BANCA D'ITALIA

Temi di discussione

del Servizio Studi

Contemporaneous aggregation of GARCH processes

by Paolo Zaffaroni

Number 449 - July 2002

CONTEMPORANEOUS AGGREGATION OF GARCH PROCESSES

by Paolo Zaffaroni*

Abstract

In this paper we study the effect of contemporaneous aggregation of heterogeneous GARCH processes as the cross-sectional size diverges to infinity. A complete statistical characterization of the limit aggregate is provided under general assumptions on the form and degree of heterogeneity of the micro GARCH processes. Implications on the memory and on modelling issues of the limit portfolios are also developed. The key features which characterize the results are the shape of the cross-sectional distribution of micro parameters, their degree of cross-sectional dependence and the degree of cross-sectional dependence of the rescaled innovations. These features provide a set of testable implications with respect to the relationship between the micro and aggregate statistical properties.

JEL classification: C32, C43.

Keywords: contemporaneous aggregation, GARCH, conditionally heteroskedastic newline factor models, common and idiosyncratic risk, nonlinearity, nonstationarity, memory.

Contents

* Banca d'Italia, Economic Research Department.

1 Introduction

The autoregressive conditional heteroskedasticity (ARCH) model of Engle (1982) and the generalized ARCH (GARCH) development of Bollerslev (1986) represent the most popular approaches used to describe the conditional heteroskedasticity observed in many financial time series, within the class of nonlinear time series models. The chief motivation underlying this popularity is represented by their excellent widespread performance in fitting the empirical distribution of financial asset returns, synthesized in a number of well known stylized facts, see Bollerslev, Chou, and Kroner (1992).

Given that a large number of securities are traded in financial markets, practical use of GARCH as models for asset returns has prompted the need for analyzing the effect of contemporaneous aggregation (henceforth aggregation) of GARCH, in the sense of summing or averaging across assets. Nijman and Sentana (1996) note that when GARCH models are fitted to returns on individual stocks or to returns on two exchange rates, this implies a precise parametric structure for the return of the portfolio of those stocks, and of the cross exchange rate respectively. In particular, Nijman and Sentana show that the sum of a finite number of GARCH belongs to the class of weak GARCH, defined by replacing the conditional expectation operator with the linear projection operator within the definition of GARCH (see Drost and Nijman (1993)). This has induced Meddahi and Renault (1996) to introduce a more general class of volatility models which include both GARCH and stochastic volatility (SV) models as particular cases, and establish the conditions which ensure closeness under both temporal and contemporaneous aggregation. These conditions are satisfied by weak GARCH but violated by (strong) GARCH.

In general, it turns out that the number of parameters of the exact aggregate model, based on n heterogeneous units, increases such as $O(n)$, except for particular cases of no-heterogeneity across parameters. For instance, the sum of two GARCH$(1,1)$, with a different persistence parameter, yields a weak GARCH$(2,2)$, echoing well-known results on aggregation of linear ARMA (Granger and Morris 1976). The same problem could arise when modelling the return for individual assets. For instance, Ding and Granger (1996) argue that a suitable volatility model should be viewed as the arithmetic average of several heterogenous GARCH components, each characterized by a different degree of persistence of their conditional variance. Second, according to certain models of speculative trading (see e.g. Tauchen and Pitts (1983)), observed asset prices represent the arithmetic average of traders reservation prices.

In this paper we propose a different approach, based on these considerations. Given the aggregate

$$X_{n,t} = \frac{1}{n}\sum_{i=1}^{n} x_{i,t}, \quad t \in \mathbb{Z}, \tag{1}$$

of n heterogeneous $x_{i,t}$, each parameterized as a GARCH, we characterize the asymptotic statistical properties of $X_{n,t}$ as $n \to \infty$. Therefore, this paper complements the finite n results of Nijman and Sentana (1996) and of Meddahi and Renault (1996), when focusing on GARCH.

The asymptotic limit as $n \to \infty$ (in a suitable norm) of the aggregate $X_{n,t}$, henceforth the limit aggregate, provides a valid approximation to the exact aggregate model. We establish the statistical properties of the limit aggregate, under various assumptions on the form and degree of heterogeneity of the $x_{i,t}$. For a sufficiently strong degree of cross-correlation between the $x_{i,t}$, the limit aggregate maintains the GARCH nonlinearity, uncorrelated levels and correlated squares, conveying the basic features of a volatility model. The coefficients of the limit aggregate have a semiparametric specification which could be parsimoniously parameterized and the model estimated by pseudo-maximum likelihood. As a by-product, a test for heterogeneity of the micro parameters is obtained.

The limit aggregate could be used for forecasting volatility and, possibly, employed as the basis for developing a model for pricing derivative securities with stochastic volatility. This would be empirically relevant given that the large majority of derivative contracts traded worldwide, such as options and futures on stock and interest rate indexes, is based on an aggregate as the underlying asset (such as e.g. the Standard & Poor's 500 index or the London Inter Bank Offer Rate). Alternatively, when considering single asset returns, both the heterogeneous volatility components and the equilibrium interpretations requires a large yet unknown n, and establishing the limit aggregate is instrumental when testing specific intra-day trading models (Tauchen and Pitts (1983) consider the case of homoskedasticity). Such equilibrium models are coupled with the so-called mixture of distribution hypothesis, whereby for the i-th asset the daily return is given by $\sum_{h=1}^{I} x_{i,t_h}$, summing the intra-day equilibrium return $x_{i,t}$. The sum is dictated by the mixing random variable (r.v.) I which expresses the number of news occurring during the day. It should be noticed, though, that the daily return so defined represents a temporal aggregation of the process $\{x_{i,t}\}$, at random instants, and cannot be formally interpreted as the outcome of contemporaneous aggregation, in the sense of (1).

The characterization of the asymptotic properties of $X_{n,t}$, as a by-product, permits to address other, relevant, issues that induce interest for aggregation of GARCH.

Since Ding, Granger, and Engle (1993), relatively recent empirical research shows that the effect of shocks to the conditional variance of asset returns is very persistent but is eventually absorbed as time passes, consistent with the notion of long memory when theoretical autocovariances are not summable. Several long memory volatility models have been proposed to account for this recent stylized fact of asset return dynamics (see, among others, Robinson (1991), Baillie, Bollerslev, and Mikkelsen (1996) and Robinson and Zaffaroni (1997) for ARCH-type models and Harvey (1998), Comte and Renault (1998), Breidt, Crato, and de Lima P. (1998) and Robinson and Zaffaroni (1998) for SV-type models) but its foundation is not yet well understood.

It is well known that long memory can be obtained by aggregation of heterogeneous ARMA processes, as noted in Robinson (1978) and developed in Granger (1980). The results available for ARMA are not readily applicable to a nonlinear time series framework except when ARMA or, more generally, linear models represent a good approximation. This does not apply to GARCH as they exhibit correlated squares and uncorrelated levels unlike linear processes. Indeed, in order to employ the linear aggregation results in a GARCH$(1,1)$ setting, Ding and Granger (1996) consider a recursive definition of the aggregate, different from $X_{n,t}$. They show that, under a Beta distributional assumption and a particular form of negative dependence between the micro GARCH$(1,1)$ parameters, the limit conditional variance is characterized by hyperbolically decaying coefficients, a necessary condition for long memory. Recently, Leipus and Viano (1999) study the asymptotic properties of the average of $x_{i,t}^2$, where the $x_{i,t}$ are ARCH(∞) (see Robinson (1991)) with random coefficients. They show that long memory is ruled out under weak assumptions and that for GARCH$(1,1)$ $x_{i,t}$ the limit of $1/n \sum_{i=1}^{n} x_{i,t}^2$ exhibits a summable yet hyperbolically decaying autocorrelation function (ACF) under a set of assumptions analogous to Ding and Granger (1996). We show that under no condition will the limit aggregate of GARCH exhibit long memory squares as this is ruled out by the conditions for covariance stationary levels of the limit aggregate. Moreover, except for a particular case of negative cross-sectional dependence between micro parameters (examples of which are represented by Ding and Granger (1996) and Leipus and Viano (1999)), which allows for hyperbolically decaying ACF, the limit of $X_{n,t}^2$ will be characterized by an approximately exponentially decaying ACF (imposing covariance stationarity).

Since the development of the arbitrage pricing theory (see Ross (1976)), common shocks, represented by latent common factors, have played a key role in asset pricing theory when facing a large number of assets, but there is little doubt that idiosyncratic shocks are an important determinant of assets dynamics. Suitable assumptions, typically expressed by uniform boundedness of the eigenvalues of the idiosyncratic variance-covariance matrix (see Chamberlain and Rothschild (1983)), however, allow neglecting idiosyncratic shocks in arbitrarily large portfolios. This is a crucial prerequisite for any method of estimation of the latent factors. GARCH factor models have been increasingly popular in empirical finance since King, Sentana, and Wadhwani (1994). It is crucial to understand the conditions that ensure portfolio full diversification of idiosyncratic-driven risk when GARCH are used to parameterize the idiosyncratic component of each asset so that Chamberlain and Rothschild's (1983) results could be applied. Assuming that $X_{n,t}$ represents the aggregate of the idiosyncratic component of the factor structure, it turns out that the bounded eigenvalue condition could be relaxed although bounded variance and independence of the $x_{i,t}$ will not guarantee that $X_{n,t}$ converges to its unconditional expectation in mean square. Formally, we show the fastest possible rate at which the maximum eigenvalue could diverge to infinity, consistent with the notion of factor structure. On the other hand, for certain shapes of the cross-sectional distribution of the micro parameters, still consistent with covariance stationary $x_{i,t}$, mean square convergence of $X_{n,t}$ fails, in disagreement with the definition of factor structure.

The plan of the paper is as follows. In section 2 we focus on aggregation of ARCH(1). Definitions and assumptions are introduced in section 2.1. Sections 2.2 and 2.3 focus on independent and common rescaled innovations respectively. Section 3 presents a number of generalizations, including aggregation of GARCH(p, q), and discusses at lengths the memory implications. Section 4 describes the implications of these results for dynamic GARCH factor models. Concluding remarks are in section 5. All results are formally stated in theorems and the proofs reported in appendix.

2 Aggregation of ARCH(1)

In this section we focus on ARCH(1) micro heterogeneous units $x_{i,t}$, when both the parameters and the rescaled innovations are potentially varying across units

$$x_{i,t} = z_{i,t}\,\sigma_{i,t}, \;\; i \in \mathbb{N},\, t \in \mathbb{Z}, \tag{2}$$

with

$$\sigma^2_{i,t} = \omega_i + \alpha_i x^2_{i,t-1}, \;\; a.s., \tag{3}$$

where a.s. means almost surely. We assume that the α_i and the ω_i are independent and identically distributed ($i.i.d.$) draws from a bivariate distribution with support included in $\mathbb{R}^2_+$. It should be noted that α_i and ω_j are in general cross-sectionally dependent for i=j. The case of no-heterogeneity across parameters, e.g. $\alpha_i = \alpha$, represents a simple particular case of our setting and will not be discussed.

Assuming random coefficients represents a particularly suitable way to describe heterogeneity across an arbitrary large number units. Despite the interpretation of the $x_{i,t}$ as returns for assets with random pay-off, note that random coefficients do not represent an additional source of risk. In fact, ω_i and α_i are once and for all drawn when asset i is originally priced, and could be consistently estimated using a span of data of the return for the i-th asset.

2.1 Definitions and assumptions

A key ingredient of the impact of aggregation is represented by the type of cross-sectional dependence of the rescaled innovation $z_{i,t}$. We will consider two cases. In one the $z_{i,t}$ are perfectly independent across units, expressing a source of time-varying heterogeneity, and, in the other they are perfectly correlated across units. For the sake of clarity, we will write $z_{i,t} = \epsilon_{i,t}$ and $z_{i,t} = u_t$ in the two cases. We assume throughout this paper that $\epsilon_{i,t}$ and u_t, called respectively the idiosyncratic the common rescaled innovation, satisfy the following assumption I without stating this explicitly.

Assumption I

(i) The u_t are i.i.d. across $t \in \mathbb{Z}$ and the $\epsilon_{i,t}$ are i.i.d. across $t \in \mathbb{Z}$ and $i \in \mathbb{N}$, satisfying $E(u_t) = E(\epsilon_{i,t}) = 0$, $0 < E(\mid u_t \mid^r) = E(\mid \epsilon_{i,t} \mid^r) = \mu_r < \infty$, $(r = 1, 2, 4)$ and $\mu_0 = E log\, \epsilon^2_{i,t} = E log\, u^2_t$ is well defined.

(ii) The $\{u_t,\, \epsilon_{i,t}\}$ and the $\{\omega_i,\, \alpha_i\}$ are mutually independent.

Remarks.
(a) We will mostly focus on the case when $\mu_2 = 1$ and $\mu_4 = 3$ which are the usual normalizations considered for estimation of GARCH. Indeed, these normalizations will be made for simplicity's sake and can be relaxed in most circumstances without harm. Also, generalizing I to allow for heterogeneity across moments of the idiosyncratic innovation such as $E(|\epsilon_{i,t}|^r) = \mu_r^{(i)}$, has no substantial impact.
(b) The *i.i.d.*-ness assumption implies that the micro processes are strong GARCH (Drost and Nijman 1993, Definition 1).

Henceforth $\sim$ denotes asymptotic equivalence: $a(x) \sim b(x)$, as $x \to x_0$, when $a(x)/b(x) \to 1$, and c, C bounded positive constants (not always the same). Given real γ, η with $0 \leq \gamma, \eta < \infty$, we assume that the ARCH(1) coefficients satisfy the following.

Assumption II(γ)
(i) The ω_i are i.i.d. with $0 < \omega_i$ almost surely and $E(\omega_i^2) < \infty$ for any $i \in \mathbb{N}$.
(ii) The α_i are i.i.d. with absolutely continuous distribution in the interval $[0, \gamma)$, depending upon the real parameter b_γ ($b_\gamma > -1$), whose density, denoted by $B(\cdot; b_\gamma)$ behaves

$$B(\alpha_i; b_\gamma) \sim C\,(\gamma - \alpha_i)^{b_\gamma}, \ \ \alpha_i \to \gamma^-. \tag{4}$$

Remarks.
(a) Part *(ii)* describes a mild semiparametric specification of the density function of the α_i, only imposing its behaviour in a neighbourhood of γ. The constraint $b_\gamma > -1$ is the obvious integrability condition. Indeed, it can be alternatively expressed as

$$B(\alpha_i; b_\gamma) = D(\alpha_i)(\gamma - \alpha_i)^{b_\gamma},$$

for any integrable function $D(\cdot)$ defined on $[0, \gamma]$ with $D(\alpha) \sim C$, as $\alpha \to \gamma^-$. An extremely wide variety of parametric specifications $B(\cdot; \theta)$ for $\theta \in \Theta \subset R^p$ is allowed for by (4).
(b) (4) could be generalized to

$$B(\alpha_i; b_\gamma) \sim C\,(\gamma - \alpha_i)^{b_\gamma} L(\frac{1}{\gamma - \alpha_i}), \tag{5}$$

where $L(\cdot)$ denotes a slowly varying function, defined by $L(tx)/L(t) \to 1$ as $t \to \infty$ for any $x > 0$ (see Zygmund (1977)). Qualitatively, considering

(5) has no consequence on the results and thus, for simplicity's sake, we will assume throughout this paper that $L(\cdot) = 1$.
(c) Nelson (1990) shows that the probabilistic properties of GARCH$(1,1)$ crucially depend on whether (for an arbitrary unit i) ω_i is greater or equal to zero. (i) rules out the possibility that $\sigma^2_{i,t} = 0$ almost surely for all units satisfying $\omega_i = 0$.
(d) We are focusing, by and large, on the case $\gamma \leq 1$, implying covariance stationary $x_{i,t}$. This is much stronger than considering strictly stationary $x_{i,t}$, defined by $\alpha_i \, e^{Elog(z^2_{i,t})} < 1$ (Nelson 1990, Theorem 2), which reduces to $\alpha_i < 2\,e^E$ for Gaussian $z_{i,t}$ ($E = 0.577...$ is the Euler constant). We will comment only briefly on the possibility of $\alpha_i \geq 1$. In fact, this might induce explosive behaviour for the limit aggregate, e.g. in terms of conditional moments, hiding relevant implications of the aggregation mechanism.

Assumption III(η)
The ω_i and α_i satisfy

$$\omega_i = \tilde{\omega}_i \mid 1 - \mu_2 \alpha_i \mid^{\eta}, \tag{6}$$

for an i.i.d. sequence $\tilde{\omega}_i$, mutually independent from the α_i, with $0 < \tilde{\omega}_i$ almost surely and $E(\tilde{\omega}_i^2) < \infty$ for any $i \in \mathbb{N}$.

Remarks.
(a) Case $\eta = 0$ implies that the ω_i and the α_i are mutually independent. When $\eta > 0$, we impose negative cross-sectional dependence between the ω_i and the α_i. The case of (local) positive dependence as well as of negative dependence, with the ω_i having support bounded away from zero, will not be discussed, as the same results obtained for case $\eta = 0$ apply.
(b) Assumption $III\,(\eta)$ could be substantially weakened to

$$E(\omega_i \mid \alpha_i) \sim c \; |1 - \mu_2 \alpha_i|^{\eta}, \;\; \alpha_i \to \gamma^-,$$

without specifying the degree of dependence when α_i is well below γ, but we consider (6) for simplicity's sake.
(c) Case $\eta = 1$ represents a typical reparameterization used when considering univariate ARCH models for estimation. It implies, when $\mu_2 \alpha_i < 1$, that $E_n(x^2_{i,t})$ is independent from the α_i. However, in a multivariate framework such as ours, it has strong implications for the aggregation mechanism as, in general, does the degree of cross-sectional dependence across parameters.

We will denote, for clarity's sake, the aggregate (1) by ${}^{E}X_{n,t}$ in the idiosyncratic case ($z_{i,t} = \epsilon_{i,t}$) and by ${}^{U}X_{n,t}$ in the common case ($z_{i,t} = u_t$).

Note that no distinction needs to be made between stock and flow variables unlike the temporal aggregation case, see Drost and Nijman (1993).

Some of the results of this paper (such as the behaviour of the conditional variance of $X_{n,t}$) are stated as exact rates of convergence or divergence as n, the cross-sectional dimension, grows. Therefore, we do not approximate sums (averages) with integrals but formally establish the rate at which the statistic of interest grows with n. This implies that the results are not linked to the definition of the aggregate as an arithmetic mean (cf. definition of $X_{n,t}$) but they apply to the sum $\sum_{i=1}^{n} x_{i,t}$ or to any other affine transformation such as $c_n + d_n X_{n,t}$ for given deterministic sequences $\{c_n, d_n, n \in \mathbb{N}\}$. Dealing with exact rates is often a necessity as, say, the unconditional variance of $X_{n,t}$ could be unbounded, although $X_{n,t}$ converges to zero (in a suitable norm). Other results (e.g. asymptotic distribution) necessarily requires an appropriate normalization (such as $d_n = \sqrt{n}$). These results would still provide a useful approximation when different normalizations are imposed by the problem at hand.

From now on, we will denote the conditional expectation and conditional variance operators, given the GARCH coefficients, by $E_n(\cdot)$ and $\mathrm{var}_n(\cdot)$ respectively. $\to_p$, $\to_r$ and $\to_d$ denote convergence in probability, in rth mean, and convergence in the sense of the finite-dimensional distribution, respectively.

2.2 Idiosyncratic innovations

Within this section, we assume that $\mu_2 = 1$, $\mu_4 = 3$ and $\gamma \leq 1$. The aggregate

$$^{E}X_{n,t} = \frac{1}{n}\sum_{i=1}^{n} \epsilon_{i,t}\sigma_{i,t},$$

is given by a sum of purely idiosyncratic components with $\sigma_{i,t}^2$ given in (3). Simple calculation yields

$$\mathrm{var}_n(^{E}X_{n,t}) = \frac{1}{n^2}\sum_{i=1}^{n} \frac{\omega_i}{1-\alpha_i}. \tag{7}$$

Theorem 1 *Assume $II(\gamma)$ and $III(\eta)$. As $n \to \infty$:*
Set $\eta = 0$.

(i) When $\gamma < 1$, uniformly in b_γ,

$$\mathrm{var}_n({}^E X_{n,t}) \to 0, \quad a.s.$$

(ii) When $\gamma = 1$,
if $b_1 > 0$, then a positive constant C exists such that a.s.

$$\mathrm{var}_n({}^E X_{n,t}) \sim C n^{-1};$$

if $b_1 = 0$, then a positive constant C exists such that a.s.

$$\mathrm{var}_n({}^E X_{n,t}) \sim C n^{-1} \log n;$$

if $b_1 < 0$, setting $\delta = b_1 + 1$, a.s.

$$\mathrm{var}_n({}^E X_{n,t}) \sim n^{\frac{1}{\delta}-2}\, S_\delta,$$

*with $S_\delta > 0\,(0 < \delta < 1)$ a.s., where $S_\delta\,(0 < \delta \leq 2)$ denotes a δ-stable r.v. with zero location parameter *.*

When $\eta = 1$, (i) applies for any $\gamma \leq 1$ and any b_γ.

Remarks.
(a) Let us consider the case of independent ω_i and α_i ($\eta = 0$). When $b_1 > -1/2$, one gets the usual result by which ${}^E X_{n,t}$ converges to zero in mean-square as n tends to infinity. However, for $b_1 < -1/2$ we obtain the rather striking result that the variance of ${}^E X_{n,t}$ tends to infinity at rate $n^{\frac{1}{\delta}-2}$ a.s., violating the usual result on the vanishing importance of idiosyncratic risk at the aggregate level.
(b) When $-1/2 < b_1 \leq 0$, the unconditional second moment of the aggregate is infinite even though ${}^E X_{n,t}$ goes to zero in mean-square. Therefore, misleading information would have been obtained when looking at unconditional moments rather than considering exact rates of conditional moments.
(c) When $\alpha_i = \alpha$ for any i, that is, when one allows only for time-varying heterogeneity through the $\epsilon_{i,t}$, it easily follows that ${}^E X_{n,t} \to 0$ in mean-square for any $\gamma \leq 1$.

*Using Samorodnitsky and Taqqu (1994) notation, S_δ refers to $S_\delta(\sigma, \beta, 0)$ for real parameters $\sigma \geq 0$ (scale parameter) and $-1 \leq \beta \leq 1$ (skewness parameter). We leave the values for σ, β unspecified and make them explicit only when needed.

(d) When $\gamma > 1$, including the case of micro IGARCH(1), $\text{var}_n({}^E X_{n,t})$ is unbounded and Theorem 1 does not apply. These cases are briefly discussed in section 3.
(e) When $\eta = 1$, ${}^E X_{n,t}$ converges to zero in mean-square as n tends to infinity for any γ and any shape of the distribution of α_i. We could easily extend the results to the case of an arbitrary non-negative η (see Lemma 1 in the appendix) but, for clarity's sake, we focus only on the case of η equal to zero and one.

The ${}^E X_{n,t}$ are martingale differences for any n whereas, due to the ARCH nonlinearity, the squared aggregate process

$$ {}^E Y_{n,t} = ({}^E X_{n,t})^2 = \frac{1}{n^2} \sum_{i,j=1}^{n} \epsilon_{i,t} \epsilon_{j,t} \sigma_{i,t} \sigma_{j,t}, $$

will exhibit non trivial memory properties. As shown by Engle (1982), bounded fourth moment for the micro ARCH(1) processes requires $3\,\alpha_i^2$ be smaller than 1. However, when the α_i are bounded below $1/\sqrt{3}$, the ${}^E X_{n,t}$ converge to zero in mean-square for any value of b_γ by Theorem 1. Thus, by Slutsky's theorem, the ${}^E Y_{n,t}$ converge to zero in probability, implying that looking at the memory properties of the squared aggregate is irrelevant. We summarize this result as follows.

Corollary *The covariance stationarity condition for ${}^E Y_{n,t}$ ($\gamma \leq 1/\sqrt{3}$) implies that*

$$ {}^E Y_{n,t} \rightarrow_p 0, $$

as $n \rightarrow \infty$, uniformly in b_γ.

To investigate the effect of relaxing $\gamma \leq 1/\sqrt{3}$, we study the asymptotic distribution of the $X_{n,t}$, using the suitable normalization suggested by Theorem 1. Given the possibility of asymptotic nonstationarity, one can also look at the behaviour of the aggregate of the truncated processes, obtained setting $\epsilon_{i,s} = 0\,(i = 1, ..., n)$ for all $s \leq 0$:

$$ \tilde{x}_{i,t} = \epsilon_{i,t} \tilde{\sigma}_{i,t}, $$

with

$$ \tilde{\sigma}_{i,t}^2 = \omega_i \left(\sum_{k=0}^{t-1} \alpha_i^k \prod_{j=1}^{k} \epsilon_{i,t-j}^2 \right), \tag{8} $$

yielding

$$^E\tilde{X}_{n,t} = \frac{1}{n}\sum_{i=1}^{n}\tilde{x}_{i,t}.$$

(8) is equivalent to the conditional model of Nelson (1990, eq. (6)) with the initial distribution of $\tilde{\sigma}^2_{i,0}$ equal to the Dirac mass at zero. Let us set its conditional variance equal to

$$\tilde{V}_{t,n} = var_n(^E\tilde{X}_{n,t}) = \frac{1}{n^2}\sum_{i=1}^{n}\omega_i\frac{1-\alpha_i^t}{1-\alpha_i}.$$

Theorem 2 *Assume $II(\gamma)$ and $III(\eta)$. As $n \to \infty$:*
Set $\eta = 0$.
(i) When $\gamma < 1$, uniformly in b_γ, or when $\gamma = 1$, with $b_1 > 0$,

$$\sqrt{n}\ \ ^E X_{n,t} \to_d S_2(t),$$

where the $S_2(t)$ are uncorrelated and distributed like a normal r.v. $N(0,V)$ with $V = E(\omega/(1-\alpha))$.
(ii) When $\gamma = 1$, with $b_1 < 0$,

$$\sqrt{n}\ \ ^E\tilde{X}_{n,t} \to_d \tilde{S}_2(t), \tag{9}$$

where the $\tilde{S}_2(t)$ are uncorrelated and distributed like a normal r.v. $N(0,V_t)$ with $V_t \sim c\,t^{-b_1}$ as $t \to \infty$.
Assume further $\mu_0 < 0$ and $E[\max_{k\geq 1}\prod_{s=1}^{k}\epsilon^2_{i,t-s}] < \infty$ for any $t \in \mathbb{Z}$, $i \in \mathbb{N}$. Then

$$n^{1-\frac{1}{\delta}}\ \ ^E X_{n,t} \to_d S_\delta(t), \tag{10}$$

setting $\delta = 2(b_1 + 1)$, where the $S_\delta(t)$ are distributed like a δ-stable r.v. $(0 < \delta < 2)$.

When $\eta = 1$, (i) applies for any $\gamma \leq 1$ and any b_γ.

Remarks.
(a) This result has sound implications. When the micro processes are mutually independent, the (suitably normalized) aggregate will converge to a δ-stable process, Gaussian in the stationary case ($b_1 > 0$). Hence, the ARCH parametric structure characterizing the micro processes is lost through aggregation as the limit aggregate is not a volatility model. This is caused

simply by the combination of (2) which imposes uncorrelatedness, and the independence of the $\epsilon_{i,t}$ (and of the α_i) which permits the standard central limit theorem (henceforth CLT) for *i.i.d.* variates. Indeed, the result still holds also when the $\epsilon_{i,t}$ or the α_i are cross-sectionally correlated, as long as the degree of dependence is not too strong so that the CLT applies.

The limit process satisfies $ES_\delta(t)\, S_\delta(v) = 0$ for any $t \neq v$ and any $0 < \delta \leq 2$, although note that $E(S_\delta(t))^2$ is unbounded for $\delta < 2$. When $\delta = 2$, this implies independence of $S_2(t)$ and $S_2(v)$ when $t \neq v$ whereas this is not guaranteed when $\delta < 2$.

(b) In the nonstationary case ($b_1 < 0$), we have illustrated the asymptotic distribution of both the non-truncated ${}^{E}X_{n,t}$ and truncated aggregate ${}^{E}\tilde{X}_{n,t}$ as n goes to infinity. Interestingly, these results can be viewed as sequential limits of the truncated aggregate ${}^{E}\tilde{X}_{n,t}$ normalized by $\sqrt{\tilde{V}}_{t,n}$, as t and then n go to infinity and vice versa. See Phillips and Moon (1999, Appendix B(1)) for the probability arguments necessary for sequential asymptotics. Theorem 2 shows that only in the stationary case will the limit distribution and the rate of convergence be the same. On the contrary, when $b_1 < 0$, both the rate and the asymptotic distribution will depend on the order at which n and t go to infinity.

In fact, consider first the case where t and then n go to infinity. ${}^{E}\tilde{X}_{n,t}$ converges in distribution to ${}^{E}X_{n,t}$ and $\tilde{V}_{t,n}$ converges a.s. to (7) as $t \to \infty$. Therefore, letting first t go to infinity, one obtains precisely the left hand side of (10), the non-truncated normalized aggregate, except for the random denominator, of order $n^{\frac{1}{\delta}-1}$, by Theorem 1. Finally, by letting $n \to \infty$, the limit distribution will be $S_\delta(t)/\sqrt{S_{\frac{\delta}{2}}}$ when $b_1 < 0$ (recall that the $S_{\frac{\delta}{2}}$, defined in Theorem 1, are positive a.s., unlike $S_\delta(t)$), combining (10) and Theorem 1.

Let us now consider the other type of sequential limit, letting first n and then t go to infinity. When $b_1 < 0$, the first limit yields, in distribution, $\tilde{S}_2(t)/\sqrt{V_t}$, equal to the right hand side of (9) but with $\sqrt{V_t}$ in the denominator. Then, writing $\tilde{S}_2([rt])/\sqrt{V_t}$, for any $0 \leq r \leq 1$, and letting $t \to \infty$, one easily obtains a sequence of r.vs, calling them $\bar{S}_2(r)$ $(0 \leq r \leq 1)$, normally distributed $N(0, r^{-b_1})$ and mutually independent for any $r \neq r'$.

Phillips and Moon (1999), in a general multi-index framework, discuss the relation between sequential and joint limit and establish conditions under which they give equivalent results. The stationary case, $b_1 > 0$, represents a particular example of those results whereas the nonstationary case, $b_1 < 0$, represents a situation where their conditions do not apply and, in fact, the

outcome of sequential asymptotics will depend on the order at which n and t go to infinity. (Taqqu, Willinger, and Sherman (1997) provides another example where the equivalence between sequential and joint limits fail.)
(c) There is a strong analogy of Theorem 2 with certain results of temporal aggregation of GARCH. When the x_t are univariate ARCH(1), given by (2) and (3) for non-random $\omega_i = \omega$, $\alpha_i = \alpha$ and setting $\epsilon_{i,t} = \epsilon_t$ (assume they satisfy I with $\mu_2 = 1$), Diebold (1988) shows that $\sum_{t=1}^{T} x_t/\sqrt{T} \to_d N(0, \omega/(1-\alpha))$, as $T \to \infty$, when $\alpha < 1$. This result could be generalized to the case of $\alpha \geq 1$, exploiting recent findings on the strong mixing property and regular variation of the distribution of x_t (see Davis and Mikosch (1998)). Let δ satisfy the equation $E(\alpha\, u_t^2)^{\frac{\delta}{2}} = 1$ (see Davis and Mikosch (1998, Table 1)), yielding $\delta \geq 2$ when $\alpha \leq 1$ and $0 < \delta < 2$ when $1 < \alpha < exp(-Elog\epsilon_t^2)$. Then, setting $\delta' = min[\delta, 2]$, $\sum_{t=1}^{T} x_t/T^{\frac{1}{\delta'}}$ converges in distribution to a δ'-stable r.v., as $T \to \infty$ (Diebold (1988) results is re-obtained for $\delta' = 2$).

2.3 Common innovations

In this section the aggregate will be denoted by

$$^{U}X_{n,t} = \frac{u_t}{n} \sum_{i=1}^{n} \sigma_{i,t}.$$

Due to the dependence between $\sigma_{i,t}$ and $\sigma_{j,t}$, induced by the u_t,

$$\mathrm{var}_n(^{U}X_{n,t}) = \frac{1}{n^2} \sum_{i,j=1}^{n} E_n(\sigma_{i,t}\sigma_{j,t}), \tag{11}$$

whose behaviour is described as follows. We assume $\mu_2 = 1$, $\mu_4 = 3$ and $\gamma \leq 1$ unless we specify differently.

Theorem 3 *Assume $II(\gamma)$ and $III(\eta)$. As $n \to \infty$:*
Set $\eta = 0$.
(i) When $\gamma < 1$, uniformly in b_γ, for some constant C,

$$\mathrm{var}_n(^{U}X_{n,t}) \to C, \quad a.s.$$

(ii) When $\gamma = 1$,
if $b_1 > -1/2$, a.s.

$$\mathrm{var}_n(^{U}X_{n,t}) \sim C;$$

if $b_1 = -1/2$, *then positive constants* c, C *exist such that a.s.*

$$c(\log n) \leq \mathrm{var}_n({}^U X_{n,t}) \leq C\,(\log n)^2;$$

if $b_1 < -1/2$, *setting* $\delta = -(b_1+1)/b_1$, *a.s.*

$$\mathrm{var}_n({}^U X_{n,t}) \sim n^{\frac{1}{\delta}-1}\, S_\delta,$$

with $S_\delta > 0\,(0 < \delta < 1)$ *a.s.*

When $\eta = 1$, *(i) applies for any* $\gamma \leq 1$ *and any* b_γ.

Remarks.
(a)The variance of the ${}^U X_{n,t}$ is always bounded away from zero for any values of b_γ. However, when $b_1 < -1/2$, the variance explodes asymptotically, at exactly the same rate as for the variance of the ${}^E X_{n,t}$, equal to $n^{-\frac{2b_1+1}{b_1+1}}$.
(b) The asymptotic distribution of the aggregate ${}^U X_{n,t}$ is not degenerate even in the stationary case. This suggests looking at the asymptotic behaviour of the square aggregate ${}^U Y_{n,t} = ({}^U X_{n,t})^2$. Recall that for $\gamma \leq 1/\sqrt{3}$ each micro ARCH(1) has bounded fourth moment. By using arguments similar to Theorem 3, it turns out that the the limit of ${}^U X_{n,t}$ has bounded fourth moment when $\gamma < 1/\sqrt{3}$ or $\gamma = 1/\sqrt{3}$, $b_\gamma > -3/4$. In contrast, the limit aggregate exhibits unbounded kurtosis when $\gamma = 1/\sqrt{3}$ with $b_\gamma < -3/4$. Thus, the distribution of the limit aggregate could exhibit fatter tails than the ones of the distribution of the micro ARCH(1) processes.

We now characterize the asymptotic distribution of the ${}^U X_{n,t}$. Set, for any real k,

$$\phi_k = E[|\,(1-\alpha_i)\,|^{\frac{\eta}{2}}\ \alpha_i^k],\ \ \rho_k = E[\tilde{\omega}_i^k].$$

Let us relax the assumption that $\mu_2 = 1$, $\mu_4 = 3$ and $\gamma \leq 1$.

Theorem 4 *For any* $n \in \mathbb{N}$, *there exist processes* $\{\underline{X}_{n,t}, \overline{X}_{n,t}, t \in \mathbb{Z}\}$ *such that*

$$\min[\underline{X}_{n,t}, \overline{X}_{n,t}] \leq {}^U X_{n,t} \leq \max[\underline{X}_{n,t}, \overline{X}_{n,t}],\ a.s., \tag{12}$$

satisfying the following.

(i) Assume II (γ) *and* $III(\eta)$ *and set* $\delta = b_\gamma + \eta/2$.
As $n \to \infty$, *when* $max[\gamma^{\frac{1}{2}}\mu_1, \gamma\,\mu_2] < 1$ *or* $\max[\gamma^{\frac{1}{2}}\mu_1, \gamma\,\mu_2] = 1$, $b_\gamma + \eta > 1$,

$$\underline{X}_{n,t} \to_1 \underline{X}_t = u_t\,\rho_{\frac{1}{2}} \left(\sum_{k=0}^{\infty} \phi_{\frac{k}{2}}^2 \prod_{j=1}^{k} u_{t-j}^2 \right)^{\frac{1}{2}} \tag{13}$$

$$\overline{X}_{n,t} \to_1 \overline{X}_t = u_t\,\rho_{\frac{1}{2}} \left(\sum_{k=0}^{\infty} \phi_{\frac{k}{2}} \prod_{j=1}^{k} |u_{t-j}| \right). \tag{14}$$

(ii) For any real positive γ, $\underline{X}_{n,t}$ *and* $\overline{X}_{n,t}$ *are bounded (in modulus) a.s., strictly stationary and ergodic when* $\gamma e^{\mu_0} < 1$, *uniformly in* b_γ. *When* $\gamma e^{\mu_0} = 1$, *it is required that* $\delta > -1/2$ *and* $\delta > 0$ *for* $\underline{X}_{n,t}$ *and* $\overline{X}_{n,t}$ *respectively.*

(iii) The asymptotic covariance stationarity conditions of $\underline{X}_{n,t}$, $\overline{X}_{n,t}$ *and* ${}^U X_{n,t}$, *for levels and squares, coincide (cf. Theorem 3 and second remark). Moreover, for the limit squares, when* $\sqrt{\mu_4}\,\gamma < 1$ *or when* $\sqrt{\mu_4}\,\gamma = 1$, $\delta > -3/4$, *as* $h \to \infty$,

$$cov(\underline{X}_t, \underline{X}_{t+h}) \sim c\,(\mu_2\gamma)^h\, f_{\gamma,\sqrt{\mu_4}^{-1};\delta}(h), \quad cov(\overline{X}_t, \overline{X}_{t+h}) \sim c'\,(\mu_2\gamma)^h\, f_{\gamma,\sqrt{\mu_4}^{-1};\delta}(h),$$

setting

$$f_{\gamma,\gamma';b}(h) = \begin{cases} h^{-2(b+1)}, & \gamma < \gamma', \\ h^{-2(b+1)}(1 + h^{-(2b+1)}), & \gamma = \gamma'. \end{cases} \tag{15}$$

Remarks.
(a) We have characterized the limit of the 'envelope' processes $\underline{X}_t$ and $\overline{X}_t$, which appears relatively tractable, rather than looking directly at the limit of the ${}^U X_{n,t}$. The 'envelope' seems sufficiently tight as $\underline{X}_t$ and $\overline{X}_t$ share the same covariance stationarity condition up to the fourth order, as well as the strict stationarity condition for nearly all circumstances. In contrast, the limit of ${}^U X_{n,t}$ has a very cumbersome expression, as stochastic expansion arguments (e.g. Hermite expansions for Gaussian u_t) must necessarily be used. This would make uneasy practical use, as e.g. for estimation of the limit aggregate.
(b) By means of the results developed in this paper, the exact asymptotic properties of $1/n \sum_{i=1}^{n} x_{i,t}^2$ can be easily characterized without using envelope arguments nor stochastic expansions, in contrast to the aggregate ${}^U X_{n,t}$.

However, it would represent a poor approximation for the square of the aggregate, ${}^{U}X^2_{n,t}$, as for instance they are characterized by different stationarity conditions.
(c) The strong degree of cross-sectional dependence of the $x_{i,t}$, induced by the u_t, implies that standard CLTs fail, so that the limit distribution of the aggregate is not Gaussian. In fact, both $\underline{X}_t$ and $\overline{X}_t$ are uncorrelated but not independent, displaying dynamic conditional heteroskedasticity. Note that strict stationarity of the micro ARCH(1) (cf. remark (*d*) of $II(\gamma)$) does not always guarantee that the limit aggregate is well defined. In fact, when $\gamma e^{\mu_0} = 1$, the limit aggregate could be unbounded with probability one for a sufficiently dense distribution of the α_i near $\gamma = 1/e^{\mu_0} > 1$.
(d) In contrast to the case of aggregation of GARCH processes for finite n, considered in Nijman and Sentana (1996), it turns out that the asymptotic limit of the aggregate will not belong to the class of weak GARCH. In fact, under $II(\gamma)$ and $III(\eta)$,

$$\phi_k \sim c\,\gamma^k\, k^{-(b_\gamma+\eta/2+1)}, \;\; k \to \infty, \tag{16}$$

(cf. (32) in the appendix). Thus, the ϕ_k cannot be obtained by expanding the ratio of finite order polynomials in the lag operator L such as $a(L)/b(L) = (1 + a_1 L + \ldots + a_q L^q)/(1 + b_1 L + \ldots + b_p L^p)$ for given integers $p, q \geq 0$ and constants $a_1, \ldots, a_q, b_1, \ldots, b_p$ where $a(L), b(L)$ have roots outside the unit circle in the complex plane (see definitions 1,2 and 3 in Drost and Nijman (1993) for strong, semi-strong and weak GARCH). Note that the orders (p and q) of $a(L)$ and $b(L)$ can be arbitrarily large and yet their ratio be of finite order. However, the meaning of (16) is stronger implying that the (multivariate) Markovian structure of GARCH is lost by aggregation, as $n \to \infty$.
(e) By suitably parameterizing the ϕ_k, say $\phi_k = \phi_k(\theta)$ for some $p \times 1$ vector θ, we can easily estimate the models induced by (13) and (14) by a pseudo maximum-likelihood approach as follows. Consider $\underline{X}_t$. Given a sample $X = (X_1, ..., X_T)'$ of data with sample size T, set $\mu_2 = 1$ and

$$\begin{aligned}
\bar{u}_s &= 0,\; s \leq 0, \\
\bar{u}_1 &= \frac{X_1}{\rho_{\frac{1}{2}}}, \\
\bar{u}_2 &= \frac{X_2}{\rho_{\frac{1}{2}} \left(1 + \phi^2_{\frac{1}{2}}(\theta)\bar{u}_1^2\right)^{\frac{1}{2}}},
\end{aligned}$$

$$\bar{u}_3 = \frac{X_3}{\rho_{\frac{1}{2}}\left(1+\phi^2_{\frac{1}{2}}(\theta)\bar{u}_2^2+\phi^2_1(\theta)\bar{u}_1^2\bar{u}_2^2\right)^{\frac{1}{2}}},$$
$$\bar{u}_4 = \ldots,$$

and

$$\sigma_t = \rho_{\frac{1}{2}}\left(\sum_{k=0}^{\infty}\phi^2_{\frac{k}{2}}\prod_{j=1}^{k}u^2_{t-j}\right)^{\frac{1}{2}}, \;\; \bar{\sigma}_t(\theta) = \rho_{\frac{1}{2}}\left(\sum_{k=0}^{t}\phi^2_{\frac{k}{2}}(\theta)\prod_{j=0}^{k}\bar{u}^2_{t-j}\right)^{\frac{1}{2}}.$$

Setting $u = (u_1, .., u_T)'$ and given the Jacobian

$$det[\frac{\partial u}{\partial X'}] = \prod_{t=1}^{T}\sigma_t^{-1},$$

$det[A]$ indicating the determinant of a square matrix A, the conditional log likelihood can be easily obtained, e.g. for Gaussian u_t,

$$log\,pdf\;(X) = -\frac{T}{2}log(2\pi) - \sum_{t=1}^{T}log\,\sigma_t - \frac{1}{2}\sum_{t=1}^{T}\frac{X_t^2}{\sigma_t^2}.$$

Therefore, by substituting σ_t with $\bar{\sigma}_t(\theta)$, one gets the maximum likelihood estimator (MLE) of $(\rho_{\frac{1}{2}}, \theta)$ and, more generally, the pseudo MLE if the u_t are not Gaussian. A formal development is beyond the focus of this paper but, under suitable regularity conditions, the MLE will be characterized by the usual asymptotic properties.
(f) When parameterizing the ϕ_k, one can allow ϕ_0, ϕ_1 and ϕ_2 to be free parameters, letting $\phi_k = \phi_k(\theta)$ only for $k \geq 3$. This implies that one can perform a test of heterogeneity of the α_i by means of estimation of the limit aggregate model based on aggregate data. In fact, $\alpha_i = \alpha$ under the hypothesis of no-heterogeneity implying

$$\phi_0\,\phi_2 = \phi_1^2, \tag{17}$$

which can be easily verified using estimated values of ϕ_0, ϕ_1 and ϕ_2. Note that when $\eta = 0$, then $\phi_0 = 1$ yielding $var(\alpha_i) = \phi_2 - \phi_1^2$, zero when (17) holds. This test represents an extension of Lewbel (1994) test for heterogeneity of micro ARMA.

(g) The limit processes $\underline{X}_t$, $\overline{X}_t$ differ from all the ARCH-type long memory volatility models introduced in the relevant literature, in particular from the ARCH(∞) of Robinson (1991). Moreover, imposing covariance stationary squares ($\gamma \leq 1/\sqrt{\mu_4}$ and $b_{\sqrt{\mu_4}^{-1}} > -3/4$), implies short memory for the limit process, with an near-exponentially decaying autocovariance function (ACF). The reason is that, for ARCH(1), the set of parameter values consistent with bounded fourth moment, $\alpha_i\sqrt{\mu_4} < 1$, is strictly smaller than the set of parameter values consistent with bounded second moment, $\alpha_i\mu_2 < 1$, given that $\mu_2 < \sqrt{\mu_4}$ (the u_t^2 are not degenerate). Relaxing the constraint $\mu_4 = 3$ is irrelevant. In fact, setting $\mu_4 = 1$ permits $\gamma = 1$, but then the ACF, evaluated at lag h, of the limit aggregate is $O(\mu_2^h)$ as $h \to \infty$, where $\mu_2 < 1$. Note that the degree of cross-sectional dependence between the α_i and the ω_i is irrelevant for the memory properties of the squared limit aggregate.

3 Generalizations

3.1 Aggregation of GARCH

Relevant implications arise with respect of the memory properties of the squares of the limit aggregate when considering GARCH(p,q) with $p > 0$. These implications are ruled out by ARCH(1) and, more generally, by ARCH(q) structure.

We now discuss aggregation of micro GARCH(1,1). Let

$$\begin{aligned} x_{i,t} &= z_{i,t}\sigma_{i,t},, \quad i \in \mathbb{N},\, t \in \mathbb{Z}, \\ \sigma_{i,t}^2 &= \omega_i + \alpha_i x_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2, \;\; a.s. \end{aligned} \tag{18}$$

Given real $0 < \bar{\alpha}$, $\bar{\beta} < \infty$, assume that the α_i, β_i, ω_i satisfy the following assumptions, replacing $II(\gamma)$ and $III(\eta)$.

Assumption IV($\bar{\alpha}, \bar{\beta}$)
(i) The ω_i are i.i.d. with $0 < \omega_i$ almost surely and $E(\omega_i^2) < \infty$ for any $i \in \mathbb{N}$.
(ii) The α_i and the β_j are mutually independent for any $i, j \in \mathbb{N}$.
(iii) The α_i and the β_i are i.i.d. with absolutely continuous distribution in the interval $[0, \bar{\alpha})$ and $[0, \bar{\beta})$, respectively, depending upon the real parameters $b_{\bar{\alpha}}, c_{\bar{\beta}} > -1$, whose densities behave,

$$B(\alpha_i; b_{\bar{\alpha}}) \sim C\,(\bar{\alpha} - \alpha_i)^{b_{\bar{\alpha}}}, \;\; \alpha_i \to \bar{\alpha}^-, \tag{19}$$

$$B(\beta_i; c_{\bar{\beta}}) \sim C\,(\bar{\beta} - \beta_i)^{c_{\bar{\beta}}}, \;\; \beta_i \to \bar{\beta}^-. \tag{20}$$

Assumption V(η)
The ω_i, α_i and β_i satisfy

$$\omega_i = \tilde{\omega}_i \mid 1 - \mu_2 \alpha_i - \beta_i \mid^{\eta}, \tag{21}$$

for an i.i.d. sequence $\tilde{\omega}_i$, mutually independent from the α_i and β_i, with $0 < \tilde{\omega}_i$ almost surely and $E(\tilde{\omega}_i^2) < \infty$ for any $i \in \mathbb{N}$.

Remarks.
(a) Set

$$\pi_i = (\alpha_i + \beta_i), \tag{22}$$
$$\nu_i = (\alpha_i + \beta_i)^2 + 2\alpha_i^2 = \pi_i^2 + 2\alpha_i^2. \tag{23}$$

It turns out the statistical properties of the aggregate in levels $X_{n,t}$, and squares $X_{n,t}^2$, could be characterized by the shape of the cross-sectional distribution of the π_i (the so-called 'persistence' parameter) and of the ν_i, respectively. In particular, for finite n, covariance stationary levels requires

$$\bar{\pi} = \bar{\beta} + \mu_2 \bar{\alpha} \leq 1,$$

and covariance stationary squares

$$\bar{\nu} = (\bar{\beta} + \mu_2 \bar{\alpha})^2 + (\mu_4 - \mu_2^2)\bar{\alpha}^2 \leq 1.$$

Asymptotic covariance stationarity, as $n \to \infty$, will also require that the distribution of the π_i and the ν_i be not too dense around $\bar{\pi}$ and $\bar{\nu}$ respectively.
(b) Mutual independence between the α_i and the β_i is assumed for simplicity's sake. Below, the implications of cross-sectional dependence are explored.

Based on these assumptions, the results of section 2 could be generalized to the case of aggregation of GARCH$(1,1)$. Details are skipped for easy exposition and we only develop the more delicate extension of Theorem 4.

Theorem 5 *For any $n \in \mathbb{N}$, there exist processes $\{\overline{X}_{n,t}, \underline{X}_{n,t}, t \in \mathbb{Z}\}$ such that*

$$\min[\overline{X}_{n,t}, \underline{X}_{n,t}] \leq^U X_{n,t} \leq \max[\overline{X}_{n,t}, \underline{X}_{n,t}], \ a.s.$$

satisfying the following.

(i) Assume $IV(\bar\alpha,\bar\beta)$ *and* $V(\eta)$ *and set* $\delta = min[b_{\bar\alpha}, c_{\bar\beta}] + \eta/2$.
As $n \to \infty$, *when* $\max\left[\bar\alpha^{\frac12}\mu_1 + \bar\beta^{\frac12}, \bar\alpha\mu_2 + \bar\beta\right] < 1$,

$$\underline{X}_{n,t} \to_1 \underline{X}_t, \qquad \overline{X}_{n,t} \to_1 \overline{X}_t.$$

When $\max\left[\bar\alpha^{\frac12}\mu_1 + \bar\beta^{\frac12}, \bar\alpha\mu_2 + \bar\beta\right] = 1$, $min[b_{\bar\alpha}, c_{\bar\beta}] + \eta > 1$ *is needed (see (45) and (46) in the appendix for the definition of* $\{\underline{X}_t, \overline{X}_t, t \in \mathbb{Z}\}$*).*

(ii) For any real positive $\bar\alpha, \bar\beta$, $\underline{X}_t$ *and* $\overline{X}_t$ *are bounded (in modulus) a.s., strictly stationary and ergodic when* $(\bar\alpha e^{\frac{\mu_0}{2}})^{\frac12} + \bar\beta^{\frac12} < 1$ *or* $(\bar\alpha e^{\frac{\mu_0}{2}})^{\frac12} + \bar\beta^{\frac12} = 1$, $\delta > 0$, *for* $\overline{X}_t$ *and when* $\bar\alpha e^{\frac{\mu_0}{2}} + \bar\beta < 1$ *or* $\bar\alpha e^{\frac{\mu_0}{2}} + \bar\beta = 1$, $\delta > -1/2$, *for* $\underline{X}_t$.

(iii) The asymptotic covariance stationarity conditions are the following. For $\underline{X}_{n,t}$, $\bar\pi < 1$ *or* $\bar\pi = 1$, $\delta > -1/2$, *for levels and* $\bar\nu < 1$ *or* $\bar\nu = 1$, $\delta > -3/4$, *for squares. For* $\overline{X}_{n,t}$, $\bar\pi' = \bar\alpha^{\frac12}\mu_2^{\frac12} + \bar\beta^{\frac12} < 1$ *or* $\bar\pi' = 1$, $\delta > -1/2$, *for levels and* $\bar\nu' = \bar\alpha^{\frac12}\mu_4^{\frac14} + \bar\beta^{\frac12} < 1$ *or* $\bar\nu' = 1$, $\delta > -3/4$, *for squares.*

For the limit squares, under those conditions,

$$0 < cov(\underline{X}_t^2, \underline{X}_{t+h}^2) \le C\,\bar\pi^h\, f_{\bar\nu,1;\delta}(h),$$
$$0 < cov(\overline{X}_t^2, \overline{X}_{t+h}^2) \le C'\,\bar\pi'^h\, f_{\bar\nu',1;\delta}(h),$$

as $h \to \infty$, *with* $f_{\gamma,\gamma';b}(h)$ *defined in (15).*

Remark. Formal analysis of aggregation of heterogeneous higher order GARCH(p,q) gets more elaborate. Reparameterizing GARCH(p,q) as GARCH(m,m), for $m = max[p,q]$, yields

$$\sigma_{i,t}^2 = \omega_i + \sum_{k=1}^{m}(\alpha_{i,k} z_{i,t-k}^2 + \beta_{i,k})\sigma_{i,t-k}^2, \quad a.s.$$

It follows that the limit levels could be characterized by looking at the cross-sectional distribution of the $\pi_i = \sum_{j=1}^m(\alpha_{i,j} + \beta_{i,j})$ and for the squares by the cross-sectional distribution of the $\nu_i = \sum_{j=1}^m(\pi_{i,j}^2 + 2\alpha_{i,j}^2)$.

Having developed the effect of aggregation for GARCH$(1,1)$, it is useful to compare certain aspects of our results with previous related ones. Let us first summarize the findings of Ding and Granger (1996) and Leipus and Viano (1999).

Ding and Granger (1996) consider

$$\begin{aligned} X_{n,t}^{DG} &= u_t \left(\sum_{i=1} w_i \tau_{i,t}^2 \right)^{\frac{1}{2}}, \quad \sum_{i=1}^{n} w_i = 1, \\ \tau_{i,t}^2 &= \sigma^2(1-\alpha_i-\beta_i) + \alpha_i (X_{n,t-1}^{DG})^2 + \beta_i \tau_{i,t-1}^2, \end{aligned}$$

where $\mu_2 = 1$ and σ^2 is a constant parameter. Assume that I holds. Note that $X_{n,t}^{DG}$ differ from $X_{n,t}$ in (1) and, moreover, that the $\tau_{i,t}^2$ are not univariate GARCH(1, 1), given by (18). This special structure allows Ding and Granger to apply Granger (1980) linear aggregation results, suggesting that, as $n \to \infty$, $\sum_{i=1}^{n} w_i \tau_{i,t}^2$ converge (in some norm) to a special case of Robinson (1991) ARCH(∞), with hyperbolically decaying coefficients. It is assumed that the β_i have a Beta distribution, over $[0, 1]$, and that $\alpha_i = \alpha_i^*(1-\beta_i)$, where the α_i^* are independent from the β_i and have an unspecified distribution. Note that $E(\alpha_i \mid \beta_i) = E(\alpha_i^*)(1-\beta_i)$, a case of perfect negative cross-sectional dependence.

Leipus and Viano (1999) study the asymptotic behaviour of

$$Y_{n,t}^{LV} = \frac{1}{n} \sum_{i=1}^{n} x_{i,t}^2,$$

where the $x_{i,t}$ are random coefficients ARCH(∞), and establish a sufficient condition for covariance stationarity of the limit that rules out long memory, in the sense of imposing absolute summability of the ACF. Moreover, they provide an example with GARCH(1, 1) $x_{i,t}$ (cf. (18)) where the ACF of the limit of $Y_{n,t}^{LV}$ is hyperbolically decaying. They assume $\alpha_i = \beta_i(1-\beta_i)$ and a Beta parametric distributional assumption for the β_i over $[0, 1]$. Note how this implies another case of perfect negative cross-sectional dependence, given by $E(\alpha_i \mid \beta_i) = \beta_i(1-\beta_i)$, very similar to Ding and Granger (1996).

From a methodological point of view, we have shown that analyzing the effect of aggregation of GARCH is a different mathematical problem than aggregation of ARMA and the extension is not straightforward. The adopted definition of the aggregate represents a crucial aspect. With this respect, the definition used by Leipus and Viano (1999) is closer to our definition than the one used by Ding and Granger (1996), although $Y_{n,t}^{LV}$ differs from $X_{n,t}^2$ for two important reasons, independently from the specification of the $x_{i,t}$. In fact, given

$$X_{n,t}^2 = A + B,$$

$$A = \frac{1}{n^2}\sum_{i=1}^{n} x_{i,t}^2, \quad B = \frac{1}{n^2}\sum_{\substack{i,j=1 \\ i\neq j}}^{n} x_{i,t}x_{j,t}.$$

$Y_{n,t}^{LV}$ equals A/n, involving only the sum of the squares. B involves the double products. Therefore, the limit of $Y_{n,t}^{LV}$ provides only an approximation to the squares of the limit aggregate $X_{n,t}^2$ as, specifically, $Y_{n,t}^{LV}$ converges to zero in first mean, under the condition for covariance stationary $X_{n,t}^2$ (cf. Theorem 5-(iii)). Secondly, the double products term B represent precisely the key ingredient in order to evaluate the effect of the degree of cross-sectional dependence of the rescaled innovations $z_{i,t}$. In fact, when $z_{i,t} = \epsilon_{i,t}$, term B is asymptotically negligible whereas for $z_{i,t} = u_t$ term B defines the behaviour of the limit aggregate and term A is negligible (or at most of the same order of magnitude in the nonstationary case, see Theorem 1-(iii) and 3-(iii)).

With respect to the memory properties, Ding and Granger (1996) and Leipus and Viano (1999) obtain precisely the same type of results, an hyperbolic decaying yet summable ACF for the limit squares (by summability of the ACF, formally, this is not long memory although practically there could be little difference.). This is not surprising as they share a set of crucial assumptions. We recall that this result regarding the memory of the squared limit aggregate is not explicit in Ding and Granger (1996) but it follows considering recent findings on the memory of ARCH(∞), where summability of the ACF of squares is obtained when imposing covariance stationary squares (see Giraitis, Kokoska, and Leipus (1998)).

The key aspect, defining the degree of memory of the limit squared aggregate, is represented by the degree of cross-sectional dependence between the α_i and the β_i. To make this clear, it is useful to compare the covariance stationarity conditions for levels and squares of GARCH$(1,1)$ $x_{i,t}$, in turn equal to

$$\alpha_i\mu_2 + \beta_i < 1, a.s., \tag{24}$$

for levels and to

$$(\alpha_i\mu_2 + \beta_i)^2 + \alpha_i^2(\mu_4 - \mu_2^2) < 1, a.s., \tag{25}$$

for squares. Given $\mu_2^2 < \mu_4$ (u_t^2 are not degenerate), then (25) is strictly stronger than (24), for independent α_i, β_i and, more generally, for any shape of the cross-sectional dependence such that $E(\alpha_i \mid \beta_i) \sim C$, with $C > 0$, as

$\beta \to \bar{\beta}^-$. This certainly includes the case of positive dependence or even negative dependence as long as the marginal distribution of the α_i has support $[c, \bar{\alpha})$ for some $c > 0$.

Set $\eta = 0$ for simplicity's sake. Imposing covariance stationary squares (cf. part (iii) of Theorem 5), the asymptotic behaviour of the ACF of limit aggregate square ($\underline{X}_t^2$, say) is given by $[E\pi_i^{\frac{u}{2}}]^2$, as $u \to \infty$. Now, replace $IV(\bar{\alpha}, \bar{\beta})$-(ii) with

$$E(\alpha_i \mid \beta_i) \sim c(\bar{\beta} - \beta_i)^\delta, \;\; \beta_i \to \bar{\beta}^- \tag{26}$$

for some $\delta \geq 1$. Ding and Granger (1996) and Leipus and Viano (1999) assumption represents a particular case of (26).

Under (26), the left hand side of (24) and (25) goes to $\bar{\beta}^-$ and to $\bar{\beta}^{2-}$, respectively, as $\beta_i \to \bar{\beta}^-$. Moreover, $\bar{\beta} = 1$ is allowed for so that (25) will not be strictly stronger than (24) anymore. Under these conditions the π_i will have support $[0, 1)$ or, equivalently, $\bar{\pi} = 1$. This leads to $E(\pi_i^u) \sim c\bar{\pi}^u u^{-(c_{\bar{\beta}}+1)} = cu^{-(c_{\bar{\beta}}+1)}$ as $u \to \infty$, yielding an hyperbolically decaying ACF of the limit square aggregate, $c' u^{-2(c_{\bar{\beta}}+1)}$ as $u \to \infty$. Summability of the ACF follows from the covariance stationarity condition for levels, $c_{\bar{\beta}} > -1/2$, equal to $2(c_{\bar{\beta}} + 1) > 1$. This condition is in general weaker but certainly not stronger than the covariance stationarity conditions for squares. Therefore, long memory for the squared limit aggregate is ruled out, independently from the shape of the joint cross-sectional distribution of the micro parameters. This type of result is not new to ARCH-type models. In fact for ARCH(∞) also the covariance stationarity condition for levels rules out long memory squares (see Zaffaroni (1999, Theorem 5)).

Note that for the opposite case of negative dependence, $E(\beta_i \mid \alpha) \sim c(\bar{\alpha} - \alpha_i)$, as $\alpha_i \to \bar{\alpha}^-$, (25) will still be strictly stronger than (24), given that $\mu_2^2 < \mu_4$, implying an exponentially decaying ACF. This is not a surprising outcome, as the $x_{i,t}$ behave in this case (locally) like an ARCH(1), for which the case hyperbolical decaying ACF was ruled out.

3.2 Further extensions

(a) We have so far considered only aggregation of covariance stationary units. However, this is much more restrictive than imposing only strictly stationary units (cf. remark (d) of $II(\gamma)$). Aggregation of IGARCH is an important and particular case, where the π_i will have a degenerate distribution at 1.

Let us focus on micro ARCH(1). Although case $\gamma > 1$ does not appear to be empirically relevant (see e.g. Bollerslev, Chou, and Kroner (1992)), our framework can easily account for such a possibility. It is obvious that now one needs to evaluate conditional moments, not only with respect to the GARCH coefficients but also with respect to past rescaled innovations, considering the truncated aggregates ${}^{E}\tilde{X}_{n,t}$, ${}^{U}\tilde{X}_{n,t}$. For example, focusing on the idiosyncratic case, one obtains that as $t, n \to \infty$, under $II(\gamma)$, $III(0)$ (skipping details for simplicity's sake)

$$\mathrm{var}_n({}^{E}\tilde{X}_{n,t}) \sim \gamma^t \left(c\,(log\,log(n\,\phi_{2t}))^{\frac{1}{2}} \frac{t^{-b_\gamma/2+1/2}}{n^{3/2}} + c' \frac{t^{-b_\gamma}}{n} \right).$$

Note that, irrespective of the value of b_γ, ${}^{E}\tilde{X}_{n,t}$ converges to zero in mean-square as $t/n \to \infty$ suggesting that the usual result arises when n is large compared with t. A fixed t is an important particular case. When $n \sim ct^{b_\gamma+1}$ and $\gamma = 1$, the results of Theorem 1 are re-obtained. On the other hand, when $\gamma > 1$, the rate of divergence is exponential with respect to the time dimension. Parallel results can be obtained for the common innovations case.
(b) We can allow for cross-correlation across units not only through a common rescaled innovation but also by assuming dependence across the α_i. Indeed, the limit laws on which this paper is based, have been extended to the case of stationary dependent sequences satisfying some form of mixing condition (see references in Samorodnitsky and Taqqu (1994, p.575)) and are therefore fairly easily adaptable to our framework. We acknowledge that, in this case, a problem of interpretation arises by adapting the time series meaning of dependence to a cross-sectional framework.
(c) Analogous results would be obtained when a number m ($m < n$) of units exhibits different properties from $II(\gamma)$, e.g. IGARCH(1), as long as these units are bounded a.s. and $1/m + m/n \to 0$ a.s. for $n \to \infty$, meaning that they make a degenerate fraction of units. In this case, the properties of the limit aggregate will be entirely determined by the non-degenerate fraction of units $x_{i,t}$, whose properties are defined by the shape of the cross-sectional distribution of the parameters dictated by $II(\gamma)$. The case of a non-degenerate fraction of units, not characterized by $II(\gamma)$, or which are not uniformly bounded a.s. will not be discussed but it can nonetheless be accounted for by a suitable generalization of our framework.

4 Application: aggregation of conditionally heteroskedastic factor models

This framework can be used to evaluate the impact of aggregation of the components of conditionally heteroskedastic factor models:

$$x_{i,t} = \beta_{i,1} f_{1,t} + \beta_{i,2} f_{2,t} + .. + \beta_{i,K} f_{K,t} + w_{i,t},\ i = 1, ..., n, \tag{27}$$

where $f_t = (f_{1,t}, ..., f_{K,t})'$ is a vector of K $(n > K)$ unobserved common factors, the $\beta_{i,j}\,(j = 1, .., K)$ are the associated factor loadings and the $w_{i,t}\,(i = 1, .., n)$ indicate idiosyncratic r.vs, orthogonal to the $f_{j,t}\,(j = 1, .., K)$. In particular, consider a situation where a large number of assets exist, so that n is arbitrarily large, and our random coefficients framework applies. The portfolio, made by $1/n$th of each asset, would then be

$$\frac{1}{n} \sum_{i=1}^{n} \beta_i' f_t + \frac{1}{n} \sum_{i=1}^{n} w_{i,t},$$

setting $\beta_i = (\beta_{i,1}, ..., \beta_{i,K})'$.

For $w_t = (w_{1,t}, ..., w_{n,t})'$ assume $E_{t-1}(w_t) = 0$, $E_{t-1}(w_t w_t') = \Gamma_t$ and $E_{t-1}(f_t) = 0$, $E_{t-1}(f_t f_t') = \Lambda_t$ with $E_{t-1}(w_t f_t') = 0$, $E_{t-1}(\cdot)$ denoting the expectation operator conditionally on $\{f_s,\, x_{i,s}, s < t, i \in \mathbb{N}\}$. The time variation of Λ_t and Γ_t motivates the denomination of conditionally heteroskedastic factor model. Sentana (1998) shows that several multivariate volatility models are described by (27), such as the latent factor model with ARCH factors of Diebold and Nerlove (1989) and the factor GARCH model of Engle (1987).

Depending on whether the $w_{i,t}$ are assumed conditionally mutually orthogonal (diagonal Γ_t) or mildly correlated (non-diagonal Γ_t) across units, (27) is referred to as a conditional exact or approximate K factor structure (see Hansen and Richard (1987)), generalizing the definition of factor structure introduced by Chamberlain and Rothschild (1983).

Consider the case when the idiosyncratic component have a ARCH(1) parameterization:

$$\begin{aligned} w_{i,t} &= z_{i,t}\sigma_{i,t},\ \ i \in \mathbb{N}, t \in \mathbb{Z}, \\ \sigma_{i,t}^2 &= \omega_i + \alpha_i w_{i,t-1}^2, \ \ a.s. \end{aligned}$$

Setting $z_{i,t} = \epsilon_{i,t}$ or, alternatively, $z_{i,t} = u_t$, suggests that a conditional exact or approximate K factor structure for the the $x_{i,t}$ could be obtained.

In Chamberlain and Rothschild (1983) the maximum degree of cross-sectional dependence allowed for the $w_{i,t}$ is expressed by boundedness of the maximum eigenvalue of $E(\Gamma_t)$, uniformly in n. This clearly collapses

$$lim_{n\to\infty} \max_{i=1,..,n} \text{var}(w_{i,t}) < \infty,$$

when Γ_t is diagonal, in turn when $z_{i,t} = \epsilon_{i,t}$. Although this is a sufficient condition for

$$\frac{1}{n^2}\sum_{i=1}^{n} var(w_{i,t}) \to 0, \quad n \to \infty, \tag{28}$$

it could be substantially relaxed. In fact, in our random coefficient framework, the equivalent condition to (28) would then be (consider case $\mu_2 = 1$ and $\eta = 0$)

$$\frac{1}{n^2}\sum_{i=1}^{n} \frac{\omega_i}{1-\alpha_i} \to 0, \;\; a.s., \quad n \to \infty.$$

Under $II(\gamma)$, this certainly holds when $\gamma < 1$. However, when $\gamma = 1$, one then needs $b_1 > -1/2$. Note that for $\gamma < 1$ or $\gamma = 1$, $b_1 > 0$, then $var(w_{i,t}) = E(\frac{\omega}{1-\alpha}) < \infty$ whereas $var(w_{i,t})$ is unbounded for $-1/2 < b_1 < 0$. The former situation expresses the bounded eigenvalue condition whereas the latter describes the case where maximum eigenvalue, $max_{i=1,..,n}\omega_i/(1-\alpha_i)$, of $E_n(\Gamma_t)$ could go to infinity although not too quickly, diverging at rate $n^{\frac{1}{b_1+1}}$ (cf. Theorem 1).

To summarize, although $\omega_i/(1-\alpha_i) < \infty$ a.s., this is not sufficient to ensure that idiosyncratic risk is fully diversifiable when trading a possibly infinite number of assets. Moreover, the aggregate of the idiosyncratic component will be described by a stable model, Gaussian in the stationary case or when the number of assets is much larger than sample size, and not a proper volatility model (cf. Theorem 2). This is particularly important when considering exact zero factor structures, when there are no common factors ($K = 0$). These results must be born in mind when using statistical inference methods on such nonlinear factor models based on a large cross-section, such as when extracting the common component (factors). This is because the crucial assumption of a vanishing importance of the idiosyncratic part of the portfolio might fail.

When $z_{i,t} = u_t$ the $w_{i,t}$ are correlated across units with $E(w_{i,t}w_{j,t}) = \mu_2 E(\sigma_{i,t}\sigma_{j,t})$. Now the degree of cross-sectional dependence is too strong as one cannot fully diversify the risk induced by the $w_{i,t}$, viz. $var(1/n \sum_{i=1}^{n} w_{i,t}) \geq c > 0$, for any n (cf. Theorem 3). Hence, although setting $z_{i,t} = u_t$ delivers an interesting case of a factor model with cross-sectional correlated idiosyncratic risk, this rules out the case of (conditional) approximate factor models.

5 Concluding remarks

The asymptotic properties of the aggregate $X_{n,t}$, as n goes to infinity, are fully characterized by the shape of the cross-sectional distribution of micro parameters, by their degree of cross-sectional dependence and by the degree of cross-sectional dependence of the rescaled innovations.

Only for a sufficient degree of cross-sectional dependence of the rescaled innovations will the limit aggregate maintain the GARCH nonlinearity, uncorrelatedness in levels and dependence in squares. In this case, the model can be parameterized and estimated by pseudo-maximum likelihood. A test for heterogeneity is proposed. When the rescaled innovations are independent across units, the limit aggregate will not maintain the nonlinearity anymore, exhibiting a stable asymptotic distribution.

Imposing bounded fourth moment of the limit aggregate rules out long memory in its squares, in particular yielding an (approximately) exponentially decaying ACF, for any shape of the cross-sectional distribution. When considering aggregation of GARCH(p, q), with $p \geq 1$, however, although long memory is still ruled out, one could obtain an hyperbolically decaying ACF of the squared limit, under a restrictive form of negative cross-sectional depen dence between the GARCH(p, q) parameters.

When the cross-sectional distribution of the micro parameters is still defined over the covariance stationary region for $x_{i,t}$ but is sufficiently dense around their upper limit, then mean square convergence toward zero of the limit aggregate fails even for *i.i.d.* $x_{i,t}$. An application of this result to aggregation of conditionally heteroskedastic factor models is presented.

These aspects, related to the cross-sectional distribution of the parameters and of the rescaled innovations, describe a rich set of testable implications linking the statistical properties of the cross-section to the aggregate ones. These implications could be verified empirically and subjected to hypothesis testing, using a panel of data.

Long memory at the aggregate level is ruled out by the different restrictions imposed on the parameters space by the covariance stationary condition for levels and squares respectively. This represents an undesirable feature of GARCH. In Zaffaroni (2001), the issue of inducing long memory squares by aggregation is analyzed in a more general framework of models of changing volatility, which include both ARCH-type and SV-type volatility models.

Appendix

We should recall that c, C denote arbitrary positive constants, always bounded and not necessarily the same, the symbol $\sim$ denotes asymptotic equivalence and $P(A)$, 1_A, respectively, the probability and the indicator function of any event A. We first introduce two preliminary lemmas, then present the proof of the theorems.

Lemmata

We begin with the following lemma, proved in Lippi and Zaffaroni (1998, Lemma 1), which adapts to our framework known results on convergence of normed sums in *i.i.d.* r.vs in the domain of attraction of a possibly non-Gaussian stable distribution.

Lemma 1 *Consider n i.i.d. draws of a positive r.v. α with probability density $B(\cdot; b)$ defined in the interval $[0, \gamma)$ for real $\gamma > 0$ such that for a real b $(-1 < b < \infty)$, as $\alpha \to 1^-$,*

$$B(\alpha; b) \sim C\,(\gamma - \alpha)^b\, L(\frac{1}{(\gamma - \alpha)}), \tag{29}$$

where $L(\cdot)$ denotes a slowly varying function. Set $\delta = (b+1)/k$. Then a.s., as $n \to \infty$,

(1) If $2 \leq \delta$,

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{(\gamma - \alpha_i)^k} \sim E(\frac{1}{(\gamma - \alpha)^k}) + n^{-\frac{1}{2}} S_2,$$

(2) If $1 < \delta < 2$,

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{(\gamma - \alpha_i)^k} \sim E(\frac{1}{(\gamma - \alpha)^k}) + \tilde{L}(n) n^{1/\delta - 1} S_\delta,$$

(3) If $\delta = 1$,

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{(\gamma - \alpha_i)^k} \sim \hat{L}(n) + \tilde{L}(n) S_1,$$

(4) If $0 < \delta < 1$,

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{(\gamma - \alpha_i)^k} \sim \tilde{L}(n)\, n^{1/\delta - 1} S_\delta,$$

where $\tilde{L}(n), \hat{L}(n)$ are slowly varying functions. We recall that $S_\delta\,(0 < \delta \le 2)$ defines a δ-stable r.v. with zero location parameter (cf. footnote 1), including the case of the normal distribution (S_2). In case (4) S_δ will be a totally skewed to the right δ-stable r.v. with zero location parameter, implying $S_\delta > 0$ a.s.

Remark. When $L(\cdot) = 1$ then $\tilde{L}(\cdot) = 1$ and $\hat{L}(\cdot) = log(\cdot)$.

Lemma 2 *Under the assumptions of Lemma 1 with $\gamma = 1$, for any integer $p = 1, 2, ..$ and real k, as $n \to \infty$:*
(i)

$$\frac{1}{n^p} \sum_{i_1,..,i_p=1}^{n} \frac{1}{(1-\alpha_{i_1}...\alpha_{i_p})^k} \sim c + C\frac{1}{n}\sum_{i=1}^{n}(1-\alpha_i)^{(p-1)b+(p-1-k)}.$$

The boundedness condition is $pb + (p-k) > 0$.
(ii) When $pb + (p-k) > 0$ for any integer $u > 0$ and $r\,(0 \le r \le p)$ with $s = p - r$, as $n \to \infty$,

$$\frac{1}{n^p} \sum_{i_1,...i_r...,i_p=1}^{n} \frac{\alpha_{i_1}^u..\alpha_{i_r}^u}{(1-\alpha_{i_1}...\alpha_{i_p})^k} \to g_{(r,s)}^{(k)}(u), \;\; a.s.$$

where, as $u \to \infty$,

$$g_{(r,s)}^{(k)}(u) \sim c\,(E(\alpha_i^u))^r(1 + u^{-(sb+s-k)}),$$

for $0 < c < \infty$.

Proof. Cases $\gamma < 1$ or $\gamma = 1$, $k < 0$ are trivial. Let us focus on case $\gamma = 1$, $k > 0$. We discuss case $p = 2$, as the other cases follow exactly along the same lines. Set $L(\cdot)$ in (29) for simplicity's sake.
(i) As $n \to \infty$,

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{(1-\alpha_i\alpha_j)^k} \to c\int_0^1 (1-t)^b(1-\alpha_j t)^{-k}dt < \infty \;\; a.s.,$$

by (29). Using Gradshteyn and Ryzhik (1994, # 3.197-3), the integral of the right hand side equals

$$B(1, b+1)\,{}_2F_1(k, 1; 2+b; \alpha_j), \tag{30}$$

where $B(\cdot,\cdot)$ is the Beta function and ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ denotes the hypergeometric function (Gradshteyn and Ryzhik 1994, section 9.1). Hence, as $\alpha_j \to^- 1$, by Gradshteyn and Ryzhik (1994, # 9.122-1 and # 9.131-1)

$$(30) \sim \left(c\, 1_{b>k-1} + c' 1_{b=k-1} log(1-\alpha_j)\; + c'' 1_{b<k-1} (1-\alpha_j)^{b+1-k}\right),$$

yielding, as $n \to \infty$,

$$\frac{1}{n^2}\sum_{i=1}^{n} \frac{1}{(1-\alpha_i\alpha_j)^k} \sim c + C\frac{1}{n}\sum_{i=1}^{n}(1-\alpha_i)^{b+1-k},$$

with a bounded limit when $2b+1-k > -1$. (ii) When $r = 2$ ($s = 0$), by Gradshteyn and Ryzhik (1994, # 3.197-3), as $n \to \infty$,

$$\frac{1}{n^2}\sum_{i,j=1}^{n} \frac{\alpha_j^u \alpha_i^u}{(1-\alpha_i\alpha_j)^k} \sim \frac{c}{n}\sum_{i=1}^{n} \alpha_i^u B(u+1,b+1) {}_2F_1(k,u+1;u+2+b;\alpha_i).$$

Using the aforementioned results of Gradshteyn and Ryzhik (1994), as $\alpha_i \to 1^-$,

$$\begin{aligned} {}_2F_1(k,u+1;u+2+b;\alpha_i) \quad &\sim 1_{b>k-1}\ {}_2F_1(k,u+1;u+2+b;1) \\ &+1_{b<k-1}\ (1-\alpha_i)^{b+1-k} {}_2F_1(u+2-k+b,b+1;u+2+b;1), \end{aligned}$$

the logarithmic term, arising when $b = k-1$, being absorbed by the second one. Simplifying terms yields, as $n \to \infty$,

$$\begin{aligned} \frac{1}{n^2}\sum_{i,j=1}^{n} \frac{\alpha_j^u \alpha_i^u}{(1-\alpha_i\alpha_j)^k} \quad &\sim \frac{1_{b>k-1} c}{n} \frac{\Gamma(b+1-k)\Gamma(u+1)}{\Gamma(u+2-k+b)} \sum_{i=1}^{n} \alpha_i^u \\ &+ \frac{1_{b<k-1}\, c'}{n}\sum_{i=1}^{n} \alpha_i^u (1-\alpha_i)^{b+1-k}. \end{aligned} \tag{31}$$

By Stirling's formula (Brockwell and Davis 1987, p.522)

$$\int_0^s u^k (s-u)^b \sim C\, s^{k+b-1}\, k^{-(b+1)}, \;\; k \to \infty, \tag{32}$$

for $b > -1$ and $0 < s < \infty$. Thus, under (29),

$$E(\alpha_i^k) \sim c\, \gamma^k\, k^{-(b+1)}, \;\; k \to \infty. \tag{33}$$

It follows that the limit (as $n \to \infty$) of both terms on the right hand side of (31) are asymptotically equivalent to $u^{-(2b+2-k)}$, for $u \to \infty$. We recall that to impose stationarity $2b+2-k > 0$.

Finally, when $r = 1$ ($s = 1$), as $n \to \infty$,

$$\frac{1}{n^2} \sum_{i,j=1}^{n} \frac{\alpha_i^u}{(1-\alpha_i \alpha_j)^k} \sim \frac{1_{b>k-1}\, c}{n} \sum_{i=1}^{n} \alpha_i^u + \frac{1_{b<k-1}\, c'}{n} \sum_{i=1}^{n} \alpha_i^u (1-\alpha_i)^{b+1-k}. \ \square$$

Main results

Proof of Theorem 1. When $\eta = 0$ apply Lemma 1 with $k = 1$. When $\eta = 1$

$$\mathrm{var}_n({}^{E}X_{n,t}) = \frac{1}{n^2} \sum_{i=1}^{n} \tilde{\omega}_i = O(\frac{1}{n}) \ \ a.s. \ \square$$

Proof of Theorem 2. Set $\eta = 0$. (i) Given the $i.i.d.$-ness of the $x_{i,t}$, the Lindeberg-Lévy CLT applies, as $n \to \infty$. In fact $1/n \sum_{i=1}^{n} \omega_i/(1-\alpha_i)$ converges to $E\left(\omega_i/(1-\alpha_i)\right)$ a.s., bounded when $b_1 > 0$. Moreover, for any integer $u > 0$ and any n, easy calculations yield

$$cov_n(\frac{1}{n^{\frac{1}{2}}} \sum_{i=1}^{n} x_{i,t}, \frac{1}{n^{\frac{1}{2}}} \sum_{i=1}^{n} x_{i,t+u}) = 0.$$

where $cov_n(.,.)$ denotes the covariance operator, conditioning on the ω_i, α_i ($i = 1, .., n$).

(ii) For (9) we follow (i), where from (33), when $b_1 < 0$,

$$E(\tilde{x}_{i,t}^2) = V_t = E\left(\omega_i \frac{1-\alpha_i^t}{1-\alpha_i}\right) = \rho_1 \sum_{k=0}^{t-1} \phi_k \sim c\, t^{-b_1}, \ \ t \to \infty.$$

For (10), we must first show that the distribution of $\sigma_{i,t}^2$ belongs to the domain of attraction of a δ'-stable r.v., with $\delta' = (b_1+1)$. Setting $c_k(t) = \prod_{s=1}^{k} \epsilon_{i,t-s}^2$ and $d_q(t) = \max_{k=q,q+1,..} c_k(t)$,

$$\begin{aligned} P(\sigma_{i,t}^2 > u) &= P(\sum_{k=0}^{\infty} \alpha_i^k\, c_k(t) > u) \\ &\le P(d_0(t) \sum_{k=0}^{\infty} \alpha_i^k > u) = P(d_0(t)\,/(1-\alpha_i) > u) \\ &= E(P_\epsilon(d_0(t)\,/(1-\alpha_i) > u)), \end{aligned} \tag{34}$$

setting $P_\epsilon(.)$ equal to the probability distribution, conditioning upon the $\epsilon_{i,t}\,(i \in \mathbb{N}, t \in \mathbb{Z})$. Next, by Dudley (1989, Theorem 8.3.5), with probability one there exists a random integer $K < \infty$ such that for all $k > K$

$$c_k(t) = O(e^{\frac{k\,\mu_0}{2}}), \;\; a.s., \tag{35}$$

implying that $c_k(t) \to 0$ a.s. for $k \to \infty$. Therefore, for some m, such that $m \to \infty$,

$$P(\sigma^2_{i,t} > u) = E(P_\epsilon(\sigma^2_{i,t} > u)) \geq E(P_\epsilon(d_m(t)/(1-\alpha_i) > u)). \tag{36}$$

Let us now consider the $y_i = 1/(1-\alpha_i)$. We show that, under $II(\gamma)$, the y_i have distribution which belongs to the domain of attraction of a δ'-stable distribution, totally skewed to the right. In fact, denoting by $f_y(\cdot)$ the probability density function of the y_i,

$$f_y(u) = B(1-u^{-1}; b_1)u^{-2}, \;\; 1 \leq u < \infty,$$

with $f_y(u) \sim \; c\,u^{-(b_1+2)}$ as $u \to \infty$. Therefore, as $u \to \infty$

$$P(y_i \geq u) \sim C\,u^{-(b_1+1)}, \;\; P(y_i < -u) = 0, \tag{37}$$

and

$$\int_0^u t^2 f_y(t)dt \sim cu^{2-(b_1+1)},$$

and Feller (1966, Theorem IX.8.1) applies, yielding, for $n \to \infty$

$$\frac{1}{n^{\frac{1}{\delta'}}} \sum_{i=1}^{n} \frac{1}{1-\alpha_i} \to_d S_{\delta'}, \tag{38}$$

where, using Samorodnitsky and Taqqu (1994) notation, $S_{\delta'}$ refers to $S_{\delta'}(\sigma, 1, 0)$ with zero location parameter, skewness parameter equal to 1 (implying $S_{\delta'} > 0$ a.s.) and scale parameter $\sigma = (C/D_{\delta'})^{\frac{1}{\delta'}}$ (with C as in (37)) where

$$D_a = \begin{cases} \frac{1-a}{\Gamma(2-a)cos(\pi a/2)}, & a \neq 1, \\ 2/\pi, & a = 1, \end{cases}$$

(see Samorodnitsky and Taqqu (1994, Property 1.2.15 and eq.(1.2.9))).

(38) implies, under $II(\gamma)$ and $III(\eta)$, that the $\sigma^2_{i,t}$ also have distribution which belongs to the domain of attraction of a δ'-stable distribution, totally

skewed to the right. In fact, from (34), by dominated convergence theorem ($P_\epsilon(\cdot) \leq 1$ and $E\,1 = 1 < \infty$), as $u \to \infty$,

$$P(\sigma_{i,t}^2 > u) \leq E(P_\epsilon(d_0(t)\,/(1-\alpha_i) > u)) \sim c\,E\,[d_0(t)]^{\delta'}\,u^{-\delta'},$$

given $E[d_0(t)]^{\delta'} \leq [Ed_0(t)]^{\delta'} < \infty$, as $\delta' < 1$, by Jensen inequality. On the other hand, for (36), setting $m = m(u)$ and assuming that m goes to infinity with u, as $u \to \infty$,

$$P(\sigma_{i,t}^2 > u) \geq E(P_\epsilon(d_m(t)\,/(1-\alpha_i) > u)) \sim cE\,[d_m(t)]^{\delta'}\,u^{-\delta'} = g(u)\,u^{-\delta'},$$

for some positive function $g(u) \downarrow 0$ as $u \to \infty$. Therefore, by the arbitrariness of $g(u)$,

$$P(\sigma_{i,t}^2 > u) \sim\; c\,u^{-\delta'},\;\; u \to \infty,$$

c depending on both the distribution of the $\epsilon_{i,t}$ and of the α_i. Moreover, from

$$P(\sigma_{i,t}^2 > u) = P(\sigma_{i,t} > u^{\frac{1}{2}}),$$

it follows that $\sigma_{i,t}$ belongs to the domain of attraction of a δ-stable distribution, totally skewed to the right, setting $\delta = 2(b_1 + 1)$. Therefore, given $x_{i,t} = \epsilon_{i,t}\sigma_{i,t}$,

$$\frac{1}{n^{\frac{1}{\delta}}}\sum_{i=1}^{n} x_{i,t} \to_d S_\delta(t),\;\; n \to \infty,$$

where the $S_\delta(t)$ have a δ-stable marginal distribution with zero location parameter, skew parameter β that depends on the symmetry (around zero) of the $\epsilon_{i,t}$, given by $\beta = (P-Q)/(P+Q)$ (see Feller (1966, eq.(8.4)) and Samorodnitsky and Taqqu (1994, p.6)), setting

$$P = \frac{p\,E[\epsilon_{i,t}^{\delta} 1_{\epsilon_{i,t}>0}]}{p\,E[\epsilon_{i,t}^{\delta} 1_{\epsilon_{i,t}>0}] + (1-p)\,E[(-\epsilon_{i,t})^{\delta} 1_{\epsilon_{i,t}<0}]},$$

with $p = P(\epsilon_{i,t} > 0)$, $Q = 1 - P$, and scale parameter, σ, defined by

$$P(x_{i,t} > u) \sim D_\delta\,\frac{1+\beta}{2}\sigma^\delta u^{-\delta},$$
$$P(x_{i,t} < -u) \sim D_\delta\,\frac{1-\beta}{2}\sigma^\delta u^{-\delta},$$

as $u \to \infty$. When $\eta = 1$, (i) applies. □

Proof of Theorem 3. Set $\eta = 0$. By Schwarz inequality

$$\sigma_{i,t}\sigma_{j,t} \geq \omega_i^{\frac{1}{2}}\omega_j^{\frac{1}{2}} \left(\sum_{a=0}^{\infty} (\alpha_i\alpha_j)^{a/2} \prod_{h=1}^{a} u_{t-h}^2 \right),$$

and taking expectations

$$\text{var}_n({}^U X_{n,t}) \geq \frac{c}{n^2} \sum_{i,j=1}^{n} \frac{(\omega_i\omega_j)^{\frac{1}{2}}}{(1-\alpha_i\alpha_j)},$$

and, likewise,

$$\text{var}_n({}^U X_{n,t}) \leq \left(\frac{C}{n} \sum_{i=1}^{n} \frac{\omega_i^{\frac{1}{2}}}{(1-\alpha_i)^{\frac{1}{2}}} \right)^2.$$

Case $\gamma < 1$ easily follows as

$$\frac{1}{n^2} \sum_{i,j=1}^{n} \frac{1}{(1-\alpha_i\alpha_j)} = \sum_{k=0}^{\infty} \left(\frac{1}{n} \sum_{i=1}^{n} \alpha_i^k \right)^2 \leq \frac{1}{(1-\gamma^2)} < \infty, \ \ a.s.$$

When $\gamma = 1$, apply Lemma 2 with $p = 2$, $k = 1$ and then Lemma 1 with $k = -b_\gamma$. Case $\eta = 1$ easily follows. □

Proof of Theorem 4. Using suitable versions of Minkowski's inequality (Hardy, Littlewood, and Polya 1964, Theorem 24 and 25), for any sequence $\{a_{i,j},\ i = 1, \ldots,\ j = 1, \ldots, n\}$ one obtains:

$$\left(\sum_{i=0}^{\infty} (\frac{1}{n} \sum_{j=1}^{n} (a_{i,j})^{\frac{1}{2}})^2 \right)^{\frac{1}{2}} \leq \frac{1}{n} \sum_{j=1}^{n} \left(\sum_{i=0}^{\infty} a_{i,j} \right)^{\frac{1}{2}} \leq \left(\sum_{i=0}^{\infty} \frac{1}{n} \sum_{j=1}^{n} (a_{i,j})^{\frac{1}{2}} \right), \quad (39)$$

yielding (12), where

$$\underline{X}_{n,t} = u_t \left(\sum_{k=0}^{\infty} (\frac{1}{n} \sum_{i=1}^{n} \omega_i^{\frac{1}{2}} \alpha_i^{\frac{k}{2}})^2 \prod_{j=1}^{k} u_{t-j}^2 \right)^{\frac{1}{2}}$$

$$\overline{X}_{n,t} = u_t \left(\sum_{k=0}^{\infty} \frac{1}{n} \sum_{i=1}^{n} \omega_i^{\frac{1}{2}} \alpha_i^{\frac{k}{2}} \prod_{j=1}^{k} \mid u_{t-j} \mid \right).$$

(i) Setting $\bar{\chi}_a = (1/n \sum_{i=1}^{n} \omega_i^{\frac{1}{2}} \alpha_i^a)$ and $\chi_a = E(\bar{\chi}_a)$, using a version of the law of iterated logarithms (Stout 1974, Corollary 5.2.1): as $n \to \infty$

$$\bar{\chi}_a - \chi_a \sim \frac{2^{\frac{1}{2}}\, var^{\frac{1}{2}}(\omega_i^{\frac{1}{2}} \alpha_i^a)}{n^{\frac{1}{2}}} (\log \log (n\, var(\omega_i^{\frac{1}{2}} \alpha_i^a)))^{\frac{1}{2}} \quad a.s., \tag{40}$$

and given $var^{\frac{1}{2}}(\omega_i^{\frac{1}{2}} \alpha_i^a) \leq E^{\frac{1}{2}}(\omega_i \alpha_i^{2a}) \sim c\, \gamma^a\, a^{\frac{-(b_\gamma+\eta+1)}{2}}$, as $a \to \infty$ by (33), yields

$$\begin{aligned} & E_n \mid \underline{X}_{n,t} - \underline{X}_t \mid \\ \leq\ & \frac{\mu_1}{\rho_{\frac{1}{2}}\phi_0} \sum_{k=0}^{\infty} \mid (\bar{\chi}_{\frac{k}{2}})^2 - (\chi_{\frac{k}{2}})^2 \mid \mu_2^{\frac{k}{2}} = O\left((\frac{log\, log\, n}{n})^{\frac{1}{2}} \sum_{k=0}^{\infty} k^{-(\beta+\eta+1)} (\gamma\, \mu_2)^k \right), \ a.s. \\ & E_n \mid \overline{X}_{n,t} - \overline{X}_t \mid \\ \leq\ & \mu_1 \sum_{k=0}^{\infty} \mid \bar{\chi}_{\frac{k}{2}} - \chi_{\frac{k}{2}} \mid \mu_1^k = O\left((\frac{log\, log\, n}{n})^{\frac{1}{2}} \sum_{k=0}^{\infty} k^{\frac{-(\beta+\eta+1)}{2}} (\gamma^{\frac{1}{2}} \mu_1)^k \right), \ a.s., \end{aligned}$$

where the first inequality is obtained using (set $\underline{X}_{n,t} = u_t \underline{\sigma}_{n,t}$ and $\underline{X}_t = u_t \underline{\sigma}_t$)

$$\mid \underline{X}_{n,t} - \underline{X}_t \mid = \mid u_t \mid \frac{\mid \underline{\sigma}_{n,t}^2 - \underline{\sigma}_t^2 \mid}{\underline{\sigma}_{n,t} + \underline{\sigma}_t} \leq \frac{\mid u_t \mid}{\rho_{\frac{1}{2}} \phi_0} (\mid \underline{\sigma}_{n,t}^2 - \underline{\sigma}_t^2 \mid),$$

and

$$(\bar{\chi}_{\frac{k}{2}})^2 - (\chi_{\frac{k}{2}})^2 = (\bar{\chi}_{\frac{k}{2}} - \chi_{\frac{k}{2}})^2 + 2\, \chi_{\frac{k}{2}} (\bar{\chi}_{\frac{k}{2}} - \chi_{\frac{k}{2}}).$$

Thus, as $n \to \infty$, $\underline{X}_{n,t} \to_1 \underline{X}_t$ when $\gamma^{\frac{1}{2}} \mu_1 < 1$ and $\overline{X}_{n,t} \to_1 \overline{X}_t$ when $\gamma \mu_2 < 1$. When $max[\gamma^{\frac{1}{2}} \mu_1, \gamma \mu_2] = 1$, a sufficient condition for convergence in mean is $b_\gamma + \eta > 1$.

(ii) We adapt the proof of Nelson (1990, Theorem 2). Applying (35) to (13) and (14), it follows that

$$\mid \underline{X}_t \mid < \infty, \ \mid \overline{X}_t \mid < \infty, \quad a.s.$$

when $\gamma\, e^{\frac{\mu_0}{2}} < 1$. When $\gamma\, e^{\frac{\mu_0}{2}} = 1$ then boundedness is ensured when $\delta > 0$. Note that for $\overline{X}_t$ one must use $Elog \mid u_t \models \mu_0/2$. Strict stationarity and ergodicity follows using Stout (1974, Theorem 3.5.8) and Royden (1980, Proposition 5 and Theorem 3), by the same arguments used in Nelson (1990, p.329).

(iii) We discuss only the memory properties of the limit squares, as the covariance stationarity conditions for levels and squares follow easily. Consider $\underline{X}_t$. Set, for simplicity's sake, $\tilde{\omega}_i = 1$ with no harm. For integer $u > 0$,

$$cov(\underline{X}_t^2, \underline{X}_{t+u}^2) = \sum_{k,r=0}^{\infty} \phi_{\frac{k}{2}}^2 \phi_{\frac{r}{2}}^2 cov(u_t^2 \prod_{j=1}^{k} u_{t-j}^2, u_{t+u}^2 \prod_{s=1}^{r} u_{t+u-s}^2),$$

and, by means of the cumulants' theorem (see Leonov and Shiryaev (1959)) one easily obtains

$$cov(u_t^2 \prod_{j=1}^{k} u_{t-j}^2, u_{t+u}^2 \prod_{s=1}^{r} u_{t+u-s}^2) = E(\prod_{s=0}^{u-1} u_{t+u-s}^2) \times \tag{41}$$

$$\left(var(u_t^2) E(\prod_{j=1}^{k} u_{t-j}^2) E(\prod_{s=u+1}^{r} u_{t+u-s}^2) + E(u_t^4) cov(\prod_{j=1}^{k} u_{t-j}^2, \prod_{s=u+1}^{r} u_{t+u-s}^2) \right),$$

taking $r > u - 1$ for otherwise the left hand side of (41) vanishes. Easy calculations yields

$$cov(\underline{X}_t^2, \underline{X}_{t+u}^2)$$
$$= \mu_2^u \left(\sum_{k=0}^{\infty} \phi_{\frac{k}{2}}^2 \mu_2^k \right) \left(\sum_{r=0}^{\infty} \phi_{\frac{r+u}{2}}^2 \mu_2^r \right) + \mu_2^u \sum_{k=0}^{\infty} \phi_{\frac{k}{2}}^2 \phi_{\frac{k+u}{2}}^2 (\mu_4^k - \mu_2^{2k}) \tag{42}$$
$$+ \mu_2^u \sum_{k=1}^{\infty} \sum_{r=0}^{k-1} \phi_{\frac{k}{2}}^2 \phi_{\frac{r+u}{2}}^2 \mu_4^r \mu_2^{k-r} + \mu_2^u \sum_{r=1}^{\infty} \sum_{k=0}^{r-1} \phi_{\frac{k}{2}}^2 \phi_{\frac{r+u}{2}}^2 \mu_4^k \mu_2^{r-k}. \tag{43}$$

Recall that by (32)

$$\phi_k \sim c \, \gamma^k \, k^{-(\delta+1)}, \quad k \to \infty.$$

For the first term of (42), given $\sum_{r=0}^{\infty} = \sum_{r=0}^{u} + \sum_{r=u+1}^{\infty}$ and using Gradshteyn and Ryzhik (1994, # 3.381-3 and # 8.357), yields

$$\mu_2^u \left(\sum_{k=0}^{\infty} \phi_{\frac{k}{2}}^2 \mu_2^k \right) \left(\sum_{r=0}^{\infty} \phi_{\frac{r+u}{2}}^2 \mu_2^r \right) \sim c \, (\mu_2 \gamma)^u \, u^{-2(\delta+1)}, \quad u \to \infty.$$

Note that $\mu_4 \gamma^2 \le 1$ implies $\mu_2 \gamma < 1$, given $\mu_2 < \sqrt{\mu_4}$. Next, along the same lines, for the second term of (42)

$$\mu_2^u \sum_{k=0}^{\infty} \phi_{\frac{k}{2}}^2 \phi_{\frac{k+u}{2}}^2 (\mu_4^k - \mu_2^{2k}) \sim c \, (\mu_2 \gamma)^u \, u^{-2(\delta+1)} (1 + 1_{\gamma^2 \mu_4 = 1} u^{-2\delta-1}), \quad u \to \infty,$$

where now one needs to distinguish between case $\gamma^2\mu_4 < 1$ and case $\gamma^2\mu_4 = 1$. By similar calculations, it turns out that the two terms in (43) are of smaller order, bounded by

$$O\left(u^{-4(\delta+1)}((\mu_2\gamma)^{2u} + (\mu_4\gamma^2)^u\,(\mu_2\gamma)^u)\right),\ \ u \to \infty.$$

By means of tedious calculations, noting that

$$\overline{X}_t^2 = \underline{X}_t^2 + u_t^2 \sum_{\substack{k_1 \neq k_2 \\ =0}}^{\infty} \phi_{\frac{k_1}{2}}\phi_{\frac{k_2}{2}} \prod_{j=1}^{k_1}\prod_{s=1}^{k_2} |\,u_{t-j}\,||\,u_{t+u-s}\,|,$$

the same applies to the ACF of $\overline{X}_t^2$. □

Proof of Theorem 5. Using (39)

$$\min[\underbrace{X}_{n,t}, \overbrace{X}_{n,t}] \leq {}^U X_{n,t} \leq \max[\underbrace{X}_{n,t}, \overbrace{X}_{n,t}]$$

where

$$\underbrace{X}_{n,t} = u_t\left(\sum_{k=0}^{\infty}(\frac{1}{n}\sum_{i=1}^{n}\omega_i^{\frac{1}{2}}\prod_{j=1}^{k}(\beta_i + \alpha_i u_{t-j}^2)^{\frac{1}{2}})^2\right)^{\frac{1}{2}},$$

$$\overbrace{X}_{n,t} = u_t\left(\sum_{k=0}^{\infty}\frac{1}{n}\sum_{i=1}^{n}\omega_i^{\frac{1}{2}}\prod_{j=1}^{k}(\beta_i + \alpha_i u_{t-j}^2)^{\frac{1}{2}}\right).$$

Next, using (see Zaffaroni (1999, section 2))

$$\prod_{j=1}^{l}(\beta_i + \alpha_i u_{t-j}^2) = 1_{l=0} + 1_{l>0}\sum_{k=0}^{l}\alpha_i^k\beta_i^{l-k}\left(\widetilde{\sum}_{(k)}^{(l)} u_{t-j_1}^2 ... u_{t-j_1-...-j_k}^2\right), \quad (44)$$

with

$$\widetilde{\sum}_{(k)}^{(l)} = 1_{k=0} + 1_{k>0}\sum_{j_1=1}^{l-k+1}\sum_{j_2=1}^{l-k+2-j_1} ... \sum_{j_k=1}^{l-j_1-..-j_{k-1}},$$

and using (39) once again, yields

$$\underline{X}_{n,t} = u_t\left(\sum_{l=0}^{\infty}\sum_{k=0}^{l}\widetilde{\sum}_{(k)}^{(l)} u_{t-j_1}^2 ... u_{t-j_1-...-j_k}^2 (\frac{1}{n}\sum_{i=1}^{n}\omega_i^{\frac{1}{2}}\alpha_i^{\frac{k}{2}}\beta_i^{\frac{l-k}{2}})^2\right)^{\frac{1}{2}},$$

$$\overline{X}_{n,t} = u_t\left(\sum_{l=0}^{\infty}\sum_{k=0}^{l}\widetilde{\sum}_{(k)}^{(l)} |\,u_{t-j_1}\,| ... |\,u_{t-j_1-...-j_k}\,| \,(\frac{1}{n}\sum_{i=1}^{n}\omega_i^{\frac{1}{2}}\alpha_i^{\frac{k}{2}}\beta_i^{\frac{l-k}{2}})\right).$$

(i) Set

$$\underline{X}_t = u_t \left(\sum_{l=0}^{\infty} \sum_{k=0}^{l} \widetilde{\sum}_{(k)}^{(l)} u_{t-j_1}^2 ... u_{t-j_1-...-j_k}^2 (E(\omega_i^{\frac{1}{2}} \alpha_i^{\frac{k}{2}} \beta_i^{\frac{l-k}{2}}))^2 \right)^{\frac{1}{2}}, \quad (45)$$

$$\overline{X}_t = u_t \left(\sum_{l=0}^{\infty} \sum_{k=0}^{l} \widetilde{\sum}_{(k)}^{(l)} \mid u_{t-j_1} \mid ... \mid u_{t-j_1-...-j_k} \mid E(\omega_i^{\frac{1}{2}} \alpha_i^{\frac{k}{2}} \beta_i^{\frac{l-k}{2}}) \right). \quad (46)$$

We follow the proof of Theorem 4. Applying (40) to the sequence $\{\omega_i^{\frac{1}{2}} \alpha_i^{\frac{k}{2}} \beta_i^{\frac{l-k}{2}}\}$ yields, as $n \to \infty$,

$$\begin{aligned} & \mid \frac{1}{n} \sum_{i=1}^{n} \omega_i^{\frac{1}{2}} \alpha_i^{\frac{k}{2}} \beta_i^{\frac{l-k}{2}} - E(\omega^{\frac{1}{2}} \alpha^{\frac{k}{2}} \beta^{\frac{l-k}{2}}) \mid \\ &= O\left((\frac{\rho_1 \, log \, log \, n}{n})^{\frac{1}{2}} \, E^{\frac{1}{2}}[(1-\mu_2\alpha_i - \beta_i)^\eta \alpha_i^k \beta_i^{(l-k)}] \right) \\ &= O\left((\frac{log \, log \, n}{n})^{\frac{1}{2}} \bar{\alpha}^{\frac{k}{2}} \bar{\beta}^{\frac{l-k}{2}} \, l^{-\frac{1}{2}(min[b_\alpha, c_\beta]+\eta+1)} \right), \quad a.s., \end{aligned} \quad (47)$$

for $0 \leq k \leq l$. The last bound is obtained as follows.

$$\begin{aligned} & E[(1-\mu_2\alpha_i - \beta_i)^\eta \alpha_i^k \beta_i^{l-k}] \\ & \sim c \int_0^{\bar{\alpha}} \alpha^k (\bar{\alpha}-\alpha)^{b_\alpha} d\alpha \int_0^{\bar{\beta}} (1-\mu_2\alpha-\beta)^\eta \beta^{l-k} (\bar{\beta}-\beta)^{c_\beta} d\beta \\ &= c \, \bar{\beta}^{l-k} \int_0^{\bar{\alpha}} \alpha^k (\bar{\alpha}-\alpha)^{b_\alpha} d\alpha \int_0^1 (1-\mu_2\alpha-\bar{\beta}t)^\eta t^{l-k} (1-t)^{c_\beta} dt \\ &= c \, \bar{\beta}^{l-k} \int_0^{\bar{\alpha}} \alpha^k (\bar{\alpha}-\alpha)^{b_\alpha} (1-\mu_2\alpha)^\eta d\alpha \int_0^1 (1-\theta(\alpha)t)^\eta t^{l-k} (1-t)^{c_\beta} dt, \end{aligned}$$

by means of the change of variable $t = \beta/\bar{\beta}$, setting for simplicity's sake $\theta(\alpha) = \bar{\beta}/(1-\mu_2\alpha)$, where the constant c is not always the same. Note that by assumption $\theta(\alpha) < 1$ with $\theta(\alpha) \to 1^-$ as $\alpha \to \bar{\alpha}^-$. Using Gradshteyn and Ryzhik (1994, # 3.197-3) for the second integral of the right hand side (in t) gives

$$c \, \bar{\beta}^{l-k} B(l-k+1, c_\beta+1) \int_0^{\bar{\alpha}} \alpha^k (\bar{\alpha}-\alpha)^{b_\alpha} (1-\mu_2\alpha)^\eta {}_2F_1(-\eta, l-k+1; l-k+c_\beta+2; \theta(\alpha)) d\alpha, \quad (48)$$

where $B(\cdot,\cdot)$ is the Beta function and ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ denotes the hypergeometric function (Gradshteyn and Ryzhik 1994, section 9.1). As $\alpha \to \bar{\alpha}^-$, by Gradshteyn and Ryzhik (1994, # 9.122-1),

$$ {}_2F_1(-\eta, l-k+1; l-k+c_\beta+2; \theta(\alpha)) \sim c\frac{\Gamma(l-k+2+c_\beta)}{\Gamma(l-k+c_\beta+2+\eta)}, $$

yielding, in (48), with the change of variable $t = \alpha/\bar{\alpha}$,

$$ \begin{aligned} & c\frac{\Gamma(l-k+1)}{\Gamma(l-k+c_\beta+2+\eta)}\bar{\beta}^{l-k}\bar{\alpha}^k \int_0^1 t^k(1-t)^{b_\alpha}(1-\mu_2\bar{\alpha}t)^\eta dt \\ & \sim c\frac{\Gamma(l-k+1)}{\Gamma(l-k+c_\beta+2+\eta)}\bar{\beta}^{l-k}\bar{\alpha}^k \int_0^1 t^k(1-t)^{b_\alpha} dt \\ & \leq c\bar{\beta}^{l-k}\bar{\alpha}^k k^{-(b_\alpha+1)}(l-k)^{-(c_\beta+\eta+1)}, \end{aligned} $$

using, for the last inequality, Stirling's formula (Brockwell and Davis 1987, p.522) to expand the gamma function $\Gamma(\cdot)$. Repeating the same arguments but starting to integrate with respect to α finally yields

$$ E[(1-\mu_2\alpha_i-\beta_i)^\eta \alpha_i^k \beta_i^{l-k}] = \begin{cases} O\left(\bar{\beta}^{l-k}\,\bar{\alpha}^k\,(l-k)^{-(min[b_\alpha,c_\beta]+\eta+1)}\right), & 0 \leq k < l/2, \\ O\left(\bar{\beta}^{l-k}\,\bar{\alpha}^k\,k^{-(min[b_\alpha,c_\beta]+\eta+1)}\right), & l/2 \leq k \leq l. \end{cases} \tag{49} $$

Consider $\overline{X}_{n,t}$. Then

$$ \begin{aligned} & E_n\,|\overline{X}_{n,t} - \overline{X}_t| \\ = \;& O\left((\frac{log\,log\,n}{n})^{\frac{1}{2}} \sum_{l=0}^{\infty} [\sum_{k=0}^{l} \binom{l}{k} \mu_1^k\,\bar{\alpha}^{\frac{k}{2}}\,\bar{\beta}^{\frac{l-k}{2}}\,l^{-\frac{1}{2}(min[b_\alpha+c_\beta]+\eta+1)}] \right) \\ = \;& O\left((\frac{log\,log\,n}{n})^{\frac{1}{2}} \sum_{l=0}^{\infty} l^{-\frac{1}{2}(min[b_\alpha,c_\beta]+\eta+1)} (\bar{\alpha}^{\frac{1}{2}}\,\mu_1 + \bar{\beta}^{\frac{1}{2}})^l \right) = O\left((\frac{log\,log\,n}{n})^{\frac{1}{2}} \right), \quad a.s., \end{aligned} $$

under the stated conditions. For $\underline{X}_{n,t}$, simply adapt the proof of Theorem 4, using (47).

(ii) Boundedness, strict stationarity and ergodicity easily follows by using the same arguments as in the proof of Theorem 4.

(iii) The asymptotic covariance stationarity conditions, for levels and squares, easily follow. Let us focus on the memory properties only. Consider $\underline{X}_{n,t}$. For simplicity's sake let us set $\tilde{\omega}_i = 1$ as this is completely innocuous. Then, for any integer $u > 0$, setting $\delta_{ij}(t) = (\beta_i\beta_j)^{\frac{1}{2}} + (\alpha_i\alpha_j)^{\frac{1}{2}} u_t^2$ and $\delta_{ij} = E_n(\delta_{ij}(t))$, using (44) backward and using the cumulants' theorem,

$$\mathrm{cov}_n(\underline{X}_{n,t}^2, \underline{X}_{n,t+u}^2) = \frac{\mu_2}{n^4} \sum_{i,ja,b=1}^{n} \omega_i^{\frac{1}{2}} \omega_j^{\frac{1}{2}} \omega_a^{\frac{1}{2}} \omega_b^{\frac{1}{2}} \delta_{ab}^{u-1} (A+B),$$

with

$$A = \mathrm{cov}_n(u_t^2, \delta_{ab}(t)) \sum_{l_1=0}^{\infty} E_n(\prod_{r_1=1}^{l_1} \delta_{ij}(t-r_1)) \sum_{l_2=0}^{\infty} E_n(\prod_{r_2=1}^{l_2} \delta_{ab}(t-r_2))$$

$$B = E_n(u_t^2 \delta_{ab}(t)) \sum_{l_1=0}^{\infty} \sum_{l_2=0}^{\infty} \mathrm{cov}_n(\prod_{r_1=1}^{l_1} \delta_{ij}(t-r_1)), \prod_{r_2=1}^{l_2} \delta_{ab}(t-r_2)).$$

For A,

$$A = \mu_4(\alpha_a\, \alpha_b)^{\frac{1}{2}} \frac{1}{1-\delta_{ab}} \frac{1}{1-\delta_{ij}}.$$

However, by the bounded fourth moment condition

$$\delta_{ij} \leq \bar{\pi} < 1,$$

yielding, as $n \to \infty$,

$$\frac{1}{n^4} \sum_{i,ja,b=1}^{n} \omega_i^{\frac{1}{2}} \omega_j^{\frac{1}{2}} \omega_a^{\frac{1}{2}} \omega_b^{\frac{1}{2}} \delta_{ab}^{u-1} A \sim cE\left(\omega_a^{\frac{1}{2}} \omega_b^{\frac{1}{2}} \delta_{ab}^{u-1}\right)$$

Given the binomial formula, $(a+b)^u = \sum_{s=0}^{u} \binom{u}{s} a^s b^{u-s}$, and using (49) (replacing η, k, l with $\eta/2$, $s/2, u/2$) yields

$$E\left(\omega_a^{\frac{1}{2}} \omega_b^{\frac{1}{2}} \delta_{ab}^{u-1}\right) = O\left(\bar{\pi}^u u^{-2(\delta+1)}\right),\ u \to \infty. \tag{50}$$

For the other term, involving B, it turns out that these will be characterized by the same asymptotic behaviour as for (50). Note that the ACF of $\underline{X}_{n,t}^2$ is nonnegative, given that $cov(\delta_{ij}(t), \delta_{ab}(t)) > 0$ (u_t^2 not degenerate) for any i, j, a, b, and using the following result for sequence of independent r.vs $\{C_i, D_i\}$:

$$cov(\prod_{i=1}^{m} C_i, \prod_{i=1}^{m} D_i) = \sum_{k=1}^{m} \prod_{j=1}^{k-1} E(C_j D_j) cov(C_k, D_k) E(\prod_{j=k+1}^{m} C_j) E(\prod_{j=k+1}^{m} D_j),$$

which holds when the left hand side is well-defined and bounded. Thus, given that $var(Z) \leq E(Z^2)$ for any r.v. Z with bounded second moment,

$$B \leq E_n(u_t^2 \delta_{ab}(t)) \left(\sum_{l=0}^{\infty} \kappa_{abij}^l + \sum_{l_1=1}^{\infty} \sum_{l_2=0}^{l_1-1} \delta_{ij}^{l_1-l_2} \kappa_{abij}^{l_2} + \sum_{l_2=1}^{\infty} \sum_{l_1=0}^{l_2-1} \delta_{ab}^{l_2-l_1} \kappa_{abij}^{l_1} \right), \tag{51}$$

setting

$$\kappa_{abij} = (\beta_i \beta_j \beta_a \beta_b)^{\frac{1}{2}} + \mu_4 (\alpha_i \alpha_j \alpha_a \alpha_b)^{\frac{1}{2}} + \mu_2 (\alpha_i \alpha_j \beta_a \beta_b)^{\frac{1}{2}} + \mu_2 (\alpha_a \alpha_b \beta_i \beta_j)^{\frac{1}{2}}.$$

Consider the first term on the right hand side of (51). Using

$$\begin{aligned} &\delta_{ab}^u \, \kappa_{abij}^l \\ &= \sum_{s=0}^{u} \binom{u}{s} \sum_{k=0}^{l} \binom{l}{k} \sum_{r_1=0}^{k} \binom{k}{r_1} \sum_{r_2=0}^{l-k} \binom{l-k}{r_2} \times \\ &\mu_2^{u-s+l-k} \mu_4^k (\beta_a \beta_b)^{\frac{s+r_1+l-k-r_2}{2}} (\alpha_a \alpha_b)^{\frac{u-s+k-r_1+r_2}{2}} (\beta_i \beta_j)^{\frac{r_1+r_2}{2}} (\alpha_i \alpha_j)^{\frac{l-r_1-r_2}{2}}, \end{aligned}$$

obtained by repeated use of the binomial formula, and suitably applying (49), yields, as $n \to \infty$,

$$\frac{1}{n^4} \sum_{i,j a,b=1}^{n} \omega_i^{\frac{1}{2}} \omega_j^{\frac{1}{2}} \omega_a^{\frac{1}{2}} \omega_b^{\frac{1}{2}} \delta_{ab}^{u-1} \sum_{l=0}^{\infty} \kappa_{abij}^l \sim cE \left(\omega_a^{\frac{1}{2}} \omega_b^{\frac{1}{2}} \delta_{ab}^{u-1} (1 + u^{1-2(\delta+1)}) \right).$$

Note that the covariance stationarity conditions for the $\underline{X}_t$ requires $2(\delta+1) < 1$, yielding

$$E \left(\omega_a^{\frac{1}{2}} \omega_b^{\frac{1}{2}} \delta_{ab}^{u-1} (1 + u^{1-2(\delta+1)}) \right) \sim E \left(\omega_a^{\frac{1}{2}} \omega_b^{\frac{1}{2}} \delta_{ab}^{u-1} \right), \; u \to \infty.$$

For the other two terms, on the right hand side of (51), with easy but tedious calculations, the result follows. □

References

Baillie, R. T., T. Bollerslev, and H. O. A. Mikkelsen (1996): "Fractionally integrated generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 74/1, 3–30.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31, 302–327.

Bollerslev, T., R. Chou, and K. Kroner (1992): "ARCH modeling in finance: a review of the theory and empirical evidence," *Journal of Econometrics*, 52, 5–59.

Breidt, F., N. Crato, and de Lima P. (1998): "Modelling long-memory stochastic volatility," *Journal of Econometrics*, 73, 325–334.

Brockwell, P., and R. Davis (1987): *Time series: theory and methods.* New York: Springer-Verlag.

Chamberlain, G., and M. Rothschild (1983): "Arbitrage, factor structure, and mean-variance analysis on large asset markets," *Econometrica*, 51, 1281–1304.

Comte, F., and E. Renault (1998): "Long memory in continuous-time stochastic volatility models," *Mathematical Finance*, 8/4, 291–323.

Davis, R., and T. Mikosch (1998): "The sample autocorrelations of heavy-tailed processes with applications to ARCH," *Annals of Statistics*, 26, 2049–2080.

Diebold, F. (1988): *Empirical modeling of exchange rate dynamics.* Berlin: Springer-Verlag.

Diebold, F., and M. Nerlove (1989): "The dynamics of exchange rate volatility: a multivariate latent factor ARCH model," *Journal of Applied Econometrics*, 4, 1–21.

Ding, Z., and C. Granger (1996): "Modeling volatility persistence of speculative returns: a new approach," *Journal of Econometrics*, 73, 185–215.

DING, Z., C. W. J. GRANGER, AND R. F. ENGLE (1993): "A long memory property of stock market and a new model," *Journal of Empirical Finance*, 1, 83–106.

DROST, F., AND T. NIJMAN (1993): "Temporal aggregation of GARCH processes," *Econometrica*, 61/4, 909–927.

DUDLEY, R. (1989): *Real analysis and probability.* California: Pacific Grove.

ENGLE, R. (1987): "Multivariate ARCH with factor structures-cointegration in variance," *Preprint.*

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

FELLER, W. (1966): *An introduction to probability theory and its applications*, vol. II. New York: Wiley.

GIRAITIS, L., P. KOKOSKA, AND R. LEIPUS (1998): "Stationary ARCH models: dependence structure and central limit theorem," forthcoming *Econometric Theory.*

GRADSHTEYN, I., AND I. RYZHIK (1994): *Table of integrals, series and products.* San Diego: Academic Press, 5 edn.

GRANGER, C. (1980): "Long memory relationships and the aggregation of dynamic models," *Journal of Econometrics*, 14, 227–238.

GRANGER, C., AND M. MORRIS (1976): "Time series modelling and interpretation," *Journal of the Royal Statistical Society A*, 139, 246–257.

HANSEN, L., AND S. RICHARD (1987): "The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models," *Econometrica*, 55, 587–614.

HARDY, G., J. LITTLEWOOD, AND G. POLYA (1964): *Inequalities.* Cambridge (UK): Cambridge University Press.

HARVEY, A. (1998): "Long memory in stochastic volatility," in *Forecasting volatility in financial markets*, ed. by J. Knight,and S. Satchell. London: Butterworth-Heineman, forthcoming.

KING, M., E. SENTANA, AND S. WADHWANI (1994): "Volatility and links between national stock markets," *Econometrica*, 62/4, 901–933.

LEIPUS, R., AND M. VIANO (1999): "Aggregation in ARCH models," *Preprint*.

LEONOV, V., AND A. SHIRYAEV (1959): "On a method of calculation of semi-invariants," *Theory of Probability Applications*, 4, 319–329.

LEWBEL, A. (1994): "Aggregation and simple dynamics," *American Economic Review*, 84-4, 905–918.

LIPPI, M., AND P. ZAFFARONI (1998): "Aggregation of simple linear dynamics: exact asymptotic results," Econometrics Discussion Paper 350, STICERD-LSE.

MEDDAHI, N., AND E. RENAULT (1996): "Aggregation and marginalization of GARCH and Stochastic Volatility models," *Preprint*.

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

NIJMAN, T., AND E. SENTANA (1996): "Marginalization and contemporaneous aggregation in multivariate GARCH processes," *Journal of Econometrics*, 71, 71–87.

PHILLIPS, P., AND H. MOON (1999): "Linear regression limit theory for nonstationary panel data," *Econometrica*, 67, 1057–1112.

ROBINSON, P. M. (1978): "Statistical inference for a random coefficient autoregressive model," *Scandinavian Journal of Statistics*, 5, 163–168.

ROBINSON, P. M. (1991): "Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression," *Journal of Econometrics*, 47, 67–84.

ROBINSON, P. M., AND P. ZAFFARONI (1997): "Modelling nonlinearity and long memory in time series," in *Nonlinear Dynamics and Time Series*, ed. by C. Cutler,and D. Kaplan. Providence (NJ): American Mathematical Society.

ROBINSON, P. M., AND P. ZAFFARONI (1998): "Nonlinear time series with long memory: a model for stochastic volatility," *Journal of Statistical Planning and Inference*, 68, 359–371.

ROSS, S. (1976): "The arbitrage theory of capital asset pricing," *Journal of Economic Theory*, 17, 254–286.

ROYDEN, H. (1980): *Real analysis*. London: Macmillan.

SAMORODNITSKY, G., AND M. TAQQU (1994): *Stable non-Gaussian Processes: Stochastic Models with Infinite Variance*. New York, London: Chapman and Hall.

SENTANA, E. (1998): "The relation between conditionally heteroskedastic factor models and factor GARCH models," *Econometrics Journal*, 1, 1–9.

STOUT, W. (1974): *Almost sure convergence*. Academic Press.

TAQQU, M., W. WILLINGER, AND R. SHERMAN (1997): "Proof of a fundamental result in self-similar traffic modeling," *Computer Communication Review*, 27, 5–23.

TAUCHEN, G., AND M. PITTS (1983): "The price variability volume relationship on speculative markets," *Econometrica*, 51, 485–505.

ZAFFARONI, P. (1999): "Stationarity and memory of ARCH(∞) models," *Preprint*.

ZAFFARONI, P. (2001): "Aggregation and memory of models of changing volatility," *Preprint*.

ZYGMUND, A. (1977): *Trigonometric series*. Cambridge (UK): Cambridge University Press.

RECENTLY PUBLISHED "TEMI" (*)

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi – Divisione Biblioteca e pubblicazioni – Via Nazionale, 91 – 00184 Rome (fax 0039 06 47922059). They are available on the Internet at www.bancaditalia.it