tag{28}$$

where S_k has an $\alpha/2$ stable law, L is a slowly varying function, and $\rho_k = 0$ if $k \neq 0$ by the ARCH equation. Despite the infinite variance of $\hat{\rho}_k$ in this case, it is still possible to construct confidence intervals for ρ_k but the focus should instead be on estimating the quantiles of the limit distribution of $\hat{\rho}_k$. Due to eq. (28), subsampling [24] can be successfully used in this respect leading to robust confidence intervals and tests for ρ_k that remain valid whether EX_t^4 is finite or not. To construct the subsampling estimator, an estimate of α appearing in (28) must be used; the latter could be obtained by Hill's estimator or any other consistent method; see e.g. [20] [21] [22].

Finally, note that when applied to the squares of ARCH and SV processes, Theorems 3.1 and 3.2 of [25] do not lead to any simple expressions because the squares have nonvanishing correlations at any lags, and an analog of (24) no longer holds. To approximate the distribution of $\hat{\rho}_k$ in such a case, it is recommended to use a resampling/subsampling approach such as the ones discussed in [23] [24] [25]. These approaches can be totally nonparametric, e.g. blocking methods, or semiparametric, e.g. based on residuals from an assumed model.

References

[1] Anderson, T. W. (1971). *The Statistical Analysis of Time Series,* John Wiley, New York.

[2] Bartlett, M.S. (1946). On the theoretical specification and sampling properties of autocorrelated time series, *J. Roy. Statist. Soc. Supplement*, vol. 8, pp. 27-41. [Correction: 1948, vol. 10, p. 200].

[3] Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307-327.

[4] Bollerslev, T., Chou, R. and Kroner, K. (1992). ARCH modeling in finance: a review of theory and empirical evidence, *J. Econometrics*, 52, 5-60.

[5] Bradley, R. C. (1986). Basic properties of strong mixing conditions, in *Dependence in Probability and Statistics*, 165–192, E. Eberlein and M. S. Taqqu, eds. Birkhäuser.

[6] Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods, 2nd ed.* Springer, New York.

[7] Carasco, M and Chen, X. (2002) Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory*, 18, 17–39.

[8] Davis, R. A., and Mikosch, T. (1998). The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Ann. Statist.*, 26, 2049-2080.

[9] Davis, R. A., and Mikosch, T. (2000). The sample autocorrelations of financial time series models. In *Nonlinear and nonstationary signal processing*, 247–274, Cambridge Univ. Press, Cambridge, 2000

[10] Davis, R. A., and Mikosch, T. (2001). Point process convergence of stochastic volatility processes with application to sample autocorrelation. *J. Appl. Probab.* 38A, 93–104.

[11] Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation, *Econometrica*, 50, 987-1008.

[12] Fan, J. and Yao, Q. (2003). *Nonlinear Time Series*, Springer, New York.

[13] Fuller, W.A. (1996). *Introduction to Statistical Time Series, 2nd Ed.*, John Wiley, New York.

[14] Giraitis, L., Kokoszka, P., and Leipus, R. (2000). Stationary ARCH models: dependence structure and Central Limit Theorem, *Econometric Theory*, 16, pp. 3-22.

[15] Giraitis, L., Robinson, P., and Surgailis, D. (2000) A model for long memory conditional heteroskedasticity *The Annals of Applied Probability*, 10, pp. 1002–1024.

[16] Gouriéroux, C. (1997). *ARCH Models and Financial Applications*, Springer, New York.

[17] Granger, C. and Andersen, A. (1978). *An Introduction to Bilinear Time Series Models*, Vandenhoeck and Ruprecht, Göttingen.

[18] Hannan, E.J. and Deistler, M. (1988). *The Statistical Theory of Linear Systems*, John Wiley, New York.

[19] Kallenberg, O. (1997) *Foundations of Modern Probability*, Springer, New York.

[20] McElroy, T. and Politis, D. N. (2007). Moment-Based Tail Index Estimation, *J. Statist. Plan. Infer.*, vol. 137, no. 4, pp. 1389-1406.

[21] McElroy, T. and Politis, D. N. (2007). Computer-intensive rate estimation, diverging statistics, and scanning, *Annals Statist.*, vol. 35, no. 4, pp. 1827-1848.

[22] McElroy, T. and Politis, D. N. (2008). A fine-tuned estimator of a general convergence rate, *Australian/New Zealand J. Statistics*, vol. 50, no. 1, pp. 1-12.

[23] Politis, D. N. (2003). The impact of bootstrap methods on time series analysis, *Statistical Science*, Vol. 18, No. 2, 219-230.

[24] Politis, D. N., Romano, J.P. and Wolf, M. (1999). *Subsampling*, Springer, New York.

[25] Romano, J.P. and Thombs, L. (1996). Inference for autocorrelations under weak assumptions, *J. Amer. Statist. Assoc.*, 91, 590-600.

[26] Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. in *Time Series Models in Econometrics, Finance and Other Fields*, D.R. Cox, David V. Hinkley and Ole E. Barndorff-Nielsen (eds.), London: Chapman & Hall, pp. 1-67.

[27] Shao, X. and Wu, W. B. (2007). Asymptotic spectral theory for nonlinear time series, *Annals of Statistics*, 35, 1773-1801.

[28] Wu, W. B. (2008). An asymptotic theory for sample covariances of Bernoulli shifts, *Stochastic Processes and their Applications*, Forthcoming.