Contemporaneous Aggregation of GARCH Processes

Paolo Zaffaroni *
Imperial College London
This draft: June 2006

Abstract

In this paper the effect of contemporaneous aggregation of heterogeneous GARCH processes as the cross-sectional size diverges to infinity is studied. We analyze both cases of cross-sectionally dependent and independent individual processes. The limit aggregate does not belong to the class of GARCH processes. Dynamic conditional heteroskedasticity is only preserved when the individual processes are sufficiently cross-correlated, although long memory for the limit aggregate volatility is not attainable. We also explore more general forms of cross-sectional dependence and various types of aggregation schemes.

Keywords: contemporaneous aggregation, GARCH, cross-section asymptotic, common and idiosyncratic risk, memory, factor models, value-weighted portfolio

1 Introduction

The autoregressive conditional heteroskedasticity (ARCH) model of Engle (1982) and the generalized ARCH (GARCH) development of Bollerslev (1986) are the most popular approaches used to describe the conditional heteroskedasticity observed in many financial time series.

Given that a large number of securities are traded in financial markets, a practical use of GARCH as models for asset returns has led to the need
of analyzing the effect of contemporaneous aggregation (henceforth aggregation) of GARCH, in the sense of summing or averaging across assets. See Nijman and Sentana (1996) and Meddahi and Renault (2004). The number of parameters of the exact aggregate model, based on $n$ units, increases as $O(2n)$, except for particular cases of non-heterogeneity across parameters, e.g. the sum of two GARCH(1,1) yields (weak) GARCH(2,2). Thus, estimation of the exact volatility process for the aggregate is cumbersome, if not impossible, except for small $n$. This same problem can also arise when modelling individual asset returns. See Ding and Granger (1996). In this paper we propose a different approach, based on these considerations.

Given the aggregate

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

(1)
of $n$ heterogeneous $x_{i,t}$, each parameterized as a GARCH, we establish the asymptotic limit (in a suitable norm) as $n \to \infty$ of the aggregate $X_{n,t}$, 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 (henceforth LA) maintains the GARCH nonlinearity, uncorrelated levels and correlated squares, conveying the basic features of a volatility model. In general, the LA will not be a GARCH. In contrast, the GARCH nonlinearity is lost for a weak degree of cross-correlation between the $x_{i,t}$. The (suitably normalized) LA has a stable distribution, equal to a Gaussian noise when stationarity occurs. We also characterize separately the asymptotic behaviour of the variance of $X_{n,t}$. Contrary to the commonly held view, even with perfectly stationary and mutually independent $x_{i,t}$, $X_{n,t}$ does not necessarily converge to zero (in mean-square). Unlike the approach of this paper, which looks at the limit of (1), Leipus and Viano (2002) and Kazakevičius,
Leipus, and Viano (2004) characterize the limit in mean and mean-square of $\sum_{i=1}^{n} w_i(n) x_{i,t}^2$, for \( \text{ARCH}(\infty) \) \( x_{i,t}^2 \) and coefficients of aggregation \( w_i(n) \).

\( \text{ARCH}(\infty) \) processes, extending the \( \text{ARCH}(q), q < \infty, \) process and the \( \text{GARCH}(p, q) \) were considered by Robinson (1991) as a class of parametric alternatives in testing for dynamic conditional heteroskedasticity.) Both papers assume that the ratio of \( x_{i,t}^2 \) to its conditional mean is constant across \( i \), a case here defined of ‘common innovations’.

The plan of the paper is as follows. In Section 2 we focus on aggregation of \( \text{GARCH}(1, 1) \). Definitions and assumptions are introduced in Section 2.1. Sections 2.2 and 2.3 focus on independent and common rescaled innovations respectively. A numerical example is reported in Section 2.4. Section 3.1 focuses on the volatility implication of the LA, invalidating Ding and Granger (1996)’s conjecture according to which the squares of the LA of \( \text{GARCH}(1, 1) \) exhibit the technical condition for long memory, in the sense of non-summable autocovariance function. More general forms of cross-sectional dependence are described in Section 3.2 focusing on dynamic conditionally heteroskedastic factor models. Section 3.3 considers forms of aggregation other than the equally weighted (1). For instance, we consider an extension of our results to value-weighted portfolios. Section 4 contains concluding remarks. All results are formally stated in theorems with proofs reported in the final appendix.

## 2 Aggregation of heterogeneous \( \text{GARCH}(1, 1) \)

In this section we focus on \( \text{GARCH}(1, 1) \) units \( x_{i,t} \), when both the parameters and the rescaled innovations are potentially varying across units:

\[
x_{i,t} = z_{i,t} \sigma_{i,t}, \quad i \in \mathbb{N}, \ t \in \mathbb{Z},
\]  

(2)
with
\[ \sigma_{i,t}^2 = \omega_i + \alpha_i x_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2 \text{ a.s.,} \] (3)
where a.s. means ‘almost surely’. Model (2) can be interpreted as a nonlinear dynamic panel data model with random effects. No parameter is assumed constant across units. As shown below, such random effects are assumed to be i.i.d. across units, draw from a common underlying distribution. Conditionally on the random parameter values, each \( x_{i,t} \) is a strong GARCH(1,1) for i.i.d. (across time) \( z_{i,t} \); see Drost and Nijman (1993, Definition 1). To better focus on the volatility implications, the assumption of martingale difference \( x_{i,t} \) is maintained throughout the paper.

### 2.1 Definitions and assumptions

We will consider the two ‘canonical’ cases: perfectly independent across units innovation \( z_{i,t} = \epsilon_{i,t} \), called idiosyncratic innovation; perfectly correlated across units innovation \( z_{i,t} = u_t \), called common innovation. These cases represent the building blocks used to evaluate the effect of aggregation for more general cases of cross-sectional dependence, examined in Section 3.2. Note that even in the common innovation case the \( x_{i,t} \) are not perfectly cross-correlated when \( \omega_i, \alpha_i, \beta_i \) are heterogeneous.

**Assumption I**

(i) The \( u_t \), called the common innovations, are i.i.d. across \( t \in \mathbb{Z} \) and the \( \epsilon_{i,t} \), called the idiosyncratic innovations, are i.i.d. across \( t \in \mathbb{Z} \) and \( i \in \mathbb{N} \), satisfying \( E(u_t) = E(\epsilon_{i,t}) = 0, 0 < E \left| u_t \right|^r = E \left| \epsilon_{i,t} \right|^r = \mu_r < \infty, \ (r = 1, 2, 4) \)
and \( \mu_0 = E \log \epsilon_{i,t}^2 = E \log u_t^2 \) is well defined.

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

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 $\gamma, \eta$ with $0 \leq \gamma, \eta < \infty$, we assume that the GARCH(1, 1) parameters satisfy the following.

**Assumption II($\gamma$)**

(i) The $\omega_i$ are i.i.d. with $\omega_i \geq \omega > 0$ a.s., constant $\omega$, and $E(\omega_i^2) < \infty$.

(ii) The $\alpha_i$ and the $\beta_j$ are mutually independent for any $i, j \in \mathbb{N}$.

(iii) The $\alpha_i$ and the $\beta_i$ are i.i.d. with absolutely continuous distribution in the interval $[\bar{\alpha}, \alpha)$ and $[\bar{\beta}, \beta)$, respectively, where

$$\mu_2 \bar{\alpha} + \bar{\beta} = \gamma,$$

depending upon the real parameters $b_{\bar{\alpha}}, b_{\bar{\beta}} > -1$, with densities

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

$$B(\beta_i; b_{\bar{\beta}}) \sim C(\bar{\beta} - \beta_i)^{b_{\bar{\beta}}}, \quad \beta_i \to \bar{\beta}^-.$$  

(iv) For some $0 < p < 1$ and $c_p = \sup \left\{ 0 < c < \infty \text{ such that } E \ln(c + |\epsilon_{i,0}|^{1+p}) < 0 \right\}$

$$\left( \left( \frac{\beta_i}{(c_p)^{1-p}} \right)^{\frac{1}{p}} + \alpha_i^{\frac{1}{p}} \right)^p \leq (\beta_i + \alpha_i) \text{ a.s.} \quad (6)$$

**Assumption III($\eta$)**

The $\omega_i, \alpha_i$ and $\beta_i$ satisfy $\omega_i = \bar{\omega}_i \mid 1 - \mu_2 \alpha_i - \beta_i \mid^\eta$, for an i.i.d. sequence $\bar{\omega}_i \geq \tilde{\omega} > 0$ a.s., constant $\tilde{\omega}$, mutually independent from the $\alpha_i$ and $\beta_i$, and such that $E(\tilde{\omega}_i^2) < \infty$.

**Remarks.**

(a) Part (iii) of Ass. II($\gamma$) describes a mild semiparametric specification of the density function of the $\alpha_i, \beta_i$. Conditions $b_{\bar{\alpha}} > -1, b_{\bar{\beta}} > -1$ are required for integrability. An extremely wide variety of parametric specifications $B(\cdot; \theta)$ for $\theta \in \Theta \subset \mathbb{R}^p$ is allowed for (4) and (5), such as the Beta distribution a special case of which is the uniform distribution.
(b) Robinson (1978) considered conditions such as (4) and (5), including the 
Beta distribution case, in order to develop the statistical properties of a cer-
tain estimation procedure for a random coefficient autoregressive model.

(c) (4) could be generalized, with no effect on the results, to 
\[ B(\alpha_i; b_{\bar{\alpha}}) \sim C (\bar{\alpha} - \alpha_i)^{b_{\bar{\alpha}}} L\left(\frac{1}{\alpha - \alpha_i}\right), \quad \alpha_i \to \bar{\alpha} \] 
where \( L(\cdot) \) denotes a slowly varying function: 
\[ L(tx)/L(t) \to 1 \text{ as } t \to \infty, \text{ any } x > 0 \text{ (see Zygmund (1977)).} \] 
The same applies to \( B(\beta_i; b_{\bar{\beta}}) \).

(d) Condition (i) of Ass. II(\( \gamma \)) rules out the possibility that \( \sigma^2_{i,t} = 0 \) a.s., 
which in turn holds when \( \omega_i = 0 \) (see Nelson (1990)).

(e) We focus on covariance stationary \( x_{i,t} \). This is much stronger than con-
sidering strictly stationary \( x_{i,t} \), which just require \( E\log(\beta_i + \alpha_i z^2_{i,t}) < 0 \) 
(Nelson 1990, Theorem 2).

(f) Assumption (6) is a technical condition, depending on the support of the 
distribution of the \( \alpha_i, \beta_i \) and on the distribution of the \( \epsilon_{i,t} \). For example, for 
Gaussian \( \epsilon_{i,t} \) with \( \mu_2 = 1 \), then \( c_p = 0.4095 \ldots \) for \( p = 0.9 \) and (6) is satisfied 
for \( [\alpha_i, \bar{\alpha}] \times [\beta_i, \bar{\beta}] = [0.30, 1] \times [0, 0.60] \). Set 
\[ \mu^{(p)}_0 = E\log(c_p + |\epsilon_{i,0}|^{2/p}). \]
Then by construction \( \sup_c E\ln(c + |\epsilon_{i,0}|^{2/p}) \leq \mu^{(p)}_0 < 0. \)

(g) Case \( \eta = 0 \) of Ass. III(\( \eta \)) implies that the \( \omega_i \) and the \( \alpha_i, \beta_i \) are mutually 
independent. When \( \eta > 0 \), we impose a form of negative contemporaneous 
dependence between the \( \omega_i \) and the \( \pi_i \). In all other possible cases of depen-
dence, the same results obtained for case \( \eta = 0 \) apply.

(h) Ass. III(\( \eta \)) could be substantially weakened to 
\[ E(\omega_i | \alpha_i) \sim c \left|1 - \mu_2 \alpha_i - \beta_i\right|^\eta, \quad \text{as } \mu_2 \alpha_i + \beta_i \to \gamma^- \]
without specifying the degree of dependence when \( \mu_2 \alpha_i + \beta_i \) is well below \( \gamma \).

(i) Ass. I can be generalized to allow for heterogeneity across moments of
the idiosyncratic innovation, such as $E \left| \epsilon_{i,t} \right|^r = \mu_{ri}$. When $r = 2$ the driving parameters become $(\omega^*_{i}, \alpha^*_{i}, \beta_{i})$ with $\omega^*_{i} = \omega_{i} \sqrt{\mu_{2i}}$, $\alpha^*_{i} = \alpha_{i} \mu_{2i}$ and the driving innovation $\epsilon^*_{i,t} = \epsilon_{i,t}/\sqrt{\mu_{2i}}$ but with no changes of the results.

Set

$$\pi_{i} = (\mu_{2} \alpha_{i} + \beta_{i}),$$

(7)

$$\nu_{i} = (\mu_{2} \alpha_{i} + \beta_{i})^{2} + (\mu_{4} - \mu_{2}^{2}) \alpha_{i}^{2} = \pi_{i}^{2} + (\mu_{4} - \mu_{2}^{2}) \alpha_{i}^{2}.$$  

(8)

For finite $n$, covariance stationary levels $X_{n,t}$ require $\bar{\pi} = \bar{\beta} + \mu_{2} \bar{\alpha} \leq 1$ and covariance stationary squares $X_{n,t}^{2}$ require $\bar{\nu} = (\bar{\beta} + \mu_{2} \bar{\alpha})^{2} + (\mu_{4} - \mu_{2}^{2}) \bar{\alpha}^{2} \leq 1$. Asymptotic covariance stationarity, as $n \to \infty$, will also require that the distribution of the $\pi_{i}$ and the $\nu_{i}$ be not too dense around $\bar{\pi}$ and $\bar{\nu}$ respectively. For instance, under Ass. II(\gamma) and some additional regularity conditions it is easy to see that the $\pi_{i}$ have an absolutely continuous distribution with density satisfying

$$B(\pi_{i}; b_{\pi}) \sim C(\bar{\pi} - \pi_{i})^{b_{\pi}}, \quad \pi_{i} \to \pi^{-}$$

(9)

where $b_{\pi} = b_{\alpha} + b_{\beta} + 1$ (see Zaffaroni (2000, Lemma 3)). Likewise

$$B(\nu_{i}; b_{\nu}) \sim C(\bar{\nu} - \nu_{i})^{b_{\nu}}, \quad \nu_{i} \to \nu^{-}$$

(10)

where $b_{\nu} = b_{\alpha}/2 + b_{\alpha}/2 + 1$. Some of the results of this paper depend directly on the distribution of the $\pi_{i}$ and $\nu_{i}$, with no separate roles for $\alpha_{i}$ and $\beta_{i}$, whereas others, instead, are more easily attained considering the distribution of the $\alpha_{i}$ and $\beta_{i}$ separately. We thus preferred to maintain the more primitive Ass. II(\gamma) rather than assume (9) and (10) directly.

From now on, we will denote the conditional expectation and conditional variance operators, given the GARCH coefficients, by $E_{n}(\cdot)$ and $\text{var}_{n}(\cdot)$ respectively. Finally, the $\to_{a.s.}$, $\to_{p}$, $\to_{r}$ and $\to_{d}$ denote convergence almost
sure, in probability, in rth mean, and convergence in the sense of the finite-dimensional distribution, respectively. In the following Theorems, we will always assume that Assumptions I, with \( \mu_2 = 1, \mu_4 = 3, II(\gamma) \), with \( \gamma \leq 1, \) and \( III(\eta) \), hold without stating this explicitly.

### 2.2 Idiosyncratic innovations

For the aggregate \( E X_{n,t} = n^{-1} \sum_{i=1}^{n} \epsilon_{i,t} \sigma_{i,t} \), simple calculation yields

\[
\text{var}_n(E X_{n,t}) = \frac{1}{n^2} \sum_{i=1}^{n} \frac{\omega_i}{1 - \alpha_i - \beta_i}.
\]  

(11)

The following theorem describes the asymptotic behaviour of \( \text{var}_n(E X_{n,t}) \).

**Theorem 1** As \( n \to \infty \):

(i) When \( \gamma < 1 \) and for any \( \eta \), \( \text{var}_n(E X_{n,t}) \longrightarrow \text{a.s.} \ 0 \).

(ii) When \( \gamma = 1 \) and \( \eta = 0 \), setting \( \delta = b_{\eta} + 1 \) with \( b_{\eta} \) defined in (9):

if \( b_{\eta} > -1/2 \), \( \text{var}_n(E X_{n,t}) \longrightarrow \text{a.s.} \ 0 \);  

if \( b_{\eta} \leq -1/2 \), \( n^{2 - \frac{1}{\delta}} \text{var}_n(E X_{n,t}) \longrightarrow \text{d} S_{\delta}. \)

(iii) When \( \gamma = 1 \) and \( \eta = 1 \), (i) applies.

**Remarks.**

(a) Consider case \( \gamma = 1, \eta = 0 \): when \( b_{\eta} > -1/2 \) the aggregate variance goes to zero asymptotically, that is, idiosyncratic uncertainty is fully diversified when aggregating. By contrast, when \( b_{\eta} < -1/2 \), then \( n^{2 - \frac{1}{\delta}} = n^{\frac{2b_{\eta} + 1}{\delta}} \downarrow 0 \) as \( n \to \infty \) implying that for any \( 0 < c < \infty \) and \( 0 < d < 1/\delta - 2 \)

\[
P(n^{-d} \text{var}_n(E X_{n,t}) < c) \to 0 \text{ as } n \to \infty,
\]  

(12)

meaning that \( \text{var}_n(E X_{n,t}) \) diverges to infinity in probability at rate \( 1/\delta - 2 \). Therefore the ‘usual result’ fails.
(b) The result is non-trivial. In fact, for $-1/2 < b_{\pi} \leq 0$ $E(X_{n,t})$ is infinite even though $E X_{n,t}$ goes to zero in probability.

c) When $\pi_i = \pi = \alpha + \beta$ for any $i$, $E X_{n,t} \to 0$ in mean-square for any $\gamma \leq 1$.

d) When $\gamma > 1$, including the case of individual IGARCH(1,1), $\text{var}_n(E X_{n,t})$ is unbounded and Theorem 1 does not apply. A generalization of Theorem 1 exists and is available upon request from the author.

e) Theorem 1 easily extends to case $0 < \eta < 1$ using Zaffaroni (2004a)[Lemma 1].

As shown by Bollerslev (1986), bounded fourth moment for the individual GARCH(1,1) processes requires $\nu_i < 1$ a.s. In this case, the $E X_{n,t}$ converge to zero in mean-square for any value of $b_{\pi}$ by Theorem 1, implying the following.

**Corollary** The covariance stationarity condition for $E X_{n,t}^2$ is $\bar{\nu} \leq 1$ implying $E X_{n,t}^2 \to_p 0$, as $n \to \infty$.

To investigate the effect of relaxing condition $\bar{\nu} \leq 1$, we study the asymptotic distribution of the $E X_{n,t}$, using the suitable normalization suggested by Theorem 1. Given the possibility of asymptotic nonstationarity, we also look at the behaviour of $E \tilde{X}_{n,t} = n^{-1} \sum_{i=1}^n \tilde{x}_{i,t}$, obtained by setting $\epsilon_{i,s} = 0$ $(i = 1, ..., n)$ for all $s \leq 0$, yielding $\tilde{x}_{i,t} = \epsilon_{i,t} \tilde{\sigma}_{i,t}$, with

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

(13) is equivalent to the conditional model of Nelson (1990, eq.(6)) with the initial distribution of $\tilde{\sigma}_{i,0}$ equal to the Dirac mass at zero. Its conditional variance is $\tilde{V}_{t,n} = \text{var}_n(E \tilde{X}_{n,t}) = n^{-2} \sum_{i=1}^n \omega_i (1 - \pi_i^t)(1 - \pi_i)^{-1}$.

**Theorem 2** As $n \to \infty$:

(i) When $\gamma < 1$, any $\eta$, or $\gamma = 1$, $\eta = 0$, $b_{\pi} > 0$

$$
\sqrt{n} E X_{n,t} \longrightarrow_d S_2(t),
$$
where the $S_2(t)$ are uncorrelated and distributed like a normal r.v. $N(0,V)$ with $V = E(\omega_i/(1 - \pi_i))$.

(ii) When $\gamma = 1, \eta = 0, b_\pi < 0$

$$\sqrt{n} \ E \tilde{X}_{n,t} \longrightarrow_d \tilde{S}_2(t),$$

(14)

where the $\tilde{S}_2(t)$ are uncorrelated and distributed like a normal r.v. $N(0,V_t)$ with $V_t \sim c t^{-b_\pi}$ as $t \to \infty$.

Assuming further $E[\max_{k \geq 1} (\prod_{s=1}^{k} (c_p+ |\epsilon_{i,t-s}|^{2}) (1-p)] < \infty$ for any $t \in \mathbb{Z}, i \in \mathbb{N}$

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

(15)

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

(iii) When $\gamma = 1, \eta = 1$, (i) applies.

Remarks.

(a) When the micro processes are mutually independent, the (suitably normalized) aggregate converges to a $\delta$-stable process, Gaussian in the stationary case ($b_\pi > 0$). Hence, the ARCH structure characterizing the micro processes is lost through aggregation as the LA is not a volatility model. This is caused by (2), which imposes uncorrelatedness and independence of the $\epsilon_{i,t}$ (and of the $\pi_i$), which permits the application of the standard central limit theorem (henceforth CLT) for i.i.d. random variables. This result extends to weakly cross-sectionally correlated $\epsilon_{i,t}, \pi_i$, as long as the standard CLT applies. The limit process satisfies $E S_\delta(t) S_\delta(v) = 0$ for any $t \neq v$ and any $0 < \delta \leq 2$. When $\delta = 2$, this implies independence of $S_2(t)$ and $S_2(v)$ when $t \neq v$, whereas this is not guaranteed when $\delta < 2$. Note that $E(S_\delta(t))^2$ is unbounded for $\delta < 2$.
(b) The nonstationary case results \((b_\pi < 0)\) can be viewed as sequential limits of the (normalized) truncated aggregate \(E\tilde{X}_{n,t}/\sqrt{\tilde{V}_{t,n}}\), depending on the order at which \(t, n\) go to infinity. Only in the stationary case \((b_\pi > 0)\) is the limit distribution and the rate of convergence the same.

Let us discuss case \(b_\pi < 0\). Here both the rate of convergence and the asymptotic distribution depend on the order at which \(n\) and \(t\) go to infinity. Phillips and Moon (1999, Appendix B(1)) clarify the probability arguments necessary for sequential asymptotics and, in a general multi-index framework, establish conditions under which sequential and joint limit give equivalent results. Their conditions do not apply to our nonstationary case \(b_\pi < 0\); see also Taqqu, Willinger, and Sherman (1997) for another example where the equivalence between sequential and joint limits fails. When \(t \to \infty\), \(E\tilde{X}_{n,t} \to_d E X_{n,t}\) and \(\tilde{V}_{t,n} \to_{a.s.} \tilde{S}_2(t)/\sqrt{V_t}\) as \(n \to \infty\), yielding the left-hand side of (15), the non-truncated normalized aggregate, except for the random denominator, of order \(n^{1/\delta-1}\) (by Theorem 1). Then, as \(n \to \infty\), the limit distribution will be \(S_\delta(t)/\sqrt{S_2}\) (recall \(S_2 > 0\ a.s.\)). For the other type of sequential limit \(E\tilde{X}_{n,t}/\tilde{V}_{t,n} \to_{a.s.} \tilde{S}_2(t)/\sqrt{V_t}\) as \(n \to \infty\), yielding the right-hand side of (14) but with \(\sqrt{V_t}\) in the denominator. Writing \(\tilde{S}_2([rt])/\sqrt{V_t}\) for \(0 \leq r \leq 1\), one obtains as \(t \to \infty\) a sequence of r.v.s, normally distributed \(N(0, r^{-b_\pi})\) and mutually independent for any \(r \neq r'\).

(c) A close analogy exists between Theorem 2 and certain results of temporal aggregation of GARCH. When the \(x_t\) satisfy (2) and (3) for non-random \(\omega_i = \omega, \pi_i = \pi\) and setting \(\epsilon_{i,t} = \epsilon_t\) (let Ass. I holds with \(\mu_2 = 1\)), then \(\sum_{t=1}^T x_t/\sqrt{T} \to_d N(0, \omega/(1-\pi))\), as \(T \to \infty\), when \(\pi < 1\). This was first discovered by Diebold (1988) in the ARCH(1) case. This result could be extended to the case \(\pi \geq 1\). Looking at ARCH(1), for the sake of simplicity, let \(\delta\) satisfy the equation \(E(\alpha u_t^2)\delta = 1\) (see Davis and Mikosch (1998, Table
1), yielding $\delta \geq 2$ when $\alpha \leq 1$ and $0 < \delta < 2$ when $1 < \alpha < \exp(-E \log \epsilon^2_t)$. Then, for $\delta' = \min[\delta, 2]$, $\sum_{t=1}^{T} x_t/T^\frac{1}{\delta'} \rightarrow_d \delta'$-stable r.v., as $T \rightarrow \infty$.

2.3 Common innovations

In this section the aggregate is denoted by $U X_{n,t} = n^{-1} u_t \sum_{i=1}^{n} \sigma_{i,t}$. Due to the dependence between $\sigma_{i,t}$ and $\sigma_{j,t}$ which is induced by the $u_t$, $\text{var}_n(U X_{n,t}) = n^{-2} \sum_{i,j=1}^{n} E_n(\sigma_{i,t} \sigma_{j,t})$, whose behaviour is described as follows.

**Theorem 3** As $n \rightarrow \infty$:

(i) When $\gamma < 1$ and for any $\eta$, $\text{var}_n(U X_{n,t}) \longrightarrow_{a.s.} C$.

(ii) When $\gamma = 1$ and $\eta = 0$, setting $\delta = -(b_\bar{\nu} + 1)/b_\bar{\nu}$:

if $b_\bar{\nu} > -1/2$, $\text{var}_n(U X_{n,t}) \longrightarrow_{a.s.} C$;

if $b_\bar{\nu} \leq -1/2$, $n^{1-\frac{1}{\delta}} \text{var}_n(E X_{n,t}) \longrightarrow_d S_\delta$.

(iii) When $\gamma = 1$ and $\eta = 1$, (i) applies.

Remarks.

(a) The variance of the $U X_{n,t}$ is always bounded away from zero for every value of $b_\bar{\nu}$. However, when $b_\bar{\nu} < -1/2$, the variance explodes in probability, in the sense of (12), at exactly the same rate of $\text{var}_n(E X_{n,t})$, equal to $n^{-\frac{2b_\bar{\nu}+1}{b_\bar{\nu}+1}}$.

(b) Recall that for $\bar{\nu} \leq 1$ each individual GARCH(1,1) has a bounded fourth moment. By using arguments similar to Theorem 3, for $B(\nu_t; b_\bar{\nu}) \sim C(\bar{\nu} - \nu_t)^{b_\bar{\nu}}$, $\nu_t \rightarrow \bar{\nu}^-$, it easily follows that the limit of $U X_{n,t}$ has a bounded fourth moment when $\bar{\nu} \leq 1$ with $b_\bar{\nu} > -3/4$. By contrast, the LA exhibits unbounded kurtosis when $\bar{\nu} = 1$ with $b_\bar{\nu} < -3/4$. Thus, the distribution of the LA could exhibit fatter tails than those of the distribution of the individual GARCH(1,1) processes.

We now characterize the asymptotic distribution of the $U X_{n,t}$. 

12
**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
\[
\min[X_{n,t}, \overline{X}_{n,t}] \leq U_{X_{n,t}} \leq \max[X_{n,t}, \overline{X}_{n,t}], \text{ a.s.}
\]
satisfying the following.

(i) When \( \gamma < 1 \) or \( \gamma = 1 \), \( \min\{b_{\bar{\alpha}}, b_{\overline{\beta}}\} > 0 \):
\[X_{n,t} \rightarrow_1 X_t, \quad \overline{X}_{n,t} \rightarrow_1 \overline{X}_t \text{ as } n \rightarrow \infty.\]

\( \{X_t, \overline{X}_t\} \) and \( \{X_{n,t}, \overline{X}_{n,t}\} \) are defined in (32)-(33) in the appendix.

(ii) When \( \gamma < 1 \) or \( \gamma = 1 \), \( \min\{b_{\bar{\alpha}}, b_{\overline{\beta}}\} > -1/2 \), then \( X_t, \overline{X}_t \) are covariance stationary. Under the same conditions \( X_t \) and \( \overline{X}_t \) are strictly stationary and ergodic. When \( \mu_4^{\frac{1}{2}} \bar{\alpha} + \bar{\beta} < 1 \) or \( \mu_4^{\frac{1}{2}} \overline{\alpha} + \overline{\beta} = 1 \), \( \min\{b_{\bar{\alpha}}, b_{\overline{\beta}}\} > -1/2 \) then \( X^2_t, \overline{X}^2_t \) are covariance stationary.

**Remarks.**

(a) We have characterized the limit of the ‘envelope’ processes \( \underline{X}_{n,t} \) and \( \overline{X}_{n,t} \) rather than looking directly at \( U_{X_{n,t}} \). The ‘envelope’ seems tight enough as \( X_t \) and \( \overline{X}_t \) share the same covariance stationarity condition up to the fourth order in nearly all circumstances. The LA would certainly have a cumbersome expression, requiring stochastic expansion arguments (e.g. Hermite expansions for Gaussian \( u_t \)). This would make the LA difficult to be used in applications (e.g. for estimation).

(b) Based on these results, the exact limit of \( S_{n,t} = n^{-1} \sum_{i=1}^{n} x_{i,t}^2 \) can be easily established without using the ‘envelope’ processes. Note, however, that by simply looking at \( S_{n,t} \) would in part mask the effect of aggregation. In fact \( X^2_{n,t} = n^{-1} S_{n,t} + \left( n^{-2} \sum_{i \neq j \in \mathbb{N}} x_{i,t} x_{j,t} \right) \) and the second term on the right-hand side (in brackets) is asymptotically negligible only when \( z_{i,t} = \epsilon_{i,t} \).

(c) Unlike the finite \( n \) case of Nijman and Sentana (1996) and Meddahi and
Renault (2004), the LA does not belong to the class of weak GARCH although it displays dynamic conditional heteroskedasticity. Under Ass. II(γ) and III(η), the coefficients driving the limit aggregate depend on (see (33) in appendix)

$$E(\alpha_i^k) \sim c \bar{\alpha}^k k^{-(b_{\alpha}+1)}, E(\beta_i^k) \sim c \bar{\beta}^k k^{-(b_{\beta}+1)}, \text{ as } k \to \infty,$$

by Zaffaroni (2000, eq.(27)). (17) is not compatible with the coefficients obtained by expanding the ratio of finite-order polynomials in the lag operator (see Definition 1, 2 and 3 in Drost and Nijman (1993) for strong, semi-strong and weak GARCH). More generally (17) implies that the (multivariate) Markov structure of GARCH is lost by aggregation as \( n \to \infty \).

(d) The limit processes \( X_t \), \( X_i \) differ from all the GARCH-type long memory volatility models introduced in the relevant literature, in particular from the ARCH(∞) of Robinson (1991).

### 2.4 A numerical example

Figure 1 reports simulated examples of \( X_{n,t} \) with \( n = 1,000 \) cross-section observations and \( T = 500 \) time observations. The dotted line refers to the aggregate of \( x_{i,t} \) with innovations \( u_t \) and the bold line to the aggregate of \( x_{i,t} \) with innovations \( \epsilon_{i,t} \), both obtained by (pseudo) drawing standard normal. Thus Ass. I is satisfied for \( \mu_2 = 1, \mu_4 = 3 \). As far as Ass. II(γ) is concerned, we consider case \( \gamma = 1 \). The \( \omega_i \) are drawn from a uniform distribution over \([c, 1]\), with \( c = 1e - 14 \) throughout this section. The \( \alpha_i \) are drawn from a Beta\((p_{\alpha}, q)\) distribution over the interval \([0, 0.3 - c]\) with parameters \( q \in \{0.1, 0.5, 2\} \), \( p_{\alpha} = q \mu_{\alpha}/(0.3 - \delta - \mu_{\alpha}) \) where \( \mu_{\alpha} = E(\alpha_i) = 0.27 \). The \( \beta_i \) are drawn from a Beta\((p_{\beta}, q)\) distribution over the interval \([0, 0.7 - c]\) with \( \mu_{\beta} = 0.6, b_{\alpha} = b_{\beta} \). This implies \( b_{\epsilon} = 2q - 1 \). We assume \( \eta = 0 \) for Ass. III(η). The distribution of the aggregate with common innovations
(dotted line) is always non-degenerate and exhibiting conditional dynamic heteroskedasticity for any \( q \) (cf. Theorem 4). Instead, the distribution of the aggregate with idiosyncratic innovations (bold line) appears degenerate except when \( q = 0.1 \), implying \( b\bar{\pi} = -4/5 \) and thus below \(-1/2\) (cf. Theorem 2).

Table 1 reports the realizations of \( \text{var}_n(E X_{n,t}) \) and of the ratio \( \text{var}_n(E X_{n,t}) / \text{var}_n(U X_{n,t}) \) for \( n \in \{10, 10, 1000\} \). We now consider two cases for the \( \omega_i \). The first panel refers to \( \omega_i \) drawn from an inverted Gamma with parameters \((2, 1)\) ensuring \( E\omega_i^2 < \infty \). The second panel refers to \( \omega_i \) drawn from a uniform distribution over \([c, 1]\). The \( \pi_i \) are Beta\((p\pi, q\pi)\) over the interval \([0, 1 - c]\) with \( q\pi \in \{0.1, 0.2, 0.3, 0.7, 1, 3\} \), \( p\pi = q\pi \mu\pi / (1 - \delta - \mu\pi) \) where \( \mu\pi = E(\pi_i) = 0.97 \).

When \( q\pi = 1 \) the distribution of the \( \pi \) behaves locally (around unity) as the uniform distribution. We set \( \eta = 0 \) in Ass. III(\( \eta \)). The form of the distribution of the \( \omega_i \) has no effect as no relevant differences appear between the two panels. For values of \( q\pi < 1/2 \), implying \( b\pi < -1/2 \), one can see how \( \text{var}_n(E X_{n,t}) \) is, on average, very large, tending to diverge as \( n \) increases (cf. Theorem 1). The only exception is for \( q\pi = 0.1 \), the closest case to nonstationarity, for which the variance is already sizeable even for \( n = 10 \). In contrast, it is small, converging to zero as \( n \) increases, for \( q\pi > 1/2 \). The last six columns of Table 1 refer to the ratio of \( \text{var}_n(E X_{n,t}) / \text{var}_n(U X_{n,t}) \). It is stable for \( q\pi < 1/2 \), since they both diverge at the same rate \( n^{-2q\pi - 1 / q\pi} \), but converging to zero otherwise (cf. Theorems 1 and 3) as \( \text{var}_n(U X_{n,t}) \) is always bounded away from zero. The numerical difficulties associated with the quasi-nonstationary case \( q\pi = 0.1 \) lead to a small ratio. All the computations have been carried out in MatLab.
3 Generalizations and implications

This framework can be generalized in many directions. In the first place, we analyze in detail the memory implications for the volatility of the LA. Next, we consider various, more realistic, forms of cross-sectional dependence between the $x_{i,t}$ to which our results apply. Finally, we discuss how our results can be used for aggregation schemes other than the simple equally weighted scheme one.

3.1 Memory of aggregate volatility

In this section we focus exclusively on case $z_{i,t} = u_t$ as Section 2.2 shows that dynamic heteroskedasticity is cancelled at the aggregate level for the case of idiosyncratic innovations. Ding and Granger (1996) suggested that a long memory volatility model could be obtained by aggregating heterogeneous GARCH(1,1). Their aggregate is defined by $X_{n,t}^{DG} = u_t \left( \sum_{i=1}^{n} w_i \tau_{i,t}^2 \right)^{1/2}$ with $\tau_{i,t}^2 = \sigma^2 (1 - \alpha_i - \beta_i) + \alpha_i \left( X_{n,t-1}^{DG} \right)^2 + \beta_i \tau_{i,t-1}^2$, where the deterministic weights $w_i$ satisfy $\sum_{i=1}^{n} w_i = 1$ and $\sigma^2$ is a constant parameter. Note that $X_{n,t}^{DG}$ differ from $X_{n,t}$ in (1) and that $\tau_{i,t}^2$ differs from GARCH(1,1). Their structure allows them to apply Granger (1980) linear aggregation results, suggesting that, as $n \to \infty$, the $\sum_{i=1}^{n} w_i \tau_{i,t}^2$ converge (in some norm) to a special case of Robinson (1991) ARCH($\infty$), with hyperbolically decaying coefficients and a bounded fourth moment (stationary squares). Using the results of our paper, we re-consider Ding and Granger (1996)’s set-up and investigate the implications of aggregation for the memory of aggregate volatility. We look at

$$\Sigma_{n,t} = \frac{1}{n} \sum_{i=1}^{n} \sigma_{i,t}^2,$$

with strong GARCH(1,1) $\sigma_{i,t}^2$ (cf. (3)) and common innovations $u_t$. 

16
Theorem 5  (i) When $\bar{\pi} < 1$ or $\bar{\pi} = 1$, $b_{\bar{\pi}} > 0$, as $n \to \infty$,

$$E_n \Sigma_{n,t} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{1 - \pi_i} \to \text{a.s.} \ C < \infty,$$

and explodes in probability when $\bar{\pi} = 1$, $b_{\bar{\pi}} < 0$.

(ii) For any $u = 0, \pm 1, ...$

$$\text{cov}_n(\Sigma_{n,t}, \Sigma_{n,t+u}) > 0 \ \text{a.s.}$$

Let $\text{cov}_n(\Sigma_{n,t}, \Sigma_{n,t+u}) \to \text{a.s.} \ a(u), \ u = 0, \pm 1, ..., \ for \ n \to \infty$, when the limit exists. $a(u)$ is not necessarily finite.

When $\bar{\nu} < 1$

$$a(0) < \infty, \ a(u) = O(\bar{\nu}^{\frac{u}{2}}) \ as \ u \to \infty.$$

When $\bar{\nu} = 1$ and $0 < \alpha \leq \alpha_i$ a.s., for some $0 < \alpha < 1$, then for $d_{\rho} > -1/2$

$$a(0) < \infty, \ a(u) = O \left( \left( 1 - \alpha \right)^{\frac{u}{2}} \right) \ as \ u \to \infty.$$

and $a(0)$ is unbounded when $d_{\rho} < -1/2$.

Assume $\bar{\nu} = 1$ and $\alpha_i = \gamma_i(\bar{\beta} - \beta_i)$, where $0 \leq \gamma_i \leq \bar{\gamma} < 1$ where the $\gamma_i$ can be a function of the $\beta_i$ (this nests Ding and Granger (1996) assumption), for some $\bar{\gamma}$. Then, for $b_{\bar{\pi}} < 0$, $a(0)$ is unbounded whereas for $b_{\bar{\pi}} > 0$

$$a(0) < \infty, \ a(u) \sim c u^{-(2b_{\bar{\pi}} + 1)} \ as \ u \to \infty.$$

Remarks.

(a) Long memory is ruled out in all cases ($\sum_{u=0}^{\infty} a(u) < \infty$), even under the Ding and Granger (1996) assumptions. In this latter case, however, the autocorrelation function (acf) of the squared LA decays like a power law (not exponentially). This result agrees with Kazakeviųs, Leipus, and Viano (2004) who showed that any fourth-order stationary ARCH(\infty) cannot
exhibit long memory.

(b) The key feature, ruling out long memory, is the relationship between the conditions for the bounded second and fourth moments of GARCH. The former is (for GARCH(1, 1))

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

(19)

and the latter

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

(20)

Since $\mu_2 < \mu_4$ ($u_t^2$ are not degenerate), (20) is strictly stronger than (19) as long as $\alpha_i \geq \alpha > 0 \text{ a.s.}$ Therefore, $\bar{\pi} < 1$, yielding an exponentially decaying acf. Conditions (19) and (20) are nearly equivalent when $\alpha_i \downarrow 0^+$. Since this happens when $\beta_i \uparrow 1^-$ and $\pi_i \uparrow 1^-$, $\nu_i \uparrow 1^-$, and the aggregation results depend precisely on the behaviour near 1 of the distribution of $\pi_i$ and $\nu_i$, an hyperbolically decaying acf is obtained. The rate of decay of the acf cannot be too slow however, owing to the stationarity condition of the levels ($b_\pi > 0$). This type of result is not new for ARCH models. For instance, the covariance stationarity condition for levels rules out long memory squares for ARCH($\infty$) (see Zaffaroni (2004b, Theorem 3)).

(c) For the parallel case of negative dependence, $\beta_i = \gamma_i (\bar{\alpha} - \alpha_i)$, (20) will still be strictly stronger than (19), given $\mu_2^2 < \mu_4$, implying an exponentially decaying acf. This is not a surprising outcome, as in this case the $x_{i,t}$ behave (locally) like ARCH(1) or, more generally, like ARCH($q$), $q \geq 1$, to which Theorem 5 does not apply.

3.2 Other forms of cross-sectional dependence

Our results can be readily used to analyze the effect of aggregation for other forms of cross-sectional dependence between the units $x_{i,t}$. A convenient way
is through a (conditionally heteroskedastic) factor structure. For the sake of simplicity, we consider one-factor structures only. The simplest factor structure case is

\[ x_{i,t} = b_i u_t \sigma^u_{i,t} + \epsilon_{i,t} \sigma^\epsilon_{i,t}, \]  

(21)

where \( b_i \) are the random factor loadings, assumed \( i.i.d. \) across units with \( Eb_i \neq 0 \). Note that \( \sigma^u_t \) is homogeneous across units, where hereafter \( \sigma^z_{i,t} \) denotes the GARCH(1,1) conditional variance with generic rescaled innovation \( z_{i,t} \) (cf. (2)-(3)). When the idiosyncratic component \( \epsilon_{i,t} \sigma^\epsilon_{i,t} \) vanishes in mean square (cf. Theorem 1), then \( X_{n,t} \) would assume an exact GARCH structure as \( n \) approaches infinity. A factor structure more general than (21) is

\[ x_{i,t} = b_i u_t \sigma^u_{i,t} + \epsilon_{i,t} \sigma^\epsilon_{i,t}, \]  

(22)

where now the coefficients of the factor conditional variance are heterogeneous across units. Model (22) is equivalent to assume \( x_{i,t} = \phi_i u_t \sigma^\phi_{i,t} + \epsilon_{i,t} \sigma^\epsilon_{i,t} \). The common component, viz. the part that involves the \( u_t \), is simply (2)-(3) but with rescaled innovation \( \phi_i u_t \), with \( E \phi_i \neq 0 \). Model (22) follows setting \( b_i = \phi_i \) and replacing \( \alpha_i \) by \( \alpha_i^* = \alpha_i \phi_i^2 \) in (3). Theorems 1 and 2 apply to the idiosyncratic component and Theorems 3 and 4 to the common component providing a complete characterization of the limit of \( X_{n,t} \). For instance, the limit would not be a GARCH, even when \( \epsilon_{i,t} \sigma^\epsilon_{i,t} \) vanishes. Despite their simplicity, models (21) and especially (22) allow the description of an extremely general form of cross-sectional dependence between the \( x_{i,t} \); see King, Sentana, and Wadhwani (1994).

### 3.3 Other aggregation schemes

When \( x_{i,t} \) are returns of assets with random pay-off, the aggregate (1) defines the return of the portfolio made by \( 1/n \)th of each asset. Our results extend
immediately, under suitable regularity conditions, to the case of weighted average portfolios \( \sum_{i=1}^{n} w_i(n) x_{i,t} \) with stochastic weights \( w_i(n) \) behaving as \( 1/n \) almost surely asymptotically. Our focusing on (1) is not merely due to its mathematical simplicity. The constant weights \( 1/n \) of (1) correspond exactly to the weights of the globally minimum-variance efficient portfolio based on the so-called (unconditionally) constant variance-correlation model (see Elton and Gruber (1995, p.195-198)). Many of the stock indexes most commonly used as benchmark portfolios are equally weighted such as the Dow Jones Industrial Average, the FT30 Index, the Major Market Index, the Nikkei 225 Index, the Standard & Poor’s 500 Equal Weight Index and Value Line Index.

We now show that our results are also approximately valid for value weighted indexes. Let us assume that we observe a sample \( \{P_{1t}, \ldots, P_{nt}\} \) of \( n \) positive r.v.s. with \( P_{it} = e^{\mu_t + \sigma_{i,t}z_{i,t}} \) with non degenerate r.v. \( |z_{i,t}| < \infty \) a.s. Unless \( \sigma_t = 0 \) (see Hardy, Littlewood, and Polya (1964, Theorem 9))

\[
\left( \prod_{i=1}^{n} P_{it} \right)^{\frac{1}{n}} < \frac{1}{n} \sum_{i=1}^{n} P_{it} \quad \text{a.s.} \tag{23}
\]

By a second-order Taylor expansion of the geometric mean of the \( P_{it} \) around \( \sigma_t = 0 \) one gets \( \left( \prod_{i=1}^{n} P_{it} \right)^{\frac{1}{n}} \approx e^{\mu_t} \left[ 1 + \sigma_t \hat{m}_{it1} + \frac{\sigma_t^2}{2} \hat{m}_{it2} + \frac{\sigma_t^3}{3!} \sum_{i=1}^{n} e^{\sigma_t \hat{m}_{i(t)}(\hat{m}_{it})^3} \right] \quad \text{a.s.} \)

for some \( 0 < \sigma_t < \sigma_{i,t} \), setting \( \hat{m}_{itj} = \frac{1}{n} \sum_{i=1}^{n} z_{itj} \), \( j \geq 0 \). Likewise \( n^{-1} \sum_{i=1}^{n} P_{it} = e^{\mu_t} \left[ 1 + \sigma_t \hat{m}_{it1} + \frac{\sigma_t^2}{2} \hat{m}_{it2} + \frac{\sigma_t^3}{3!} \sum_{i=1}^{n} e^{\sigma_t \hat{m}_{i(t)}(\hat{m}_{it})^3} \right] \quad \text{a.s.} \), some \( 0 < \sigma_t'' < \sigma_t \). No moment conditions on the \( z_{it} \) are required since \( n < \infty \) and the expansion is of a finite order. By relatively simple manipulations (details are available upon request to the author) \( \left( \prod_{i=1}^{n} \frac{P_{it}}{P_{it-1}} \right)^{\frac{1}{n}} = \left( \sum_{i=1}^{n} \frac{P_{it}}{P_{it-1}} \right)^{\frac{1}{n}} B_{tn} \), a.s. with

\[
B_{tn} \approx \frac{1 - \left( \frac{\sigma_t^2}{2} \hat{m}_{it2} + \frac{\sigma_t^3}{3!} \sum_{i=1}^{n} e^{\sigma_t \hat{m}_{i(t)}(\hat{m}_{it})^3} \right) \left( \frac{1}{n} \sum_{i=1}^{n} e^{\sigma_t z_{it}} \right)^{-1}}{1 - \left( \frac{\sigma_t^2}{2} \hat{m}_{it-12} + \frac{\sigma_t^3}{3!} \sum_{i=1}^{n} e^{\sigma_t \hat{m}_{i(t-1)}(\hat{m}_{it-1})^3} \right) \left( \frac{1}{n} \sum_{i=1}^{n} e^{\sigma_{i-1} z_{i-1}} \right)^{-1}} \tag{24}
\]
for a sufficiently large $n$, when the $z_{it}$ are symmetric distribution around $Ez_{it} = 0$. Under the additional assumption of Gaussian $z_{it}$

$$B_{tn} \to_p B_t = e^{\frac{\sigma^2_{t-1} - \sigma^2_t}{2}}. \tag{25}$$

Both (24) and (25) indicate that, despite (23), there is no systematic bias in terms of rate of return. We can now apply this approximation to value weighted portfolios. If $S_{i,t}$ defines the number of outstanding shares of asset $i$ at time $t$, with price $P_{i,t}$, the rate of return of a value weighted index is $\log(\sum_{i=1}^{n} S_{i,t}P_{i,t}/\sum_{i=1}^{n} S_{i,t-1}P_{i,t-1})$. This is approximately equal to $X_{n,t} + n^{-1}\sum_{i=1}^{n} \log(S_{i,t}/S_{i,t-1})$, setting $x_{i,t} = \log(P_{i,t}/P_{i,t-1})$. The second component reflects circumstances such as stock issues and repurchases, mergers and bankruptcies which do not require describing time-varying conditional heteroskedasticity, unlike for $X_{n,t}$.

### 4 Concluding remarks

This paper analyzes the statistical properties of the aggregate $X_{n,t}$ of GARCH(1,1) processes, as $n$ approaches infinity. This leads on to many possibilities for further research. One can develop estimation procedures for this random coefficient GARCH model (see Robinson (1978) for estimation of a random coefficient autoregressive model based on similar assumptions). This would permit the testing of several of the implications of this paper, such as the precise relationship between the memory of the volatility of the LA and the cross-sectional distribution of the individual GARCH parameters. It would also represent the necessary step for developing an estimation procedure for factor model (22).
Appendix

Let us 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 a preliminary lemma (see Zaffaroni (2000, Lemma 2) for its proof), then present the proof of the theorems.

Lemma 1 Let $\{z_i\}$ be a sequence of i.i.d. positive r.vs with probability density $B(\cdot; b)$, defined in the interval $[0, 1)$, such that for real $b \in (-1, \infty)$

$$B(z; b) \sim c(1-z)^b \quad \text{as } z \to 1^-.$$  

For any integer $p = 1, 2, \ldots$ and real $k$, as $n \to \infty$:

(i) $\frac{1}{n^p} \sum_{i_1, \ldots, i_p=1}^n \frac{1}{(1-z_{i_1} \cdots z_{i_p})^k} \sim c + C \frac{1}{n} \sum_{i=1}^n (1-z_i)^{(p-1)b+(p-1-k)}$ a.s.

The boundedness condition is $pb + (p-k) > 0$.

(ii) When $pb + (p-k) > 0$, integer $u > 0$, $r (0 \leq r \leq p)$ with $s = p - r$

$$\frac{1}{n^p} \sum_{i_1, \ldots, i_p=1}^n \frac{z_{i_1}^u \cdots z_{i_p}^u}{(1-z_{i_1} \cdots z_{i_p})^k} \longrightarrow a.s. \quad g_{(r,s)}^{(k)}(u) \quad \text{as } n \to \infty,$$

where $g_{(r,s)}^{(k)}(u) \sim c \left( E(z_u^r) \right)^r (1 + u^{-(sb+s-k)})$ as $u \to \infty$.

Proof of Theorem 1. When $\eta = 0$ apply Zaffaroni (2004a)[Lemma 1] with $k = 1$. When $\eta = 1$ $\text{var}_n(E X_{n,t}) = n^{-2} \sum_{i=1}^n \tilde{\omega}_i = O(n^{-1})$ 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 $n^{-1} \sum_{i=1}^n \omega_i/(1 - \pi_i)$ converges a.s. to $E(\omega_i/(1 - \pi_i))$, bounded when $b_\bar{\pi} > 0$. Moreover, for any integers $n, u > 0$ easy calculations yield $\text{cov}_n(n^{-\frac{1}{2}} \sum_{i=1}^n x_{i,t}, n^{-\frac{1}{2}} \sum_{i=1}^n x_{i,t+u}) = 0$ a.s. where $\text{cov}_n(., .)$ denotes the covariance operator, conditional on the $\omega_i, \pi_i (i = 1, \ldots, n)$. 22
(ii) For (14) we follow (i), where by Stirling’s formula (Brockwell and Davis 1987, p.522) \( E(x_i^2) = \nu_i = E(\omega_i \frac{i}{\pi_i}) = E(\omega_i \sum_{k=0}^{t-1} E_{\pi_i}^k \sim c t^{-b_i}, \quad \text{as } t \to \infty \) when \( b_\pi < 0 \). For (15), setting \( c_k^{(p)}(t) = \prod_{s=1}^{k} (c_p + |\epsilon_{i,t-s}|^{2/i-p}) \) and \( d_k^{(p)}(t) = \max_{k=q,q+1,...} c_k^{(p)}(t), 0 < p < 1 \),

\[
P(\sigma_{i,t}^2 > u) = P(\sum_{k=0}^{\infty} \prod_{j=1}^{k} (\beta_i + \alpha_i \epsilon_{i,t-j}^2 > u) \leq E(P_e(\frac{(d_k^{(p)}(t))^{1-p}}{(1 - \pi_i)} > u)), \quad (27)
\]

using the inequality \( a^e_1 b_1^{-e} + a^e_2 b_2^{-e} < (a_1 + a_2)^e (b_1 + b_2)^{1-e} \) for some \( 0 < e < 1 \) and positive \( a_1, a_2, b_1, b_2 \) (see Hardy, Littlewood, and Polya (1964, Theorem 38)), with \( \beta_i = a^e_1 b_1^{-e}, \alpha_i \epsilon_{i,t}^2 = a^e_2 b_2^{-e}, e = p \) and using (6), where \( P_e(.) \) denotes the probability operator, conditioning on 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 \( c_k^{(p)}(t) = O(e^{ -\frac{k}{2}}) \) a.s., for all \( k > K \), implying that \( c_k^{(p)}(t) \to \text{a.s.} \) 0 for \( k \to \infty \). Therefore, for some \( m \), such that \( m \to \infty \),

\[
P(\sigma_{i,t}^2 > u) = E(P_e(\sigma_{i,t}^2 > u)) \geq E(P_e(\frac{(d_m^{(p)}(t))^{1-p}}{(1 - \pi_i)} > u)). \quad (28)
\]

Set \( \delta' = (b_\pi + 1) \). Under Ass. II(\( \gamma \)), the \( y_i = (1 - \pi_i)^{-1} \) are in the domain of attraction of a \( \delta' \)-stable distribution, totally skewed to the right. In fact, denoting by \( f_y(.) \) the probability density function of the \( y_i \), this equals \( f_y(u) = B(1 - u^{-1}; b_\pi)u^{-2}, \quad 1 \leq u < \infty \), and satisfies \( f_y(u) \sim c u^{-(b_\pi+2)} \) as \( u \to \infty \). Therefore, as \( u \to \infty \)

\[
P(y_i > u) \sim C u^{-(b_\pi+1)}, \quad P(y_i <-u) = 0, \quad (29)
\]

and \( \int_0^u t^2 f_y(t) dt \sim cu^{2-(b_\pi+1)} \). Therefore Feller (1966, Theorem IX.8.1) applies, yielding, for \( u \to \infty \) \( n^{-1/\delta'} \sum_{i=1}^{n} (1 - \pi_i)^{-1} \to_d S_{\delta'} \), where \( S_{\delta'} > 0 \) a.s. with scale parameter \( \sigma = (C/D_{\delta'})^{\frac{1}{\delta'}} \) with \( C \) as in (29) and \( D_a = (1 - a)(\Gamma(2 - a) \cos(\pi a/2))^{-1}, a \neq 1 \) or \( D_1 = 2/\pi \) (see Samorodnitsky and

23
nitsky and Taqqu (1994, Property 1.2.15 and eq.(1.2.9))). Next, from (27), \( P(\sigma_{i,t}^2 > u) \leq E(P_x([d_0^{(p)}(t)]^{-p} / (1 - \pi_i) > u)) \sim c E[d_0^{(p)}(t)]^{(1-p)\delta'} u^{-\delta'} \) as \( u \to \infty \), by the dominated convergence theorem (\( P_\cdot \leq 1 \) and \( E \cdot \leq 1 < \infty \)) and \( E[d_0^{(p)}(t)]^{(1-p)\delta'} \leq [E[d_0^{(p)}(t)]]^{-p} \delta' < \infty \), as \( \delta' < 1 \), by Jensen’s inequality. For (28), setting \( m = m(u) \) with \( 1/m + 1/u \to 0 \) as \( u \to \infty \), \( P(\sigma_{i,t}^2 > u) \geq E(P_x([d_0^{(p)}(t)]^{-p} / (1 - \pi_i) > u)) \sim c E[d_0^{(p)}(t)]^{(1-p)\delta'} u^{-\delta'} \), for some positive function \( g(u) \downarrow 0 \) as \( u \to \infty \). Therefore \( P(\sigma_{i,t}^2 > u) \sim c u^{-\delta'} \), as \( u \to \infty \), since \( g(u) \) is arbitrary, with \( c \) depending on both the distribution of \( \epsilon_{i,t} \) and of the \( \pi_i \). Hence, setting \( \delta = 2(b_n + 1) \), \( \sigma_{i,t} \) is in the domain of attraction of a \( \delta \)-stable distribution, totally skewed to the right. Collecting terms \( n^{-\frac{1}{2}} \sum_{i=1}^{n} x_{i,t} \to_d S_\delta(t) \) as \( n \to \infty \), where the \( S_\delta(t) \) have a \( \delta \)-stable marginal distribution with zero location parameter, skew parameter \( \sigma = (P - Q)/(P + Q) \), setting \( P = \frac{h E[\epsilon_{i,t}^2 1_{\epsilon_{i,t} > 0}]}{h E[\epsilon_{i,t}^2 1_{\epsilon_{i,t} > 0}] + (1-h) E[(\epsilon_{i,t})^2 1_{\epsilon_{i,t} < 0}]} \), with \( h = P(\epsilon_{i,t} > 0), Q = 1 - P \) (see Feller (1966, eq.(8.4)) and Samorodnitsky and Taqqu (1994, p.6)), and scale parameter \( \theta' \), setting \( P(x_{i,t} > u) \sim D_\delta \frac{1+\sigma}{2} \theta' u^{-\delta}, P(x_{i,t} < -u) \sim D_\delta \frac{1-\sigma}{2} \theta' u^{-\delta} \) as \( u \to \infty \). When \( \eta = 1 \), (i) applies. \( \square \)

**Proof of Theorem 3.** Case \( \gamma < 1 \) is straightforward so let us focus on case \( \gamma = 1 \). Set \( \eta = 0 \). Then \( \text{var}_n(U_n, t) = \sum_{i,j=1}^{n} E_n(\sigma_{i,t}^2) \leq n^{-2} \sum_{i,j=1}^{n} E_n(\sigma_{i,t}^2) = n^{-2} \sum_{i=1}^{n} \omega_i (1 - \pi_i)^{-1} a.s. \). Moreover, by Schwarz’s inequality, \( \text{var}_n(U_n, t) \leq C \left( n^{-1} \sum_{i=1}^{n} \frac{\omega_i^2}{(1 - \pi_i)^{1/2}} \right)^2 a.s. \) and case \( b_\pi \leq -1/2 \) follows by Zaffaroni (2004a, Lemma 1). \( \square \)

**Proof of Theorem 4.** By a version of Minkowski’s inequality (Hardy, Littlewood, and Polya 1964, Theorems 24 and 25) for the left hand side term and by Jensen’s inequality for the right hand side term, for any sequence
positive \{a_{i,j}, i = 1, \ldots, j = 1, \ldots, n\} one obtains:

\[
\left(\sum_{i=0}^{\infty} \left(\frac{1}{n} \sum_{j=1}^{n} (a_{i,j})^2\right)^{\frac{1}{2}}\right)^{\frac{1}{2}} \leq \left(\frac{1}{n} \sum_{j=1}^{n} \left(\sum_{i=0}^{\infty} a_{i,j}\right)^{\frac{1}{2}}\right)^{\frac{1}{2}} \leq \left(\frac{1}{n} \sum_{j=1}^{n} \left(\sum_{i=0}^{\infty} a_{i,j}\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}, \quad (30)
\]

yielding \(|Y_{n,t}| \leq |U_{X_{n,t}}| \leq |Y_{n,t}|\) and \(\min\{Y_{n,t}, Y_{n,t}\} \leq U_{X_{n,t}} \leq \max\{Y_{n,t}, Y_{n,t}\}\) where \(Y_{n,t} = u_t \left(\sum_{k=0}^{\infty} \left(\frac{1}{n} \sum_{i=1}^{n} \omega_i \prod_{j=1}^{k} (\beta_i + \alpha_i u_{t-j}^2)^{\frac{1}{2}}\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}\),

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

Since

\[
\prod_{j=1}^{l} (\beta_i + \alpha_i u_{t-j}^2) = 1_{l=0} + 1_{l>0} \sum_{k=0}^{l} \alpha_i^k \beta_i^{l-k} \left(\sum_{(k)} u_{t-j}^2 \ldots u_{t-j_1}^2 \ldots - j_{k-1}\right), \quad (31)
\]

with \(\sum_{(k)} = 1_{k=0} + 1_{k>0} \sum_{j=1}^{l-k} 1_{j=1} \sum_{j_2=1}^{l-k-2} \ldots \sum_{j_{k-1}=1}^{l-k-1}\), and using (30) once again for the left hand site inequality only, yields (16) with

\[
\tilde{X}_{n,t} = u_t \left(\sum_{l=0}^{\infty} \sum_{l=0}^{l} \sum_{k=0}^{l} u_{t-j_1}^2 \ldots u_{t-j_{k-1}}^2 \left(\frac{1}{n} \sum_{i=1}^{n} \omega_i^2 \alpha_i^k \beta_i^{l-k}\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}, \quad (32)
\]

\[
\bar{X}_{n,t} = u_t \left(\sum_{l=0}^{\infty} \sum_{l=0}^{l} \sum_{k=0}^{l} u_{t-j_1}^2 \ldots u_{t-j_{k-1}}^2 \left(\frac{1}{n} \sum_{i=1}^{n} \omega_i \alpha_i^k \beta_i^{l-k}\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}, \quad (33)
\]

(i) Set

\[
\tilde{X}_t = u_t \left(\sum_{l=0}^{\infty} \sum_{l=0}^{l} \sum_{k=0}^{l} u_{t-j_1}^2 \ldots u_{t-j_{k-1}}^2 \left(E(\omega_i^2 \alpha_i^k \beta_i^{l-k})\right)^{\frac{1}{2}}\right)^{\frac{1}{2}},
\]

\[
\bar{X}_t = u_t \left(\sum_{l=0}^{\infty} \sum_{l=0}^{l} \sum_{k=0}^{l} u_{t-j_1}^2 \ldots u_{t-j_{k-1}}^2 \left(E(\omega_i \alpha_i^k \beta_i^{l-k})\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}.
\]

We consider \(\eta = 0\). The same results apply to \(0 < \eta \leq 1\) with tedious calculations. Using a version of the law of iterated logarithms (see Stout (1974, Corollary 5.2.1)) for the i.i.d. sequence \(\{\omega_i \alpha_i^k \beta_i^{l-k}\}\) yields, as \(n \to \infty\),

\[
\left|\frac{1}{n} \sum_{i=1}^{n} \omega_i \alpha_i^k \beta_i^{l-k} - E(\omega_i \alpha_i^k \beta_i^{l-k})\right| = O \left(\frac{\log \log n}{n}\right)^{\frac{1}{2}} \alpha_i \beta_i^{l-k} k^{-\beta_{\eta+1}/2} (l-k)^{-(\beta_{\eta+1})/2} \quad a.s.
\]

25
For $\mathbf{X}_{n,t}$

$$E_n |\mathbf{X}_{n,t} - \mathbf{X}_t| \leq \frac{E |u_t|}{2\omega^2} \sum_{l=0}^{\infty} \sum_{k=0}^{l} \binom{l}{k} \mu_2^k \left| \frac{1}{n} \sum_{i=1}^{n} \omega_i \alpha_i^k \beta_i^{l-k} - E(\omega_i \alpha_i^k \beta_i^{l-k}) \right|$$

$$= O\left(\frac{(\log \log n)}{n}\right)^{1/2} \sum_{l=0}^{\infty} (l + 1)^{-(\min\{b_\alpha, b_\beta\} + 1)} (\mu_2 \bar{\alpha} + \bar{\beta})^l \right) \text{ a.s.,}
$$

which is $O\left(\frac{(\log \log n)}{n}\right)^{1/2}$ a.s. The last bound is obtained as follows. For some $0 < \delta < 1/2$, $\sum_{k=0}^{l} \binom{l}{k} \mu_2^k E(\omega_i^2 \alpha_i^{2k} \beta_i^{2(l-k)}) = \sum_{k=[\delta]}^{l-1} \binom{l}{k} \mu_2^k E(\omega_i^2 \alpha_i^{2k} \beta_i^{2(l-k)}) + \sum_{k=[\delta]}^{l} \binom{l}{k} \mu_2^k E(\omega_i^2 \alpha_i^{2k} \beta_i^{2(l-k)})$. By Stirling’s formula, the first and third terms satisfy $\sum_{k=0}^{[\delta]-1} \binom{l}{k} \mu_2^k E(\omega_i^2 \alpha_i^{2k} \beta_i^{2(l-k)}) = O\left(\binom{l}{[\delta]} \left(\frac{1}{\log \log n}\right)^{1/2} \sum_{k=0}^{[\delta]} \mu_2^k E(\omega_i^2 \alpha_i^{2k} \beta_i^{2(l-k)}) \right) \right) \right)$

and $\sum_{k=[\delta]}^{l} \binom{l}{k} \mu_2^k E(\omega_i^2 \alpha_i^{2k} \beta_i^{2(l-k)}) = O\left(\binom{l}{[\delta]} \binom{1}{(1-\delta)l} \bar{\alpha} \binom{1}{(1-\delta)l} \right) \right)$

whereas the second satisfies $| \sum_{k=[\delta]}^{l} \binom{l}{k} \mu_2^k E(\omega_i^2 \alpha_i^{2k} \beta_i^{2(l-k)}) | = O\left(\binom{l}{[\delta]} \binom{1}{(1-\delta)l} \right) \right)$

Therefore the first and third term are of smaller order than the second since $\delta(1-\delta)^{1-\delta}$ can be made arbitrarily close to one as $\delta \to 0^+$. Along the same lines $E_n |\mathbf{X}_{n,t} - \mathbf{X}_t| = O\left(\frac{(\log \log n)}{n}\right)^{1/2} \sum_{l=0}^{\infty} (l + 1)^{-(\min\{b_\alpha, b_\beta\} + 1)} (\mu_2 \bar{\alpha} + \bar{\beta})^l \right) \text{ a.s.}
$$

(ii) Covariance stationarity of levels follows by $\mathbf{X}_t^2 \leq \mathbf{X}_t$ and

$$E\mathbf{X}_t^2 = O\left(\sum_{l=0}^{\infty} (l + 1)^{-(\min\{b_\alpha, b_\beta\} + 1)} (\mu_2 \bar{\alpha} + \bar{\beta})^l \right) \text{ a.s.}
$$

Strict stationarity and ergodicity easily follow adapting the proof of Nelson (1990, Theorem 2 and p.329). Covariance stationarity of the squares follows using Schwartz inequality and $E\mathbf{X}_t^2 = O\left(\sum_{l=0}^{\infty} (l + 1)^{-(\min\{b_\alpha, b_\beta\} + 1)} (\mu_2 \bar{\alpha} + \bar{\beta})^l \right) \text{ a.s.}
$$

Proof of Theorem 5. In view of the independence between the $\omega_i$ and the $\alpha_i, \beta_i$ we can set $\omega_i = 1$ with no loss of generality. (i) Case $\bar{\pi} < 1$ is trivial. For case $\bar{\pi} = 1$, apply Zaffaroni (2004a)[Lemma 1] with $k = 1$.
(ii) For \( u \geq 0 \) simple calculations yield

\[
\text{cov}_n(\Sigma_{n,t}, \Sigma_{n,t+u}) = \frac{1}{n^2} \sum_{i,j=1}^{n} \frac{\pi_i^u}{\pi_i} \left( \frac{2\alpha_i \alpha_j}{(1 - \delta_{i,j})(1 - \pi_j \pi_i)} \right) \left( 1 + \frac{\pi_i}{1 - \pi_i} + \frac{\pi_j}{1 - \pi_j} \right) > 0
\]

a.s. setting \( \delta_{i,j} = E_n(\beta_i + \alpha_i u_0^2)(\beta_j + \alpha_j u_0^2) = \pi_i \pi_j + 2\alpha_i \alpha_j \) (yielding \( \delta_{i,i} = \nu_i \)).

For the first case \((\bar{\nu} < 1)\) covariance stationarity easily follows. For the autocovariance use \( \pi_i \leq \bar{\pi} \leq \bar{\nu} \bar{\pi} < 1 \) a.s. For the second case \((\bar{\nu} = 1, \alpha_i \geq \alpha > 0 \) a.s.), use \( \delta_{i,j} \leq \nu_i^2 \nu_j^2 \). Then use Lemma 1 with \( z = \nu, k = 1, p = 2 \). To show that non-stationarity arises for \( d_{\bar{\nu}} < -1/2 \), use Zaffaroni (2004a)[Lemma 1], with \( z = \nu, k = 1, \) in \( E_n\Sigma_{n,t}^2 \geq n^{-2} \sum_{i=1}^{n}(1 - \nu_i)^{-1} \). For the autocovariance use \( \pi_i^2 = \nu_i - 2\alpha_i^2 \leq \nu_i - 2\alpha_i \leq 1 - 2\alpha \) a.s. For the third case, from \( \nu = (\beta + \gamma(\bar{\beta} - \beta))^2 + 2\gamma^2(\bar{\beta} - \beta)^2 \) it follows \( \nu \rightarrow \bar{\nu} = \beta^2 \) and \( \pi = \beta + \gamma(\bar{\beta} - \beta) \rightarrow \bar{\beta} \) for \( \beta \rightarrow \bar{\beta} \). Therefore \( \bar{\nu} = 1 \) implies \( \bar{\beta} = \bar{\pi} = 1 \). Since we allow \( \beta_i \) to be equal to zero, covariance stationarity requires \( 3\bar{\gamma}^2 < 1 \). Then, from \( \text{var}_n(\Sigma_{n,t}) = E_n\Sigma_{n,t}^2 - (E_n\Sigma_{n,t})^2 \), we need to evaluate the behaviour of the second moment, since the behaviour of the first moment is given in (i).

Thus \( E_n\Sigma_{n,t}^2 = n^{-2} \sum_{i,j=1}^{n} \left( 1 + \frac{\pi_i}{1 - \pi_i} + \frac{\pi_j}{1 - \pi_j} \right) \) yielding \( n^{-2} \sum_{i,j=1}^{n}(1 + \frac{\pi_i}{1 - \pi_i} + \frac{\pi_j}{1 - \pi_j})(1 - \pi_i \pi_j)^{-1} \leq E_n\Sigma_{n,t}^2 \leq C n^{-2} \sum_{i,j=1}^{n}(1 + \frac{\pi_i}{1 - \pi_i} + \frac{\pi_j}{1 - \pi_j})(1 - \pi_i \pi_j)^{-1} \) a.s., some \( 0 < C < \infty \). The first inequality uses \( \pi_i \pi_j \leq \delta_{i,j} \) a.s. and the second one \( 1 - \delta_{i,j} = 1 - \pi_i \pi_j - 2\alpha_i \alpha_j \geq (1 - \pi_i \pi_j)/C \) a.s., since \( (1 - \pi_i \pi_j)^2 \geq (1 - \pi_i^2)(1 - \pi_j^2) \) and \( (1 - x)/(1 + x) \leq 1 \) for \( 0 \leq x \leq 1 \). Next, use Lemma 1 with \( z = \pi, p = k = 1 \), yielding \( c_\pi > 0 \), as for (i). For the autocovariance \( n^{-2} \sum_{i,j=1}^{n} \pi_i \pi_j^u \gamma_i \gamma_j (1 - \pi_i \pi_j)^{-1} \left( 1 + \frac{\pi_i}{1 - \pi_i} + \frac{\pi_j}{1 - \pi_j} \right) \leq \text{cov}_n(\Sigma_{n,t}, \Sigma_{n,t+u}) \leq C n^{-2} \sum_{i,j=1}^{n} \pi_i \pi_j^u \gamma_i \gamma_j (1 - \pi_i \pi_j)^{-1} \left( 1 + \frac{\pi_i}{1 - \pi_i} + \frac{\pi_j}{1 - \pi_j} \right) \) a.s., some \( 0 < C = C(\bar{\gamma}) < \infty \). In addition to the arguments used to bound the variance, the first inequality uses \( \beta_i \leq \pi_i \) a.s., \( \pi_i \pi_j \leq \delta_{i,j} \) a.s., and the second inequality \( 1 - \pi_i = (1 - \beta_i)(1 - \gamma_i) \geq (1 - \beta_i)(1 - \bar{\gamma}) \) a.s. Finally, use Lemma 1 with \( z = \pi, p = k = 2 \) and \( r = 1 \), recalling \( b_\pi > 0 \). \( \square \)
References


Note: Simulated examples of $X_{n,t}$ with common innovation $u_t$ (dashed line) and idiosyncratic innovations $\epsilon_{i,t}$ (bold line) with $T = 500$ and $n = 1,000$. Simulation design based on pseudo standard Gaussian $u_t$, $\epsilon_{i,t}$, $\omega_i$ uniformly distributed in $[c, 1]$, $c = 1e-14$, $\alpha_i$ Beta($p, q$) distributed in $[0, \bar{\alpha}]$ with parameters $q = q_\alpha$ and $p = q_\alpha \mu_\alpha / (\bar{\alpha} - \mu_\alpha)$ where $\mu_\alpha = E(\alpha_i)$, $\beta_i$ Beta($p, q$) distributed in $[0, \bar{\beta}]$ with parameters $q = q_\alpha$ and $p = q_\alpha \mu_\beta / (\bar{\beta} - \mu_\beta)$ where $\mu_\beta = E(\beta_i)$. Top-left plot is based on $q_\alpha = 2$, $\mu_\alpha = 0.27$, $\bar{\alpha} = 0.3$, $\mu_\beta = 0.6$, $\bar{\beta} = 0.7$; top-right plot is based on same values but $q_\alpha = 0.5$; bottom-left plot is based on same values but $q_\alpha = 0.1$; bottom-right plot is based on $q_\alpha = 0.1$, $\mu_\alpha = 0.297$, $\bar{\alpha} = 0.33$, $\mu_\beta = 0.6$, $\bar{\beta} = 0.7$. 
<table>
<thead>
<tr>
<th>$n$</th>
<th>$q_\pi = 0.1$</th>
<th>$0.2$</th>
<th>$0.3$</th>
<th>$0.7$</th>
<th>$1$</th>
<th>$3$</th>
<th>$q_\pi = 0.1$</th>
<th>$0.2$</th>
<th>$0.3$</th>
<th>$0.7$</th>
<th>$1$</th>
<th>$3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>$1.69e+11$</td>
<td>$1.58e+5$</td>
<td>$1.68e+3$</td>
<td>$20.76$</td>
<td>$11.13$</td>
<td>$3.89$</td>
<td>$0.947$</td>
<td>$0.761$</td>
<td>$0.520$</td>
<td>$0.174$</td>
<td>$0.132$</td>
<td>$0.106$</td>
</tr>
<tr>
<td>100</td>
<td>$8.05e+10$</td>
<td>$2.94e+4$</td>
<td>$2.39e+2$</td>
<td>$1.58$</td>
<td>$0.75$</td>
<td>$0.55$</td>
<td>$0.439$</td>
<td>$0.176$</td>
<td>$0.091$</td>
<td>$0.017$</td>
<td>$0.013$</td>
<td>$0.012$</td>
</tr>
<tr>
<td>1000</td>
<td>$3.05e+10$</td>
<td>$2.56e+9$</td>
<td>$1.38e+6$</td>
<td>$1.94$</td>
<td>$0.20$</td>
<td>$0.04$</td>
<td>$0.019$</td>
<td>$0.307$</td>
<td>$0.483$</td>
<td>$0.010$</td>
<td>$0.003$</td>
<td>$0.001$</td>
</tr>
</tbody>
</table>

$\omega_i$ inverted Gamma (2,1)

$\omega_i$ uniform

This table reports the simulated values for $\text{var}_{n}(E_{n,t})$ (cf. formula (11)) and its ratio with $\text{var}_{n}(U_{n,t})$. The $\pi_i = \alpha_i + \beta_i$ are Beta($p_\pi$, $q_\pi$) distributed in $[0,1-c]$ with parameters $q_\pi$, varying across columns, and $p_\pi = q_\pi \mu_\pi/(1-\mu_\pi)$ where $\mu_\pi = 0.97$, $c = 1e-14$. The first and second panel refer to the case of $\omega_i$ being distributed as inverted Gamma (1,2) and uniform in the interval $[c,1]$, respectively.