AGGREGATION OF SIMPLE LINEAR DYNAMICS: EXACT ASYMPTOTIC RESULTS*

by

Marco Lippi
Dipartimento di Scienze Economiche, Roma "La Sapienza"

and

Paolo Zaffaroni
Research Department, Banca d'Italia

Contents:

Discussion Paper
No. EM/98/350
April 1998

The Suntory Centre
Suntory and Toyota International Centres
for Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 0171-405 7686

* We thank Peter M Robinson for thoroughly reading and commenting on the paper, Renato Flores and Lucien Foldes for very useful suggestions, and Mario Forni for allowing us to use the data employed in Section 7.

Abstract

The paper deals with aggregation of AR(1) micro variables driven by a common and an idiosyncratic shock with random coefficients. We provide a rigorous analysis, based on results on sums of r.v.'s with a possibly infinite first moment, of the aggregate variance and spectral density, as the number of micro units tends to infinity. If the AR coefficients are not bounded away from unity, the aggregate process may exhibit infinite variance and long memory. Surprisingly, if the key parameter of the density function of the AR coefficients lies below a critical value (high density near unity), common and idiosyncratic components have the same importance in explaining aggregate variance, whereas the usual result, i.e. a vanishing importance of the idiosyncratic component, is obtained when the parameter lies above the critical value (low density near unity). Empirical analysis relative to major U.S. macroeconomic series, both in previous literature and in this paper, provides estimates of the parameter below the critical value.

Keywords: aggregation; idiosyncratic-driven fluctuations; long memory; nonstationarity.

JEL No.: C43

1 Introduction

Linear dynamic macroeconomic models based on microeconomic theory usually contain an autoregressive polynomial. This is the case both with models derived from intertemporal maximization and with models in which some form of habit persistence contributes to explain the dynamic behaviour of the agents. Among the innumerable examples see Sargent (1978), Hansen and Sargent (1980, 1991), Nickell (1985). A famous exception is Hall (1992); however, introduction of incomplete information in Hall's model produces an autoregressive polynomial, see Goodfriend (1992), Pischke (1995), Forni and Lippi (1997).

If heterogeneity of agents is allowed, so that both the autoregressive coefficients and the shocks may vary across agents, the simplest form of the autoregressive model is

$$(1 - \alpha_i L)x_{it} = u_t + \epsilon_{it}, \tag{1}$$

where i denotes the individual agent, u_t is a macroeconomic shock, common to all agents, whereas ϵ_{it} is an idiosyncratic shock. Aggregation of model (1), or simple variants of it, is studied in Goodfriend (1992), Pischke (1995), in which α_i is constant across agents, so that only the shocks ϵ_{it} are heterogeneous, whereas α_i may vary in Robinson (1978), Granger (1980), Gonçalves and Gourieroux (1988), Lewbel (1994), Forni and Lippi (1997).

The focus of the present paper is on long memory and non-stationarity as results of the aggregation of model (1) with heterogeneity of the coefficients α_i. The possibility of long memory for macrovariables resulting from aggregation of ARMA microvariables had been firstly suggested in Robinson (1978), where the distribution for the coefficients α_i has not a parametric specification. Systematic analysis is presented in Granger (1980), and Gonçalves and Gourieroux (1988): assuming that the coefficients α_i are drawn independently from a Beta distribution over the interval $[0,\ 1]$, or $[-1,\ 1]$, these authors study the behaviour of the n-th cross-sectional average, i.e. $n^{-1}\sum_{i=1}^{n} x_{it}$, for n tending to infinity, and find that if the distribution of the coefficients α_i is sufficiently dense around unity the limit exhibits long memory, with or without infinite variance. Long memory as a result of aggregation of (1) is also noted in Lewbel (1994), in which however the focus is on different aspects of the aggregate variable.

In another line of research, long-memory processes have been shown to be a valid alternative to the standard ARIMA modelling for many macroeconomic time series: see e.g. Diebold and Rudebusch (1989), Sowell (1992) and Gil-Alaña and Robinson (1997) among others. Thus, as noted e.g. in Haubrich and Lo (1989) and, with some generalizations, in Michelacci and Zaffaroni (1997), aggregation of model (1), modified as to include a time trend, can reproduce relevant features of macroeconomic time series.

As we argue in Section 2, the results obtained so far on aggregation of model (1) are not based on fully rigorous arguments. Rather, heuristic reasoning is often applied, like substituting first moments for averages even when the former are not necessarily finite. As a consequence, the relative importance of common and idiosyncratic components in explaining the aggregate variance cannot be correctly assessed. Furthermore, the literature, with the exception of Robinson (1978), has considered only the case of the Beta distribution for the coefficients α_i.

The main purpose of this paper is a rigorous and complete treatment of the aggregation of model (1), focusing both on the relative importance of common and idiosyncratic components for aggregate variance, and on the persistence of aggregate variables. We will make the general assumption that the α's are drawn independently from a distribution on $[0,\ 1]$ whose probability density $B(\cdot\ ; b)$, depending on the real parameter $b > -1$, is asymptotically equivalent to $C\,(1-\alpha)^b$, for $C > 0$, when α approaches unity. If $b > -1/2$ then almost surely, i.e. for almost every sequence of coefficients α drawn independently from B, the n-th cross-section average of the common component converges in variance to a stationary variable as n tends to infinity, whereas the average of the idiosyncratic component vanishes. By contrast, if $b < -1/2$ then almost surely common and idiosyncratic cross-section averages explode in variance and, strikingly, at the same rate. Thus the elementary result of a vanishing idiosyncratic component does not extend to model (1) if the α's are, so to speak, dense enough near unity. However, there is a crucial difference in the asymptotic behaviours of common and idiosyncratic components for $b < -1/2$: the common component when differenced converges almost surely to a stationary non-zero variable which gives by integration an infinite-variance long-memory variable, whereas the idiosyncratic component, when differenced, tends to zero and nothing can be recovered by integration. These results are valid for any parametric and nonparametric specification of the probability density $B(\cdot\ ; b)$, the Beta representing a particular parametric case.

The basic definitions and the results are given in Sections 3 to 6. The proofs, given in Appendix B, are based partly upon the usual properties of the Hilbert space generated by the x's, partly upon a Lemma—which is based on a generalization of Kolmogorov's strong law of large numbers—on sums of independent random variables not necessarily possessing finite first moment (see Lemma 1, Appendix A). Our model, as defined in Section 3, is very close to the one analyzed in Al-Najjar (1995) and Uhlig (1996). However, here we are mainly interested with the unbounded variance case.

In Sections 3 to 5 we deal with micro and macro stochastic variables and processes, not with their realizations. In Section 6 we deal with a finite time period $1, 2, \ldots, T$ and the corresponding estimated micro and macro variances, conditional on the values assumed for $t \leq 0$. We obtain exact asymptotic rates at which estimated variances converge or diverge almost surely as T and n tend to infinity. In particular, provided that T tends to infinity fast enough with respect to n, the result of divergence at the same rate for common and idiosyncratic averages is reobtained. However, if T is fixed and n tends to infinity, then, irrespectively of whether b is greater or smaller than $-1/2$, the common component tends weakly to a non-zero random variable, whereas the idiosyncratic component variance tends to zero. Thus, when we are facing a typical macroeconomic data set, with series resulting from the aggregation of several millions of individual variables, and a T not much bigger than few hundreds, the commonly held idea that no idiosyncratic component is left in the aggregates appears as fairly sensible.

If model (1), modified as to include a time trend, is assumed as a good approximation to the micro time series, then empirical evidence produced so far points to a value of b between -1 and $-1/2$ (significantly less that $-1/2$; see Gil-Alaña and Robinson (1997) for recent results on several annual macroeconomic U.S. series and a survey of previous outcomes on fractional differencing and macroeconomic time series). In Section 7 we report on an empirical exercise that we have conducted using U.S. quarterly consumption data. Consistently with previous results we find a value of b between -1 and $-3/4$, thus well below $-1/2$. Section 8 concludes.

2 The problem

Let us rewrite here model (1) and add some specifications. We want to deal with a countable infinity of individual processes:

$$(1-\alpha_i L)x_{i,t} = u_t + \epsilon_{i,t},$$

where i is a positive integer, u_t and $\epsilon_{i,t}$ are white noises, $u_t \perp \epsilon_{i,t-k}$ for any i and any integer k, $\epsilon_{i,t} \perp \epsilon_{j,t-k}$, for $i \neq j$ and any integer k. Let σ_u^2 be the variance of u_t and assume that the variance of $\epsilon_{i,t}$ does not depend on i, i.e. that $\mathrm{var}(\epsilon_{i,t}) = \sigma_\epsilon^2$. We are interested in the limit of the process

$$X_{n,t} = \frac{1}{n}\sum_{i=1}^{n} x_{i,t},$$

for n tending to infinity under the assumption that the coefficients α are independently drawn from a distribution on [0, 1] with probability density $B(\alpha)$.

Consider firstly the aggregation of the idiosyncratic component

$$E_{n,t} = \frac{1}{n}\sum_{i=1}^{n} \frac{\epsilon_{i,t}}{1-\alpha_i L}.$$

The corresponding spectral density is

$$\frac{\sigma_\epsilon^2}{2\pi\, n^2}\sum_{i=1}^{n}\frac{1}{\mid 1-\alpha_i e^{-i\lambda}\mid^2} = \frac{\sigma_\epsilon^2}{2\pi\, n}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\mid 1-\alpha_i e^{-i\lambda}\mid^2}\right) \tag{2}$$

Then, following Granger (1980), for large n we replace the term within brackets with the expectation of $1/|1-\alpha e^{-i\lambda}|^2$, obtaining

$$\frac{\sigma_\epsilon^2}{2\pi\, n}\int_0^1 \frac{1}{\mid 1-\alpha e^{-i\lambda}\mid^2} B(\alpha)d\alpha. \tag{3}$$

The spectral density at the zero frequency and the variance of $E_{n,t}$ are respectively

$$\frac{\sigma_\epsilon^2}{2\pi\, n}\int_0^1 \frac{1}{(1-\alpha)^2} B(\alpha)d\alpha, \tag{4}$$

and

$$\frac{\sigma_\epsilon^2}{n}\int_0^1 \frac{1}{(1-\alpha^2)}B(\alpha)d\alpha. \tag{5}$$

Now, as long as the expectations of $(1-\alpha)^{-2}$ and $(1-\alpha^2)^{-1}$, appearing in (4) and (5) respectively, are finite, the replacement is correct and leads, incidentally, to the result that the idiosyncratic component tends to zero in mean-square as n tends to infinity. However, a uniform distribution for the α's is sufficient to yield an infinite integral both in (4) and (5). When this is the case, (4) and (5) no longer help to know the asymptotic behaviour of the corresponding averages. What we really need to know are the rates at which $\sum_{i=1}^n(1-\alpha_i)^{-2}$ and $\sum_{i=1}^n(1-\alpha_i^2)^{-1}$ tend to infinity. Secondly, the standard approach necessarily requires to evaluate the integrals in (4) and (5), obtainable in closed-form only for specific parameterizations, e.g. when $B(\alpha)$ is a Beta density.

Consider now the common component:

$$U_{n,t} = \frac{1}{n}\sum_{i=1}^n \frac{1}{(1-\alpha_i L)}u_t.$$

This may be rewritten as

$$U_{n,t} = u_t + \left(\frac{1}{n}\sum_{i=1}^n \alpha_i\right)u_{t-1} + \left(\frac{1}{n}\sum_{i=1}^n \alpha_i^2\right)u_{t-2} + \cdots. \tag{6}$$

Then, following Gonçalves and Gourieroux (1988), we replace (6) with

$$U_t = u_t + \mu_1 u_{t-1} + \mu_2 u_{t-2} + \cdots, \tag{7}$$

where μ_k is the k-th moment from zero of the density $B(\alpha)$ and therefore the limit, almost surely, of $\sum_{i=1}^n \alpha_i^k/n$. The spectral density at zero and the variance of U_t are

$$\frac{\sigma_u^2}{2\pi}(1+\mu_1+\mu_2+\cdots)^2 \tag{8}$$

and

$$\sigma_u^2(1+\mu_1^2+\mu_2^2+\cdots), \tag{9}$$

respectively. However, firstly we must observe that the vectors U_{nt} and U_t belong to the infinite-dimensional Hilbert space spanned by u_{t-k}, for $k =$

$0, \infty$. Convergence of each of the coefficients of (6) to the corresponding coefficient of (7) is a necessary but not sufficient condition for convergence of U_{nt} to U_t in mean-square, i.e. in the metric of the Hilbert space. Secondly, when (8) or (9) are not finite, the replacement, like in the idiosyncratic case, does not give sufficient information on the asymptotic behaviour of the finite-sample averages. Thus, we will have to deal simultaneously with two problems: firstly, the asymptotic behaviour of averages that correspond to random variables with no finite first moment; secondly, the asymptotic behaviour of stochastic variables which take values in an infinite-dimensional vector space. Moreover, the α's will be drawn from an arbitrary density function, instead of the Beta employed almost exclusively in the literature on the problem.

3 Basic definitions and assumptions

In the sequel $\sim$ will denote asymptotic equivalence, c_θ, C_θ will denote constants depending on a parameter θ.

Assumption I. *Let G be the interval $[0,\ 1]$ and B a family of absolutely continuous distributions on G, depending upon a real parameter $b \in (-1,\ \infty)$, whose densities are indicated by $B(\cdot\ ; b)$. We assume that for $\alpha \to 1^-$, there exists a $C_b > 0$ such that*

$$B(\alpha; b) \sim C_b\,(1-\alpha)^b. \tag{10}$$

Remarks. (1) We point out that Assumption I does not impose any constraint on the behaviour of the probability density $B(\cdot\ ; b)$ within any given interval $[0,\ \bar{\alpha}]$, with $\bar{\alpha} < 1$, and is therefore typically non-parametric.

(2) All our results still hold if we modify Assumption I in the following way: for any member of the family B there exist a b, with $-1 < b < \infty$, and a C such that the density is asymptotically equivalent to $C(1-\alpha)^b$ for $\alpha \to 1^-$, so that there is a subset of $B_b \subset B$ corresponding to each b. Assuming, as we have done, that there is only one element in B_b for each b has only the effect of simplifying the notation. As a particular case of the more general assumption, we could assume that the members of B depend on a vector $\theta \in \Theta \in R^s$, $s \geq 1$, provided that there exists a function $b : \Theta \to (-1, \infty)$

such that $B(\alpha;\theta) \sim C_\theta(1-\alpha)^{b(\theta)}$, as $\alpha \to 1^-$. An example with more than one parameter is that of the Beta density function.

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and let us denote by L_2^0 the Hilbert space of all real-valued, square-integrable, zero-mean random variables defined on Ω. Restricting to L_2^0 implies that var(x) =E(x^2) and that mean-square convergence is equal to convergence in variance.

We want to give a precise meaning to the idea that any 'individual' in G is endowed with an individual random process, whose variables belong to L_2^0. Random processes belonging to different individuals are orthogonal to each other and are orthogonal to a random process which is 'common' to all individuals. A rigorous definition of the individual processes requires that we consider functions defined on G with values in L_2^0, i.e. random variables associating with any $\alpha \in G$ a real random variable defined on Ω.

Assumption II *There exist: (1) a family of random variables*

$$\phi_t \,:\, G \to L_2^0,$$

for $t = -\infty, \infty$, with $\phi_t(\alpha)$ denoted by $\epsilon_{\alpha,t}$; (2) a family $u_t \in L_2^0$, $t = -\infty, \infty$, fulfilling the following properties

(i) u_t is a white-noise process.

(ii) The probability distribution of $\epsilon_{\alpha,t}$ is independent of α and t. Moreover, $\epsilon_{\alpha,t}$ is a white-noise process for any $\alpha \in G$.

(iii) The random variables u_t and $\epsilon_{\alpha,s}$ are orthogonal for any $\alpha \in G$, any t and s. For $\alpha \neq \beta$, the random variables $\epsilon_{\alpha,t}$ and $\epsilon_{\beta,s}$ are orthogonal for any t and s.

The white noises u_t and $\epsilon_{\alpha,t}$ will be called common and idiosyncratic shocks respectively.

Considering first and second moments, Assumption II implies that:

$$\begin{aligned}
E(\epsilon_{\alpha,t}) &= 0, \text{ for any } \alpha \text{ and } t. \\
cov(\epsilon_{\alpha,t}, \epsilon_{\alpha,s}) &= \begin{cases} \sigma_\epsilon^2, & s = t \text{ and any } \alpha, \quad 0 < \sigma_\epsilon^2 < \infty, \\ 0, & s \neq t, \text{ and any } \alpha. \end{cases} \\
E(u_t) &= 0. \\
cov(u_t, u_s) &= \begin{cases} \sigma_u^2, & s = t, \quad 0 < \sigma_u^2 < \infty, \\ 0, & s \neq t. \end{cases} \\
cov(u_t, \epsilon_{\alpha,s}) &= 0, \text{ for any } \alpha \text{ and any } t, s\,. \\
cov(\epsilon_{\alpha,t}, \epsilon_{\beta,s}) &= 0, \text{ for any } \alpha \neq \beta \text{ and any } t, s\,.
\end{aligned}$$

The definition of the variables that we want to aggregate is based on u_t and $\epsilon_{\alpha,t}$. Precisely, the family $\psi_t : G \to L_2^0$ is defined as

$$\psi_t(\alpha) = x_{\alpha,t} = \frac{1}{1-\alpha L} u_t + \frac{1}{1-\alpha L} \epsilon_{\alpha,t},$$

for $t = -\infty, \infty$. Thus any individual $\alpha \in G$ is now endowed with the process $x_{\alpha,t}$, whose variables belong to L_2^0, resulting from the summation of a common and an idiosyncratic autoregressive component.

Remark. The variable $x_{\alpha,t}$ is I(0). However, no important change in our results occurs if in the definition above we put $(1-L)x_{\alpha,t}$ instead of $x_{\alpha,t}$, so that the microvariable is I(1) (see Lewbel's model reported in Section 7). Obviously the order of integration of the aggregate variable would be increased by one.

Lastly, we want to consider infinite sequences of independent drawings from G. It will be convenient to formalize this by considering the cartesian product $\mathcal{G} = \prod_{i=1}^{\infty} G_i$, with $G_i = G$, with the product probability measure. If

$$A = (\alpha_1, \alpha_2, \ldots) \in \mathcal{G},$$

A_n will indicate the truncation $(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Given $A \in \mathcal{G}$ we define the sample averages

$$\begin{aligned} U_{A_n,t} &= \frac{1}{n}\sum_{i=1}^{n} \frac{1}{1-\alpha_i L} u_t \\ E_{A_n,t} &= \frac{1}{n}\sum_{i=1}^{n} \frac{1}{1-\alpha_i L} \epsilon_{\alpha_i,t} \\ X_{A_n,t} &= U_{A_n,t} + E_{A_n,t} \end{aligned}$$

The problems we want to study are: (1) whether there exists a vector in L_2^0 to which $X_{A_n,t}$ converges almost surely, i.e. almost everywhere in $\mathcal{G}$, in the metric of L_2^0, i.e. in the mean square of real-valued variables defined on Ω; (2) under what circumstances this limit can be defined as the first moment of ψ_t, meant as a variable defined on G with values in L_2^0. We shall be able to give a complete answer, depending on the parameter b, separately for the components $U_{A_n,t}$ and $E_{A_n,t}$.

Closely linked to the variables ψ_t is the variable $f : G \to L[-\pi, \pi]$, where $L[-\pi, \pi]$ is the space of real-valued absolutely integrable functions defined on $[-\pi, \pi]$, which associates with any α the spectral density of $x_{\alpha,t}$. By Assumption II,iii, f associates with any $\alpha \in G$ the function:

$$s_\alpha(\lambda) = \frac{\sigma_u^2}{2\pi} \frac{1}{|1 - \alpha e^{-i\lambda}|^2} + \frac{\sigma_\epsilon^2}{2\pi} \frac{1}{|1 - \alpha e^{-i\lambda}|^2}.$$

Moreover, for the spectral densities of $U_{A_n,t}$ and $E_{A_n,t}$ we have, by Assumption II:

$$\begin{aligned} s_{A_n}^U(\lambda) &= \frac{\sigma_u^2}{2\pi} \frac{1}{n^2} \sum_{i,j=1}^{n} \left(\frac{1}{1 - \alpha_i e^{-i\lambda}} \frac{1}{1 - \alpha_j e^{i\lambda}} \right), \\ s_{A_n}^E(\lambda) &= \frac{\sigma_\epsilon^2}{2\pi} \frac{1}{n^2} \sum_{i=1}^{n} \frac{1}{|1 - \alpha_i e^{-i\lambda}|^2}, \end{aligned}$$

where the second formula holds for all sequences A with no repetitions, and therefore a.e. in $\mathcal{G}$ (since $B(\cdot\,; b)$ has no atoms), the question being whether there exist functions in $L[-\pi, \pi]$ to which such spectral densities converge.

Lastly, by Assumption II, the variances of $U_{A_n,t}$ and $E_{A_n,t}$ are, respectively:

$$\begin{aligned} V_{A_n}^U &= \frac{\sigma_u^2}{n^2} \sum_{i,j=1}^{n} \frac{1}{1 - \alpha_i \alpha_j} \\ V_{A_n}^E &= \frac{\sigma_\epsilon^2}{n^2} \sum_{i=1}^{n} \frac{1}{1 - \alpha_i^2}, \end{aligned}$$

the second formula holding for all sequences A with no repetitions, and therefore a.e. in $\mathcal{G}$.

4 The idiosyncratic component

We begin with a theorem that gives a complete description of the asymptotic behaviour of $V_{A_n}^E$. We only need a few premises. Firstly, Assumptions I and II will be tacitly supposed to hold in all the results below. Secondly, the phrase "for A almost everywhere in $\mathcal{G}$" has been preferred to "almost surely", which is customary when stating strong laws of large numbers. Lastly, let us introduce the concept of *sufficiently fast growing* sequence :

Definition A sequence $\{f_n, n = 1, 2, \ldots\}$ of positive real numbers is sufficiently fast growing (SFG) if

$$\sum_{n=1}^{\infty} 1/(n\, f_n) < \infty.$$

Remark. For example, the sequence $f_i = \log i$ is not SFG, whereas the sequence $f_i = (\log i)^d$, for any $d > 1$ is SFG. The definition does not establish an upper bound to the speed at which an SFG sequence may grow. However, as it will be clear below, we are interested in SFG sequences that grow as slowly as possible, but not so slowly as to violate the above condition.

Theorem 1 *As* $n \to \infty$

(i) If $b > 0$*, then for* A *a.e. in* $\mathcal{G}$ *there exists a positive real* C_A *(depending on* A*) such that*

$$V_{A_n}^E \leq C_A n^{-1}.$$

(ii) If $b = 0$*, then for* A *a.e. in* $\mathcal{G}$ *there exist positive reals* c_A *and* C_A *such that*

$$c_A n^{-1} \leq V_{A_n}^E \leq C_A n^{-1} \log n.$$

(iii) Let $b < 0$*. Given an SFG sequence* $\{f_n\}$*, for* A *a.e. in* $\mathcal{G}$ *there exist positive reals* c_A *and* $C_{A,f}$ *(depending both on* A *and* $\{f_n\}$*) such that*

$$c_A n^{-\frac{2b+1}{b+1}} \leq V_{A_n}^E \leq C_{A,f}\, n^{-\frac{2b+1}{b+1}} f_n^{-\frac{b}{b+1}}.$$

Remarks. (1) It must be pointed out that for $b > -1/2$ Theorem 1 gives the usual result that the average of the idiosyncratic component vanishes in mean-square as n tends to infinity, i.e. that $E_{A_n,t}$ tends to zero in L_2^0 for A a.e. in $\mathcal{G}$.

(2) For $b = -1/2$ Theorem 1 in inconclusive because the lower bound is constant, but the upper bound is $C_{A,f} f_n$.

(3) For $b < -1/2$ we have a rather striking result: the variance of $E_{A_n,t}$ tends to infinity at least as fast as $n^{-(2b+1)/(b+1)}$, for A a.e. in $\mathcal{G}$.

(4) When $b > 0$ (case (i)), we can easily obtain an exact rate result by some

version of the law of iterated logarithm (Stout 1974). The same is true for case (i) of Theorem 2, 8 and 9.

The following two theorems describe the asymptotic behaviour of the spectral density of $E_{A_n,t}$ at $\lambda = 0$, and $\lambda \neq 0$ respectively.

Theorem 2 *As $n \to \infty$*
(i) If $b > 1$, then for A a.e. in $\mathcal{G}$ there exists a positive real C_A such that

$$s^E_{A_n}(0) \leq C_A n^{-1}.$$

(ii) If $b = 1$, then for A a.e. in $\mathcal{G}$ there exist positive reals c_A and C_A such that

$$c_A n^{-1} \leq s^E_{A_n}(0) \leq C_A n^{-1} \log n.$$

(iii) Let $b < 1$. Given the SFG sequence $\{f_n\}$, for A a.e. in $\mathcal{G}$ there exist c_A and $C_{A,f}$ such that

$$c_A n^{-\frac{2b}{b+1}} \leq s^E_{A_n}(0) \leq C_{A,f}\, n^{-\frac{2b}{b+1}} f_n^{-\frac{1-b}{b+1}}.$$

Remark. The spectral density of $E_{A_n,t}$ at $\lambda = 0$ tends to infinity when $b < 0$, irrespective of whether the variance of $E_{A_n,t}$ tends to zero, for $b > -1/2$, or to infinity, for $b < -1/2$.

Theorem 3 *As $n \to \infty$, for A a.e. in $\mathcal{G}$ the spectral density of $E_{A_n,t}$ tends to zero pointwise in $(0, \pi]$ for any $b > -1$ (and therefore irrespective of whether the variance tends to zero or to infinity). If $b > 0$ pointwise convergence to zero occurs in $[0, \pi]$.*

Remark. The result of Theorem 1 for $b < -1/2$ must be integrated with Theorem 3: when n is huge and $b < -1/2$ the contribution of $E_{A_n,t}$ to the variance of $X_{A_n,t}$ does not vanish but concentrates in a thin peak in the vicinity of $\lambda = 0$.

The usual statement that the idiosyncratic component is washed away by aggregation is valid in our model only for $b > -1/2$. However, the next theorem provides a weak generalization of the usual statement: the variance of the *innovation* to the aggregate idiosyncratic variable tends to zero irrespective of the value of b. For, consider the Wold representation

$$E_{A_n,t} = a^E_{A_n}(L)\eta_{A_n,t},$$

where $\eta_{A_n,t}$ is white noise, $a^E_{A_n}(0) = 1$, while $a^E_{A_n}(L)$ has no roots of modulus smaller that unity, i.e. is fundamental (Rozanov 1963).

Theorem 4 *For A a.e. in $\mathcal{G}$*

$$\lim_{n\to\infty} \mathrm{var}(\eta_{A_n,t}) = 0.$$

The last theorem gives a negative answer to the question whether the asymptotic behaviour of $E_{A_n,t}$, for $b < -1/2$, could be recovered by firstly differencing, taking the limit and then integrating.

Theorem 5 *For any $b > -1$, and A a.e. in $\mathcal{G}$, the first difference $(1 - L)E_{A_n,t}$ converges to zero in L^0_2.*

The results obtained deserve some comments. Firstly, it is a fairly elementary observation that for A a.e. in $\mathcal{G}$ the average $E_{A_n,t}$ either converges zo zero or does not converge. For,

$$E_{A_n,t} = \frac{1}{n}\epsilon_{\alpha_1,t} + \frac{1}{n}\epsilon_{\alpha_2,t} + \cdots + \frac{1}{n}\epsilon_{\alpha_n,t}.$$

If $E_{A_n,t}$ converges to $\tilde{E}$ then

$$\tilde{E} = c_1\epsilon_{\alpha_1,t} + c_2\epsilon_{\alpha_2,t} + \cdots = \sum_{i=1}^{\infty} c_i\epsilon_{\alpha_i,t}.$$

Since the $\epsilon_{\alpha_i,t}$ are mutually orthogonal for A a.e. in $\mathcal{G}$, then convergence of $E_{A_n,t}$ to $\tilde{E}$ implies that c_i is the limit of $1/n$, i.e. that $c_i = 0$. Now, Theorem 1 considerably improves on this observation: for $b > -1/2$, $E_{A_n,t}$ converges to zero, for $b < -1/2$, it diverges in variance and we know the rate of divergence.

Secondly, Theorems 3, 4 and 5 are very important to understand the difference between the asymptotic behaviour of common and idiosyncratic component. For, consider the average

$$U_{A_n,t} = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{1-\alpha_i L} u_t.$$

Since the polynomial on the RHS is not necessarily fundamental (it might have some roots inside the unit circle), let us write the Wold representation of $U_{A_n,t}$ as

$$U_{A_n,t} = a^U_{A_n}(L) v_{A_n,t},$$

with $a^U_{A_n}(L)$ fundamental, $a^U_{A_n}(0) = 1$ and $v_{A_n,t}$ white noise. As is well known var$(v_{A_n,t})$ $\geq$var(u_t), equality holding if u_t is fundamental for $U_{A_n,t}$. Therefore, unlike the innovation of the idiosyncratic component average, the innovation of the common component average can never tend to zero, nor does the spectral density tend to zero for any $\lambda \neq 0$. Further observations on the difference between common and idiosyncratic asymptotic behaviour are postponed to the next section.

Thirdly, one may wonder whether the consequence of Theorem 1, that $E_{A_n,t}$ tends to zero a.s. when $b > -1/2$, can be read as a strong law of large numbers, i.e. the statement that $E_{A_n,t}$ tends a.s. to the mean of the function associating with $\alpha \in G$ the variable $(1/(1-\alpha L)\epsilon_{\alpha,t} \in L^0_2$. The answer depends on the definition of the mean, i.e. of the integral of a function defined on G with values in L^0_2. This problem has been recently addressed in economic literature in Al-Najjar (1995) and Uhlig (1996). In both papers the definition of Bochner and Pettis integrals are recalled. Given a function $F : G \to L^0_2$, the Bochner integral is a straightforward generalization of the Lebesgue integral, i.e. the limit of the sum of simple functions approximating F in mean square. The Pettis integral of F is a function $y \in L^0_2$ such that $\int_\Omega gyd\mathcal{P} = \int_G \left[\int_\Omega gF(\alpha)d\mathcal{P}\right] B(\alpha,b)d\alpha$, for any $g \in L^0_2$. In both papers it is observed that the Bochner integral of an idiosyncratic variable does not exist (Al-Najjar (1995), p. 1220; Uhlig (1996), p. 46). However, both $\epsilon_{\alpha,t}$ and $(1/(1-\alpha L)\epsilon_{\alpha,t}$ are Pettis integrable and the Pettis integral is zero. Uhlig (1996) proves this result for a bounded idiosyncratic variable (p. 45), but the proof can be easily extended to an unbounded idiosyncratic variable (for instance, by using the argument employed in Al-Najjar (1995), p. 1200, footnote). Thus Theorem 1, for $b > -1/2$, implies a strong law of large numbers for the variable $(1/(1-\alpha L)\epsilon_{\alpha,t}$, with the mean defined as the Pettis integral. However, for $b < -1/2$, the Pettis mean is still zero whereas $E_{A_n,t}$ diverges in variance a.e. in $\mathcal{G}$, so that the strong law breaks down.

5 The common component

Let us denote by $\mu_{b,k}$ the k-th moment from zero of the probability density $B(\cdot\,; b)$. We recall that for any integer k and real b

$$\int_0^1 u^k(1-u)^b du = \frac{\Gamma(k+1)\Gamma(b+1)}{\Gamma(k+b+2)}.$$

By Stirling formula and Assumption I, as $k \to \infty$

$$\mu_{b,k} \sim C\,\Gamma(b+1)k^{-(b+1)},$$

for $C > 0$. Therefore $\sum_{k=0}^{\infty} \mu_{b,k}$ converges if and only if $b > 0$, while $\sum_{k=0}^{\infty} \mu_{b,k}^2$ converges if and only if $b > -1/2$. Thus for $b > -1/2$ the series

$$U_{b,t} = (1 + \mu_{b,1}L + \mu_{b,2}L^2 + \cdots)u_t = M_b(L)u_t$$

is well defined in L_2^0. Moreover, going back to the definition of $U_{A_n,t}$ (Section 3), and putting

$$M_{A_n}(L) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{1-\alpha_i L},$$

each of the coefficients of $M_{A_n}(L)$ converges, as already observed in Section 2, to the corresponding coefficient of $M_b(L)$ for A almost everywhere in $\mathcal{G}$. The following theorem gives a positive answer to the question whether there is also convergence of $U_{A_n,t}$ to $U_{b,t}$ in mean-square.

Theorem 6 *If $b > -1/2$ then for A a.e. in $\mathcal{G}$*

$$\lim_{n\to\infty} U_{A_n,t} = U_{b,t} \quad in\ L_2^0,$$

that is

$$\lim_{n\to\infty} \text{var}(U_{A_n,t} - U_{b,t}) = 0.$$

An easy consequence is the following:

Theorem 7 *If $b > -1/2$, then $s^U_{A_n}(\lambda)$, the spectral density of $U_{A_n,t}$, converges in first-mean to the spectral density of $U_{b,t}$ for A a.e. in $\mathcal{G}$.*

The next two theorems parallel Theorems 1 and 2.

Theorem 8
(i) If $b > -1/2$, then for A a.e. in $\mathcal{G}$

$$\lim_{n\to\infty} V^U_{A_n} = \mathrm{var}(U_{b,t})$$

(this is a corollary to Theorem 6 and is listed here for completeness).
(ii) If $b = -1/2$, then for A a.e. in $\mathcal{G}$ there exist positive reals c_A and C_A such that

$$c_A \log n \le V^U_{A_n} \le C_A(\log n)^2.$$

(ii) Let $b < -1/2$. Given an SFG sequence $\{f_n\}$, for A a.e. in $\mathcal{G}$ there exist positive reals c_A and $C_{A,f}$ such that

$$c_A n^{-\frac{2b+1}{b+1}} \le V^U_{A_n} \le C_{A,f}\, n^{-\frac{2b+1}{b+1}} f_n^{-\frac{2b+1}{b+1}}.$$

Remark. For $b < -1/2$ the variance of the common component average diverges. This is precisely the same result obtained for the idiosyncratic component average. Moreover, the speed at which $V^U_{A_n}$ diverges is precisely the speed at which $V^E_{A_n}$ diverges. This may be surprising because, leaving aside the difference between σ^2_u and σ^2_ϵ, given the matrix whose entries are $1/(1-\alpha_i\alpha_j)$, $V^E_{A_n}$ is obtained by summing only the n diagonal entries whereas $V^U_{A_n}$ is obtained by summing all the n^2 entries. To reach an intuitive understanding of this fact let $\tilde{\alpha}_n$ denote the maximum of A_n. The asymptotic behaviour of $V^U_{A_n}$ depends on the speed at which $\tilde{\alpha}_n$ tends to unity, i.e. on how fast an α_h, with $h > n$ and $\alpha_h > \alpha_n$, is drawn. It must be pointed out that, irrespective of how fast this occurs, when α_h is drawn the diagonal term $1/(1-\alpha_h^2)$ dominates the terms $1/(1-\alpha_i\alpha_h)$, with $i < h$ (so that $\alpha_i < \alpha_h$). Now, if $b < -1/2$ the maximum of A_n shifts so frequently towards unity that only the diagonal terms matter, so that there is no difference in the asymptotic behaviour of the two components. By contrast, when $b > -1/2$, occurrence of big α's is not so frequent as to compensate for the different number of terms in $V_{A_n}^U$ and $V_{A_n}^E$ respectively, so that the first converges to a finite limit while the second vanishes.

Theorem 9
(i) If $b > 0$, then for A a.e. in $\mathcal{G}$

$$\lim_{n\to\infty} s_{A_n}^U(0) = \frac{\sigma_u^2}{2\pi}\left(\sum_{i=1}^{\infty}\mu_{b,i}\right)^2 = s_b^U(0).$$

(ii) If $b = 0$, then for A a.e. in $\mathcal{G}$ there exist positive reals c_A and C_A such that

$$c_A \log n \le s_{A_n}^U(0) \le C_A(\log n)^2.$$

(ii) Let $b < 0$. Given an SFG sequence $\{f_n\}$, for A a.e. in $\mathcal{G}$ there exist positive reals c_A and $C_{A,f}$ such that

$$c_A n^{-\frac{2b}{b+1}} \le s_{A_n}^U(0) \le C_{A,f} n^{-\frac{2b}{b+1}} f_n^{-\frac{2b}{b+1}}.$$

As already observed, for $b > -1/2$ the sequence of the moments $\mu_{b,k}$ is square summable, so that $U_{b,t}$ belongs to L_2^0. As a consequence,

$$\sigma_u^2\left|\sum_{s=0}^{\infty}\mu_{b,s}e^{-i\lambda s}\right|^2 < \infty \tag{11}$$

for λ almost everywhere in $[0\ \pi]$ (this is an application of Riesz-Fischer Theorem). In the next theorem we prove that the LHS in (11) converges for any $b > -1$ and any $\lambda > 0$. Thus $s_b^U(\cdot)$ makes sense for any $b > -1$, denoting a spectrum for $b > -1/2$, a pseudo-spectrum for $b \le -1/2$.

Theorem 10
(i) For $b > -1$ and $\lambda \neq 0$

$$s_b^U(\lambda) = \frac{\sigma_u^2}{2\pi}\left|\sum_{s=0}^{\infty}\mu_{b,s}e^{-i\lambda s}\right|^2 = \frac{\sigma_u^2}{2\pi}\left|\int_G \frac{1}{1-\alpha e^{-i\lambda}}B(\alpha;b)d\alpha\right|^2 < \infty.$$

If $b > 0$ finiteness holds for any λ.
(ii) For $b > -1$ and A a.e. in $\mathcal{G}$ the spectral density of $U_{A_n,t}$, i.e. $s_{A_n}^U(\lambda)$, converges to $s_b^U(\lambda)$ for any $\lambda \in (0,\ \pi]$. For $b > 0$ pointwise convergence occurs in $[0,\pi]$.

The next theorem contains a result on the memory of $U_{b,t}$ for any feasible b, i.e. $b > -1$ (more rigorously, on $s_b^U(\lambda)$ in the vicinity of $\lambda = 0$), and on the memory of the limit of $(1-L)U_{A_n,t}$ for $b \leq -1/2$. The theorem neatly shows the link between the cross-sectional distribution of the coefficients α_i and the second order properties of the aggregate process: the fatter the tail of the cross-sectional distribution around unity, the stronger the memory of the aggregate process. It is worthwhile noting that since our model for the cross-sectional density is non-parametric (see the Remarks on Assumption I), this simple yet powerful result is valid for any particular specification.

Theorem 11 *As* $\lambda \to 0^+$

$$s_b^U(\lambda) \sim \begin{cases} c\lambda^{2b}, & b < 0 \\ c' \log(\frac{1}{\lambda}), & b = 0 \\ c'', & b > 0 \end{cases}$$

with c, c', c'' *positive.*
If $b \leq -1/2$ *then for* A *a.e. in* $\mathcal{G}$ $(1-L)U_{A_n,t}$ *converges in variance to a stationary non-zero limit, whose order of integration is* $-(1+b)$.

Remarks. (1) The asymptotic behaviour of $s_b^U(\cdot)$, for $\lambda \to 0^+$, has been obtained also for $b \leq -1/2$, when $s_b^U(\cdot)$ is a pseudo-spectrum. Summing up, for $b > -1/2$ the spectral density $s_b^U(\cdot)$ corresponds to a stationary variable (which is already known) with long memory for $b \leq 0$ and short memory for $b > 0$; whereas for $b \leq -1/2$ $s_b^U(\cdot)$ is a long-memory non-stationary pseudo-spectrum.
(2) An immediate implication of Theorem 11 is that the micro-parameter b can be directly estimated from the estimate of the slope near the zero frequency of the periodogram of the macro-data.

We conclude with two comments. Firstly, the dichotomy observed for the idiosyncratic average, convergence to zero or no convergence, does not hold for the common component: we have divergence for $b < -1/2$ but convergence to a non-zero vector for $b > -1/2$. The same observation holds for the spectral density of the common component, that does not vanish asymptotically for any λ, unlike what has been shown for the idiosyncratic component in Theorem 4. Lastly, comparing the second part of Theorem

11 to Theorem 5, we see that the first difference of the common component average converges to a stationary non-zero limit irrespective of even when $b < -1/2$, whereas differencing annihilates the limit of the idiosyncratic component average even when $b > -1/2$.

Secondly, going back to the framework and results in Al-Najjar (1995) and Uhlig (1996), for $b > -1/2$ the variable $U_{b,t} \in L_2^0$ is the Bochner integral of the common component, meant as the stochastic variable from G to L_2^0 associating $(1/1-\alpha L)u_t$ with α (as recalled in Section 4, the Bochner integral of the idiosyncratic does not exist). This may be shown either by applying the Measurability Theorem mentioned in Al-Najjar (1995), p. 1219 (our common component is weakly measurable and has an essentially separable range), or, more directly, by writing the common component as $u_t+\alpha u_{t-1}+\alpha^2 u_{t-2}+\cdots$, and integrating term by term. Thus our Theorem 6, for $b > -1/2$, states a strong law of large numbers for the random variable associating $(1/1-\alpha L)u_t$ with α, the mean being defined as the Bochner integral. For $b < -1/2$ neither the Bochner nor the Pettis integral of the common component exist.

6 Conditional results

In the previous sections, we have analyzed the behaviour of variance and spectral density for the common and idiosyncratic component, evaluating the expectations unconditionally with respect to time, and then letting n go to infinity.

We will now consider a different situation, when both T and n are finite. Thus we will evaluate the variance of both the idiosyncratic and common component with respect to a sample of dimension $T \times n$, hence conditioning on the observations prior to time $t = 1$. In the next theorem, on one hand, we corroborate the 'unconditional' (with respect to time) results of the previous sections and, one the other, we study the behaviour of the aggregate process when, say, n is big compared to T. Only the variance will be considered here, though the same method could be applied to the memory.

Let $\mathcal{F}_t$ be the sigma-algebra generated by $\{u_v, \epsilon_{i,s}\, i = 1, 2, \dots, \mid s, v \leq t\}$. Henceforth $\mathrm{var}(x \mid \mathcal{F}_t)$ denotes the variance of the r.v. x conditional on $\mathcal{F}_t$.

Theorem 12 *For A a.e. $i\,\mathcal{G}$, as T, $n \to \infty$*
(i) If $b \neq -1/2$:

$$\mathrm{var}(U_{A_n,T} \mid \mathcal{F}_0) \sim$$
$$\sigma_u^2 \left(\frac{c_{U,b}}{-(2b+1)} (T^{-(2b+1)} - 1) + \frac{c'_{U,b}}{-b}\, g(n,T,b) \frac{(T^{-b}-1)}{n} + \frac{4c''_{U,b}}{-(3b+1)} g(n,T,b)^{1/2} \frac{(T^{-(3b+1)/2}}{n^{1/2}} \right.$$
$$var(E_{A_n,t} \mid \mathcal{F}_0) \sim \sigma_\epsilon^2 \left(c_{E,b}\, g(n,T,b)^{1/2} \frac{T^{-b/2+1/2}}{n^{3/2}} + c'_{E,b} \frac{T^{-b}}{n} \right)$$

(ii) If $b = -1/2$:

$$\mathrm{var}(U_{A_n,T} \mid \mathcal{F}_0) \sim$$
$$\sigma_u^2 \left(c_{U,-1/2} log(T) + 2\, c'_{U,-1/2} g(n,T,-1/2) \frac{T^{1/2}}{n} + 8 c''_{U,-1/2} g(n,T,-1/2)^{1/2} \frac{T^{1/4}}{n^{1/2}} \right)$$
$$\mathrm{var}(E_{A_n,T} \mid \mathcal{F}_0) \sim \sigma_\epsilon^2 \left(c_{E,-1/2}\, g(n,T,-1/2)^{1/2} \frac{T^{3/4}}{n^{3/2}} + c'_{E,-1/2} \frac{T^{1/2}}{n} \right)$$

for positive constants $c_{U,b}$, $c'_{U,b}$, $c''_{U,b}$, $c_{E,b}$, $c'_{E,b}$, depending on b, where $g(n,T,b) =$ $\log\log(n\,\mu_{b,2T})$.

Remarks. (1) It is easily seen that the rates of convergence or divergence established in Theorem 12 apply with no modification if we consider the expectation of the estimated variances (idiosyncratic or common) of a sample over the period $1, T$.
(2) When $T/n \to 0$, we obtain a version of the commonly held statement concerning the relative importance of the common and idiosyncratic component, namely that the idiosyncratic component vanishes irrespective of b. A fixed T is an important particular case.
(3) When $b \neq -1/2$ and $n \sim cT^{b+1}$ for a $c > 0$, one obtains that either the idiosyncratic component vanishes, for $b > -1/2$, or diverges, for $b < -1/2$. In this last case it diverges exactly at the same rate as the common component, as it would be expected given the results obtained in Theorems 1 and 8.
(4) Assume that $b = -1/2$ and $n \sim cT^{b+1}$, $c > 0$. This is the only case when the variance of the idiosyncratic component has a bounded positive limit a.s. In this case we are able to find a sharper result than with the 'unconditional'

analysis, in which the result was inconclusive for $b = -1/2$ (see Remark (2) to Theorem 1). On the other hand, the common component diverges at the rate $\log(T)$, consistently with Theorem 8-(ii).

(5) When $n \sim c\,T^{b+1}$ the function $g(n, T, b)$ is asymptotically constant.

7 An empirical application: U.S. consumption expenditure

In our empirical exercise we have studied U.S. consumption expenditure. Our data is U.S. per-capita personal consumption expenditure, both including and excluding durables (source NIPA, quarterly data, I-1947:IV-1991).[1] The same time series, up to minor differences, have been employed in Lewbel (1994), whose results are obtained by fitting ARIMA$(p, 0, q)$ models, with intercept, to the first-difference of the series with $0 \leq p \leq 4, 0 \leq q \leq 1$, thus imposing a unit root plus a linear time trend. Lewbel analyses the relationship between the moments of the distribution of the α's and the coefficients of the AR(∞) representation of the aggregate series. By and large, our analysis and Lewbel's are complementary, with Lewbel's focussing on the low-order moments while our analysis concentrates on the density of the α's around unity. However, in Lewbel's paper the possibility of an order of integration different from unity is not explored. The estimated model for aggregate consumption is

$$(1-L)C_t = \tau + \frac{a(L)}{b(L)} u_t,$$

where $a(L)$ and $b(L)$ are finite polynomials, which is consistent with the micromodel

$$(1-L)c_{it} = \tau_i + \frac{1}{1-\alpha_i L} u_t + \frac{1}{1-\alpha_i L} \epsilon_{it}.$$

By contrast, we estimate the order of integration of the aggregate C_t and find a number between 3/4 and unity, significantly smaller than unity. As a consequence, aggregate consumption is consistent with the micromodel

$$c_{it} = \rho_i + \tau_i t + \frac{1}{1-\alpha_i L} u_t + \frac{1}{1-\alpha_i L} \epsilon_{it},$$

[1] For a detailed description of our data set see Forni and Lippi (1997), Chapter 13.

with the coefficient b between -1 and $-3/4$.

In the first stage we estimate ARIMA$(p, 1, q)$ models with $0 \leq p, q \leq 4$, $p+q \geq 1$ with the results reported in Tables 1 and 2. The Gaussian Pseudo Maximum Likelihood Estimator (PMLE) was employed with starting values given by the Durbin-Levinson algorithm (Brockwell and Davis 1987)[2].

In the second stage we estimate ARFIMA$(p, 1+e, q)$ models, where the reparameterization $d = 1 + e$ expresses that we use differenced data. The results are reported in Table 3 and 4 for the parameter e only together with the LM statistic to test the null hypothesis of a unit root

$$H_0 : e = 0,$$

a simple particular case of Robinson (1994) efficient test. Finally in the third stage we estimate the parameter d only by using two different semi-parametric estimators: the log-periodogram regression estimator, introduced by Geweke and Porter Hudak (1983) and formalized in Robinson (1995b), and the Gaussian local Whittle estimator introduced by Robinson (1995a). The results are contained in Tables 5 and 6. We report the estimated values for the bandwidth parameter m ranging from 60 to 90, the latter being half of the sample size $T = 180$. Furthermore we report the results obtained by using the raw data as well as detrending the data by a linear OLS trend. In fact, in order to be able to estimate the long memory parameter using the lower periodogram frequencies, we have to eliminate any deterministic component (Kunsch 1986)[3]. Alternatively, in order to eliminate the bias induced by non-increasing monotonic trends which might perfectly mimic (Bhattacharya, Gupta, and Waymire 1983) the presence of long-range dependence, we report estimates obtained trimming out the first $[T^{1/2}]$ periodogram ordinates, and thus based on the first $m - l + 1$ Fourier frequencies $(\lambda_l, \ldots, \lambda_m)$, with $\lambda_j = 2\pi j/T$ and $l = [T^{1/2}]$ instead of $l = 1$ (no trimming).

The last column of Tables 1 and 2 reports the AIC and BIC criteria, computed as in Hannan (1980) which select an ARIMA(1,1,4) and an ARIMA(4,1,4) for the time series of consumption excluding durables and an ARIMA(1,1,2) and ARIMA(3,1,1) for total consumption. Nevertheless the results of Tables 3 and 4 shows how the restriction $e = 0$, viz. a unit root,

[2]The optimization has been carried out with the GAUSS routine OPTMUM.

[3]In fact, the deterministic component would arise as a non-centrality parameter, exploding for any increasing deterministic trend, in the periodogram ordinates (Kunsch 1986).

is generally rejected both by the significance of the estimates of e as well as by the LM test statistics. The estimates of $d = 1 + e$, significantly different from unity, do not show a clear pattern: as p and q vary we find values both above and below unity. Thus we estimate the long memory parameter by semiparametric techniques which are not, by construction, affected by the short-run dynamics, which is left unspecified. In any case, in view of the theoretical results, the relevant parameter is given by the long memory parameter d only. We find estimates for d belonging to the open interval $(3/4, 1)$ and, when the bandwidth is big enough, significantly smaller than 1. Given that in this third stage we do not first-difference the data, we rely on Hurvich and Ray (1995) and related work to be formally allowed to use the log-periodogram and the Gaussian estimators in the non-stationary case.

This range mirrors the set of estimates found in the empirical literature for U.S. GDP (see the references in the Introduction), as one would expect. The estimates do not seem to be influenced by the detrending, suggesting non-appropriateness of the linear time trend. Nevertheless we might expect the presence of a (possibly nonlinear) deterministic component in the data, hence justifying the trimming based results.

Summing up, we have produced evidence in favour of a value of the parameter b between -1 and $-3/4$. For such values, as we have shown above, both the common and the idiosyncratic component contribute to explaining the variance of the aggregate variable (cf. Theorem 1 and 8). However, a more careful consideration of the data, with $T = 180$ and n representing the number of households in U.S.A., implies that we are in a fixed T and 'large' n situation, suggesting that the estimated aggregate variance is completely explained by the common component (cf. Theorem 12).

8 Conclusions

In a vast and growing macroeconomic literature variables corresponding to individual consumers or firms are modeled as the sum of a common and an idiosyncratic component. The idiosyncratic component may play a crucial role in micro modelling but disappears when aggregation over a huge number of agents is considered: see, among others, Bertola and Caballero (1990), Goodfriend (1992), Pischke (1995), Forni and Lippi (1997). This is an elementary fact under the often implicit assumption that there is an upper bound for the variances of individual variables. However, a model as simple as (1), with the autoregressive polynomial deriving from standard microeconomic theory, does not necessarily fulfil the condition of bounded variance.

We have studied aggregation of each of the two components of model (1) separately, and established rigorous asymptotic results both for the variance and the shape of the spectral density near the zero frequency. The features of the aggregate depend crucially on the parameter b, governing the distribution of the autoregressive coefficient in the vicinity of unity. As b takes on values smaller than zero we observe, firstly, the emergence of aggregate long memory, in spite of short memory micro variables; secondly, for b smaller than $-1/2$, an exploding variance for both common and idiosyncratic components, with the same rate of divergence: in other words, the relative importance of the idiosyncratic component in the aggregate does not vanish, as it does in all other cases.

Extension of our method and results to general ARMA structures can be easily obtained: irrespective of the particular parameterization, the necessary condition for the results summarized just above is that the support of at least one autoregressive root has unity as least upper bound, the MA structure being completely irrelevant. Thus our framework is applicable to a wide range of micro behaviours, including nonlinear micro reaction functions well approximated, in the proximity of an equilibrium, by an ARMA structure. On the other hand, there are many other nonlinear frameworks which are not covered.

An important simplification in our model is the assumption that the micro variables are driven by only one common shock. Ad argued in Forni and Lippi (1997), this assumption is untenable when macroeconomic time series are considered. However, our theoretical results apply with no difficulty to each common shock separately. Moreover, the task of the empirical exercise in Section 7 is limited to the order of integration of the common component in the consumption series (i.e. to the coefficient b in the probability density of the coefficients α), so that the effect of our simplification should not be dramatic.

Our result on the idiosyncratic component for $b < -1/2$, i.e. that the contribution of the idiosyncratic component to aggregate variance does not vanish as n tends to infinity, links this paper to some recent research on the relationship between individual, sectoral or regional shocks, and aggregate variation. We shall only mention here Bak, Chen, Scheinkman, and Woodford (1993), in which a non-linearity in the model is employed to generate aggregate variation from mutually orthogonal micro shocks; and a linear model in Forni and Lippi (1997, Section 1.5) in which a general autoregressive structure (each microvariable depending potentially on all the others lagged) is superimposed to mutually orthogonal individual shocks (like in an input-output model), so that the decline of the idiosyncratic component may be much slower than in the standard case. In this paper (Theorem 1) we find that aggregate fluctuations can be generated by purely idiosyncratic uncertainty even within the simplest and most common linear framework.

Table 1: **Real consumption (no durables): ARMA estimation results**

$(p,1,q)$	ϕ_1	ϕ_2	ϕ_3	ϕ_4	θ_1	θ_2	θ_3	θ_4	AIC/BIC
$(0,1,1)$	-	-	-	-	0.299	-	-	-	1360.49
					(0.071)				1409.82
$(0,1,2)$	-	-	-	-	0.298	0.041	-	-	1373.48
					(0.075)	(0.074)			1448.85
$(0,1,3)$	-	-	-	-	0.250	0.135	0252	-	1307.37
					(0.074)	(0.075)	(0.075)		1403.88
$(0,1,4)$	-	-	-	-	0.267	0.135	0.242	0.049	1319.89
					(0.074)	(0.075)	(0.075)	(0.074)	1442.79
$(1,1,0)$	0.302	-	-	-	-	-	-	-	1345.50
	(0.071)								1394.29
$(1,1,1)$	0.522	-	-	-	-0.401	-	-	-	1703.28
	(0.416)				(0.447)				1796.75
$(1,1,2)$	2.398	-	-	-	-2.274	-0.727	-	-	7058.64
	(5.898)				(6.274)	(1.917)			7579.74
$(1,1,3)$	0.412	-	-	-	- 0.289	-0.033	0.214	-	1672.50
	(0.274)				(0.269)	(0.082)	(0.076)		1828.23
$(1,1,4)$	-0.076	-	-	-	0.351	0.142	0.270	0.158	1193.66
	(0.526)				(0.519)	(0.147)	(0.093)	(0.134)	1328.25
$(2,1,0)$	0.273	0.086	-	-	-	-	-	-	1350.38
	(0.075)	(0.074)							1424.48
$(2,1,1)$	0.131	0.716	-	-	0.145	-	-	-	3277.36
	(0.077)	(0.061)			(0.109)				3519.31
$(2,1,2)$	0.138	0.805	-	-	0.122	-0.862	-	-	1439.13
	(0.061)	(0.054)			(0.066)	(0.061)			1573.13
$(2,1,3)$	0.658	-0.373	-	-	-0.551	0.396	0.267	-	1782.15
	(0.229)	(0.199)			(0.225)	(0.191)	(0.092)		1983.10
$(2,1,4)$	- 1.004	0.355	-	-	1.288	0.001	0.267	0.354	1825.26
	(6.989)	(1.802)			(6.961)	(0.443)	(0.155)	(1.966)	2067.56
$(3,1,0)$	0.251	0.025	0.228	-	-	-	-	-	1294.29
	(0.073)	(0.075)	(0.073)						1389.84
$(3,1,1)$	-0.087	-0.377	0.505	-	0.251	-	-	-	2841.04
	(0.105)	(0.061)	(0.080)		(0.119)				3105.59
$(3,1,2)$	-0.026	-0.445	0.546	-	0.219	0.678	-	-	1888.31
	(0.081)	(0.077)	(0.064)		(0.081)	(0.081)			2101.23
$(3,1,3)$	-0.007	-0.446	0.573	-	0.214	0.658	-0.327	-	1680.18
	(0.218)	(0.142)	(0.181)		(0.244)	(0.142)	(0.216)		1903.22
$(3,1,4)$	0.252	-0.596	0.552	-	-0.105	0.764	-0.292	-0.023	1827.46
	(0.309)	(0.145)	(0.227)		(0.319)	(0.206)	(0.288)	(0.142)	2107.24
$(4,1,0)$	0.256	0.025	0.233	-0.020	-	-	-	-	1308.30
	(0.074)	(0.075)	(0.075)	(0.074)					1430.13
$(4,1,1)$	-0.449	-0.689	0.217	0.546	0.848	-	-	-	2050.62
	(0.139)	(0.107)	(0.143)	(0.148)	(0.107)				2281.84
$(4,1,2)$	- 0.467	-0.707	0.223	0.561	0.800	1.019	-	-	1232.15
	(0.073)	(0.095)	(0.097)	(0.079)	(0.014)	(0.028)			1395.72
$(4,1,3)$	-0.469	-0.709	0.220	0.560	0.798	1.019	0.253	-	1196.73
	(0.242)	(0.103)	(0.184)	(0.099)	(0.244)	(0.154)	(0.191)		1379.95
$(4,1,4)$	-0.469	-0.707	0.227	0.569	0.795	1.014	0.249	-0.369	1141.11
	(1.426)	(1.299)	(0.921)	(0.662)	(1.404)	(1.510)	(0.807)	(0.410)	1339.45

The parameter ϕ_i denotes the autoregressive coefficient and θ_i the moving average coefficient at lag i. We used the Gaussian Pseudo Maximum Likelihood Estimator (PMLE) in the frequency domain with initial values given by Durbin-Levinson algorithm. Standard errors are in parentheses. The last two columns contains the AIC and BIC model selection criteria (multiplied by 10^6).

Table 2: **Total real consumption: ARMA estimation results**

$(p,1,q)$	ϕ_1	ϕ_2	ϕ_3	ϕ_4	θ_1	θ_2	θ_3	θ_4	AIC/BIC
$(0,1,1)$	-	-	-	-	0.141	-	-	-	4596.38
					(0.074)				4763.02
$(0,1,2)$	-	-	-	-	0.081	0.216	-	-	4455.53
					(0.073)	(0.073)			4700.02
$(0,1,3)$	-	-	-	-	0.131	0.232	0.185	-	4324.45
					(0.073)	(0.072)	(0.073)		4643.71
$(0,1,4)$	-	-	-	-	0.131	0.235	0.185	-0.004	4324.45
					(0.075)	(0.074)	(0.074)	(0.075)	4780.26
$(1,1,0)$	0.199	-	-	-	-	-	-	-	4543.58
	(0.073)								4708.31
$(1,1,1)$	1.293	-	-	-	-1.225	-	-	-	7629.08
	(0.477)				(0.422)				8047.71
$(1,1,2)$	0.796	-	-	-	-0.999	0.156	-	-	3692.92
	(2.453)				(2.461)	(0.544)			3965.55
$(1,1,3)$	0.199	-	-	-	- 0.337	0.234	0.152	-	4075.53
	(0.357)				(0.354)	(0.102)	(0.120)		4455.03
$(1,1,4)$	0.319	-	-	-	-0.507	0.225	0.110	-0.050	3802.23
	(32.51)				(32.52)	(6.119)	(5.354)	(5.337)	4230.96
$(2,1,0)$	0.151	0.222	-	-	-	-	-	-	4368.40
	(0.073)	(0.073)							4608.10
$(2,1,1)$	1.289	-0.893	-	-	-1.146	-	-	-	11885.3
	(0.034)	(0.033)			(0.039)				12762.7
$(2,1,2)$	1.461	-1.075	-	-	-1.414	1.408	-	-	6748.04
	(0.037)	(0.040)			(0.084)	(0.082)			7376.40
$(2,1,3)$	0.119	0.014	-	-	-0.187	0.149	0.195	-	4533.03
	(0.518)	(0.391)			(0.512)	(0.413)	(0.141)		5044.16
$(2,1,4)$	1.258	-0.207	-	-	-1.195	0.512	-0.136	-0.193	6100.81
	(0.319)	(0.326)			(0.291)	(0.332)	(0.174)	(0.127)	6910.67
$(3,1,0)$	0.119	0.199	0.145	-	-	-	-	-	4324.14
	(0.074)	(0.073)	(0.074)						4643.37
$(3,1,1)$	0.163	-0.001	0.026	-	- 0.292	-	-	-	2841.04
	(4.553)	(0.592)	(0.181)		(4.554)				4896.17
$(3,1,2)$	0.174	-0.046	0.035	-	-0.286	0.310	-	-	4072.22
	(0.702)	(0.311)	(0.211)		(0.699)	(0.351)			4531.39
$(3,1,3)$	0.171	-0.043	0.022	-	-0.286	0.305	0.074	-	4141.73
	(3.224)	(1.487)	(0.573)		(3.224)	(1.855)	(1.499)		4691.54
$(3,1,4)$	2.589	-0.492	-0.006	-	-2.760	1.109	-0.631	-0.473	14102.3
	(87.67)	(17.65)	(1.987)		(89.57)	(27.05)	(28.99)	(18.17)	16261.4
$(4,1,0)$	0.124	0.210	0.151	-0.043	-	-	-	-	4364.32
	(0.075)	(0.074)	(0.074)	(0.075)					4770.72
$(4,1,1)$	-2.719	0.582	-0.194	0.388	3.827	-	-	-	41001.4
	(8.001)	(1.428)	(0.526)	(0.802)	(10.12)				45624.5
$(4,1,2)$	-3.136	0.844	-0.043	0.512	3.876	-1.398	-	-	19963.9
	(21.56)	(7.007)	(0.392)	(2.975)	(26.37)	(11.52)			22614.1
$(4,1,3)$	-3.428	0.940	-0.362	0.574	3.893	-0.425	1.968	-	15667.8
	(51.91)	(15.03)	(6.221)	(7.103)	(57.06)	(13.74)	(27.12)		18066.5
$(4,1,4)$	-3.433	0.997	-0.390	0.626	3.892	-0.427	1.894	-0.539	15538.1
	(88.39)	(32.11)	(12.41)	(13.82)	(96.81)	(29.28)	(47.79)	(15.41)	18238.9

See description in Table 1 .

Table 3: **Real consumption (no durables): estimation of d and LM test for** $H_0 : e = 0$

$(p, 1+e, q)$	e	LM	$(p, 1+e, q)$	e	LM	$(p, 1+e, q)$	e	LM
$(0, 1+e, 0)$	0.240 (0.062)	16.49	$(1, 1+e, 0)$	-0.721 (0.085)	39.95	$(2, 1+e, 0)$	0.399 (0.110)	239.1
$(0, 1+e, 1)$	0.240 (0.103)	0.171	$(1, 1+e, 1)$	-0.698 (0.226)	0.036	$(2, 1+e, 1)$	0.363 (0.118)	298.8
$(0, 1+e, 2)$	0.280 (0.160)	0.453	$(1, 1+e, 2)$	0.481 (0.114)	0.472	$(2, 1+e, 2)$	0.357 (0.076)	11.54
$(0, 1+e, 3)$	0.091 (0.119)	18.92	$(1, 1+e, 3)$	-0.814 (0.225)	8.97	$(2, 1+e, 3)$	0.479 (0.174)	55.07
$(0, 1+e, 4)$	0.081 (0.161)	124.9	$(1, 1+e, 4)$	-0.619 (1.071)	23.36	$(2, 1+e, 4)$	0.831 (0.464)	4.15
$(3, 1+e, 0)$	0.040 (0.385)	108.3	$(4, 1+e, 0)$	-0.036 (0.473)	40.10			
$(3, 1+e, 1)$	0.636 (0.320)	171.7	$(4, 1+e, 1)$	0.227 (0.212)	115.6			
$(3, 1+e, 2)$	0.-0.749 (0.096)	208.5	$(4, 1+e, 2)$	-0.036 (0.132)	40.10			
$(3, 1+e, 3)$	0.389 (0.099)	5.596	$(4, 1+e, 3)$	-0.712 (0.304)	42.99			
$(3, 1+e, 4)$	-0.712 (0.301)	0.885	$(4, 1+e, 4)$	-0.643 (0.806)	2.851			

For each $ARFIMA(p, 1+e, q)$ we report the the estimate of the parameter e only and the LM test statistic for the null hypothesis $H_0 : e = 0$. We used the Gaussian Pseudo Maximum Likelihood Estimator (PMLE) in the frequency domain with initial values given by the Durbin-Levinson algorithm and a grid search. Standard errors are in parentheses.

Table 4: **Total real consumption : estimation of d and LM test for $H_0 : e = 0$**

$(p, 1+e, q)$	e	LM	$(p, 1+e, q)$	e	LM	$(p, 1+e, q)$	e	LM
$(0, 1+e, 0)$	0.191 (0.062)	7.33	$(1, 1+e, 0)$	-0.772 (0.087)	0.781	$(2, 1+e, 0)$	-0.124 (0.358)	47.10
$(0, 1+e, 1)$	0.302 (0.122)	0.113	$(1, 1+e, 1)$	0.298 (0.115)	10.11	$(2, 1+e, 1)$	-0.912 (0.576)	575.4
$(0, 1+e, 2)$	0.211 (0.121)	2.057	$(1, 1+e, 2)$	-0.588 (0.371)	0.858	$(2, 1+e, 2)$	-0.038 (0.099)	67.22
$(0, 1+e, 3)$	0.034 (0.117)	21.43	$(1, 1+e, 3)$	0.399 (0.139)	44.70	$(2, 1+e, 3)$	0.434 (0.142)	27.61
$(0, 1+e, 4)$	0.053 (0.168)	1704	$(1, 1+e, 4)$	-0.775 (0.584)	24.72	$(2, 1+e, 4)$	-0.667 (0.429)	1.687
$(3, 1+e, 0)$	-0.218 (0.489)	707.1	$(4, 1+e, 0)$	-1.018 (0.149)	888.7			
$(3, 1+e, 1)$	0.374 (0.191)	27.09	$(4, 1+e, 1)$	0.243 (0.186)	3.981			
$(3, 1+e, 2)$	0.-0.713 (0.311)	29.94	$(4, 1+e, 2)$	-0.597 (0.571)	27.33			
$(3, 1+e, 3)$	-0.523 (0.199)	1.068	$(4, 1+e, 3)$	-0.672 (0.385)	24.28			
$(3, 1+e, 4)$	-0.759 (0.636)	6.048	$(4, 1+e, 4)$	-0.346 (1.080)	12.43			

See description in Table 3 .

Table 5: **Real consumption (no durables): semiparametric estimation of d**

	Logperiodogram		Gaussian		
m	*raw*	*detrended*	*raw*	*detrended*	*n.obs*
No trimming : $l = 1$.					
60	0.965	0.782	0.984	0.886	60
	(0.083)		(0.064)		
65	0.949	0.817	0.969	0.909	65
	(0.079)		(0.062)		
70	0.937	0.827	0.958	0.916	70
	(0.077)		(0.059)		
75	0.921	0.868	0.944	0.935	75
	(0.074)		(0.057)		
80	0.911	0.839	0.935	0.906	80
	(0.071)		(0.056)		
85	0.895	0.852	0.921	0.914	85
	(0.069)		(0.054)		
90	0.879	0.851	0.908	0.915	90
	(0.067)		(0.053)		
Trimming : $l = [T^{1/2}]$					
60	0.875	0.896	0.899	0.871	47
	(0.083)		(0.064)		
65	0.846	0.970	0.871	0.939	52
	(0.079)		(0.062)		
70	0.828	0.972	0.853	0.953	57
	(0.077)		(0.059)		
75	0.801	0.1.053	0.826	1.101	62
	(0.074)		(0.057)		
80	0.791	0.961	0.817	0.919	67
	(0.071)		(0.056)		
85	0.766	0.976	0.792	0.939	72
	(0.069)		(0.054)		
90	0.743	0.960	0.769	0.935	77
	(0.067)		(0.053)		

The numer m represents the bandwith which may differ from the number of observations ($n.obs.$) due to the trimming. The second half of the table expresses the estimates obtained leaving out (trimming) the first $[T^{1/2}] - 1$ Fourier frequencies. We report the estimates of d using two semiparametric estimators : the Logperiodogram estimator (column 2 and 3) and the Gaussian estimator (column 4 and 5). Standard errors are in parentheses. For each case we obtain the result for the raw data as well as for the data detrended by a linear trend (OLS).

Table 6: **Total real consumption : semiparametric estimation of d**

	Logperiodogram		Gaussian		
m	*raw*	*detrended*	*raw*	*detrended*	*n.obs*
No trimming : $l = 1$.					
60	0.966	1.001	0.985	0.966	60
	(0.083)		(0.064)		
65	0.951	1.021	0.971	0.989	65
	(0.079)		(0.062)		
70	0.945	1.003	0.965	0.947	70
	(0.077)		(0.059)		
75	0.926	1.027	0.948	0.965	75
	(0.074)		(0.057)		
80	0.917	0.986	0.940	0.939	80
	(0.071)		(0.056)		
85	0.900	1.010	0.924	0.935	85
	(0.069)		(0.054)		
90	0.888	0.978	0.914	0.913	90
	(0.067)		(0.053)		
Trimming : $l = [T^{1/2}]$					
60	0.857	1.059	0.885	0.968	47
	(0.083)		(0.064)		
65	0.833	1.105	0.861	1.026	52
	(0.079)		(0.062)		
70	0.835	1.028	0.859	0.919	57
	(0.077)		(0.059)		
75	0.801	0.1.084	0.824	0.965	62
	(0.074)		(0.057)		
80	0.792	0.974	0.816	0.907	67
	(0.071)		(0.056)		
85	0.766	1.032	0.787	0.900	72
	(0.069)		(0.054)		
90	0.752	0.955	0.774	0.856	77
	(0.067)		(0.053)		

See description in Table 5 .

9 Appendix A

In this appendix C will denote an arbitrary constant, not necessarily the same, the symbol $\sim$ will denote asymptotic equivalence and $P(A)$ the probability of any event A.

Lemma 1 *Assume that the random variables β_i, $i = 1, \infty$, are i.i.d. with support $[0,\ 1]$ and density f. Assume that for $u \to 0^+$ there exist $C > 0$ and $b > -1$ such that $f(u) \sim Cu^b$. Let k be such that $2 > (b+1)/k > 0$. Let $\{a_n\}$ be a sequence of positive real numbers such that as $n \to \infty$*

$$a_n \to \infty, \ \sum_{n=1}^{\infty} a_n^{-(b+1)/k} < \infty.$$

Then as $n \to \infty$:
(i) If $\frac{b+1}{k} \neq 1$

$$\frac{1}{a_n}\sum_{i=1}^{n}\frac{1}{\beta_i^k} - \frac{1}{a_n}\sum_{i=1}^{n}\frac{1}{1-\frac{(b+1)}{k}}\left(a_i^{-(b+1)/k+1} - 1\right) \to 0 \ \ a.s.$$

(ii) If $\frac{b+1}{k} = 1$

$$\frac{1}{a_n}\sum_{i=1}^{n}\frac{1}{\beta_i^k} - \frac{1}{a_n}\sum_{i=1}^{n}\log a_i \to 0, \ \ a.s..$$

Proof: Obviously we cannot use a standard LLN because we allow for an unbounded first moment for $1/\beta_i^k$ when $(b+1)/k \leq 1$. Note that restricting to the case $2 > (b+1)/k$ is innocuous because if $(b+1)/k \geq 2$, then $1/\beta_i^k$ has a finite first moment so that standard LLN applies.

Let us set $y_i = 1/\beta_i^k$. We will use Loève (1977, Ch.5 Theorem 16.4 A) setting, in the notation there adopted, $c_n = c = c' = 1$ and $g_n(x) = 1(x > 1) + x^2\, 1(0 \leq x \leq 1)$, where $1(A)$ denotes the indicator function of the event A.

To prove (i) let us denote by f_y the probability density function of the variables y_i. We have

$$f_y(u) = \frac{1}{k} f(u^{-1/k})u^{-1/k-1}, \ \ 1 \leq u < \infty,$$

with $f_y(u) \sim Cu^{-(b+1)/k-1}$ as $u \to \infty$ for a $0 < C < \infty$. Therefore, as $n \to \infty$

$$P(y_n \geq a_n x) \sim C \int_{xa_n}^{\infty} t^{-(b+1)/k-1} dt \sim C(xa_n)^{-(b+1)/k},$$

since $(b+1)/k > 0$. Moreover, since $(b+1)/k < 2$:

$$\int_0^1 x^{-(b+1)/k+1} dx < \infty,$$

so that the above mentioned theorem (Loève (1977)) applies and

$$E(y_n^{a_n}) \sim C \int_1^{a_n} t^{-(b+1)/k} dt = C \frac{1}{-(b+1)/k+1} \left(a_n^{-(b+1)/k+1} - 1\right)$$

as $n \to \infty$.
In case (*ii*) we have $E(y_n^{a_n}) \sim C \int_1^{a_n} y^{-1} dy = C \log a_n$, as $n \to \infty$. QED
Remark. Case (*ii*) includes the uniform distribution on [0, 1] corresponding to $b = 0$.

The following Corollary is an immediate implication of Lemma 1, providing a link to the assumption and results of the paper.

Corollary 1 *Let $\{f_n\}$ be an SFG sequence (for the definition of a 'sufficiently fast growing' sequence see Section 4). Assume that the random variables α_i, $i = 1, \infty$, are i.i.d. and that their probability density $B(\cdot\,; b)$ fulfils Assumption I (see Section 3). Then for A a.e. in $\mathcal{G}$, as $n \to \infty$:*

(1) If $(b+1)/k \neq 1$,

$$\sum_{i=1}^{n} \frac{1}{(1-\alpha_i)^k} \leq \frac{k}{b-k+1} \left[n - \frac{b+1}{k} f_n^{\frac{k-b-1}{b+1}} \left(n^{\frac{k}{b+1}} - 1\right)\right] + o\left((n\, f_n)^{\frac{k}{b+1}}\right),$$

$$\sum_{i=1}^{n} \frac{1}{(1-\alpha_i)^k} \geq -\frac{b+1}{b-k+1} \left[n^{\frac{k}{b+1}} - 1\right] + o\left(n^{\frac{k}{b+1}}\right).$$

(2) If $(b+1)/k = 1$,

$$\sum_{i=1}^{n} \frac{1}{(1-\alpha_i)^k} = \sum_{i=1}^{n} \log f_i + n(\log n - 1) + o(n\, f_n).$$

It must be pointed out that all the $o(\cdot)$'s depend on the sequence A.

Proof. Apply Lemma 1 setting $\beta_i = 1 - \alpha_i$ and

$$a_n = (nf_n)^{\frac{k}{b+1}}.$$

10 Appendix B

Proof of Theorem 1. We have

$$\frac{1}{2}\sum_{i=1}^{n}\frac{1}{1-\alpha_i}\leq\sum_{i=1}^{n}\frac{1}{1-\alpha_i^2}\leq\sum_{i=1}^{n}\frac{1}{1-\alpha_i},$$

so that Corollary 1 can be applied with $k=1$. If $b>0$ we are in case (1). The term on the RHS of the first inequality is dominated, for $n\to\infty$, by n/b, i.e.

$$\sum_{i=1}^{n}\frac{1}{1-\alpha_i}\leq\frac{1}{b}n+o(n)=\frac{1}{b}n\left(1+\frac{bo(n)}{n}\right).$$

As remarked in Corollary 1, $o(n)$ depends on the sequence A. Dividing by n^2, statement (i) follows with C depending on A. Case $b=0$ requires the application of Corollary 1, statement (2). When $b<0$ we apply again Corollary 1, statement (1). In this case however the term n/b is dominated for n tending to infinity. We end up with

$$\begin{aligned}\frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{1-\alpha_i} &\leq -\frac{b+1}{b}n^{-\frac{2b+1}{b+1}}f_n^{-\frac{b}{b+1}}+o\left(n^{-\frac{2b+1}{b+1}}\right)\\ \frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{1-\alpha_i} &\geq -\frac{b+1}{b}n^{-\frac{2b+1}{b+1}}+o\left(n^{-\frac{2b+1}{b+1}}\right)\end{aligned}$$

The conclusion follows. QED

Proof of Theorem 2. Since

$$s_{A_n}^E(0)=\frac{\sigma_\epsilon^2}{2\pi n^2}\sum_{i=1}^{n}\frac{1}{(1-\alpha_i)^2},$$

we can apply Corollary 1 with $k=2$. The proof follows the lines of the proof of Theorem 1. QED

Proof of Theorem 3. If $\tilde{\lambda}>0$, the real random variable $1/|1-\alpha e^{-i\tilde{\lambda}}|^2$, has a finite expectation. Therefore

$$\frac{1}{n}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{|1-\alpha e^{-i\tilde{\lambda}}|^2}\right)$$

tends to zero for A a.e. in $\mathcal{G}$. The function $s^E_{A_n}(\lambda)$ is decreasing, being the sum of decreasing functions, and tends therefore to zero pointwise in $[\tilde{\lambda}, \pi]$ for A a.e. in $\mathcal{G}$. Since $(0, \pi]$ is the union of a countable sequence of sets $[\lambda_s, \pi]$, the result follows. QED

Proof of Theorem 4. We remind that

$$\operatorname{var}(\eta_{A_n,t}) = \exp \frac{1}{\pi} \int_0^\pi \left[\log s^E_{A_n}(\lambda)\right] d\lambda.$$

Since $s^E_{A_n}(\lambda)$ is decreasing:

$$s^E_{A_n}(\lambda) \leq g_{\tilde{\lambda}}(\lambda)$$

where

$$g_{\tilde{\lambda}}(\lambda) = \begin{cases} s^E_{A_n}(0) \text{ if } 0 \leq \lambda < \tilde{\lambda} \\ s^E_{A_n}(\tilde{\lambda}) \text{ if } \lambda > \tilde{\lambda} \end{cases}$$

For A a.e. in $\mathcal{G}$ there exists $D_{A,\tilde{\lambda}}$ such that $s^E_{A_n}(\tilde{\lambda}) \leq D_{A,\tilde{\lambda}} n^{-1}$. We have

$$\operatorname{var}(\eta_{A_n,t}) \leq \tilde{\lambda} \log s^E_{A_n}(0) - (\pi - \tilde{\lambda})(\log n - \log D_{A,\tilde{\lambda}}).$$

For $b \geq 1$ the first term on the left hand side tends to $-\infty$, so that the whole sum tend to $-\infty$. For $b < 1$, applying Theorem 2, (iii), we have

$$\begin{aligned} \operatorname{var}(\eta_{A_n,t}) \leq \;& -\tilde{\lambda}\frac{2b}{b+1} \log n - (\pi - \tilde{\lambda}) \log n - (\pi - \tilde{\lambda}) \log D_{A,\tilde{\lambda}} \\ & + \tilde{\lambda} \log C_{A,f} + \tilde{\lambda}\frac{1-b}{b+1} \log f_n. \end{aligned}$$

Taking $\{f_n\}$ slow enough and $\tilde{\lambda}$ such that

$$-\tilde{\lambda}\frac{2b}{b+1} - (\pi - \tilde{\lambda}) < 0,$$

we get the result. QED

Proof of Theorem 5. Let b assume any value greater than -1. The spectral density of $(1 - L)E_{A_n,t}$, i.e. $|1 - z|^2 s^E_{A_n}(z)$, is zero at $z = 1$ and is strictly increasing in $[0\ \pi]$. Since $s^E_{A_n}(\pi)$ tends to zero (Theorem 3) for A a.e. in $\mathcal{G}$, then $(1 - L)E_{A_n,t}$ converges in variance to zero for A a.e. in $\mathcal{G}$. QED

Proof of Theorem 6. Put $\mu_{A_n,k} = (1/n)\sum_{i=1}^n \alpha_i^k$, and

$$\begin{aligned}
U_{A_n,s,t} &= \left(1 + \mu_{A_n,1}L + \mu_{A_n,2}L^2 + \cdots \mu_{A_n,s}L^{s-1}\right) u_t \\
U_{b,s,t} &= \left(1 + \mu_{b,1}L + \mu_{b,2}L^2 + \cdots \mu_{b,s}L^{s-1}\right) u_t \\
\hat{U}_{A_n,s,t} &= \left(\mu_{A_n,s}L^s + \mu_{A_n,s+1}L^{s+1} + \cdots\right) u_t \\
\hat{U}_{b,s,t} &= \left(\mu_{b,s}L^s + \mu_{b,s+1}L^{s+1} + \cdots\right) u_t
\end{aligned}$$

We have:

$$\begin{aligned}
\mathrm{var}(U_{A_n,t} - U_{b,t}) &= \mathrm{var}(U_{A_n,s,t} - U_{b,s,t}) + \mathrm{var}(\hat{U}_{A_n,s,t} - \hat{U}_{b,s,t}) \\
&\le \mathrm{var}(U_{A_n,s,t} - U_{b,s,t}) + \left(\sqrt{\mathrm{var}(\hat{U}_{A_n,s,t})} + \sqrt{\mathrm{var}(\hat{U}_{b,s,t})}\right)^2
\end{aligned}$$

We have

$$\mathrm{var}\left(\hat{U}_{A_n,s,t}\right) = \frac{\sigma_u^2}{n^2}\sum_{i,j=1}^n \frac{\alpha_i^{2s}\alpha_j^{2s}}{1-\alpha_i\alpha_j}.$$

Since

$$(1-\alpha_i\alpha_j)^2 \ge (1-\alpha_i^2)(1-\alpha_j^2), \tag{12}$$

then

$$\mathrm{var}\left(\hat{U}_{A_n,s,t}\right) \le \frac{\sigma_u^2}{n^2}\left[\sum_{i=1}^n \frac{\alpha_i^{2s}}{(1-\alpha_i^2)^{1/2}}\right]^2 \le \left[\frac{\sigma_u}{n}\sum_{i=1}^n \frac{\alpha_i^{2s}}{(1-\alpha_i)^{1/2}}\right]^2.$$

For A a.e. in $\mathcal{G}$ the expression within the square brackets in the last expression tends to a limit $\tilde{\mu}_{b,s}$, which is asymptotically equivalent, for $s \to \infty$, to $K\mu_{b-1/2,2s}$ (remind that $b - 1/2 > -1$ by assumption). Given a positive real τ take s such that $\tilde{\mu}_{b,s} < \tau/2$ and $\mathrm{var}\left(\hat{U}_{b,s,t}\right) < \tau$. For A a.e. in $\mathcal{G}$ there exists an $n_{A,\tau}$ such that for $n > n_{A,\tau}$

$$\mathrm{var}\left(U_{A_n,s,t} - U_{b,s,t}\right) < \tau, \quad \mathrm{var}\left(\hat{U}_{A_n,s,t}\right) < \tau,$$

so that

$$\mathrm{var}\left(U_{A_n,t} - U_{b,t}\right) < \tau + (\sqrt{\tau} + \sqrt{\tau})^2.$$

Since τ is arbitrary the theorem is proved. QED

Proof of Theorem 7. In general, if $z_n \in L_2^0$ and $z_n \to z$, then $s_n \to s$ in first-mean, where s_n and s denote the spectral density of z_n and z respectively. For, denote by $\hat{s}_n$ the spectral density of $z_n - z$, by $\check{s}_n$ the $(1,2)$ entry of the joint spectral density matrix of z_n and z, and by $\tilde{s}_n$ the $(1,2)$ entry of the joint spectral density matrix of z and $z_n - z$. We have

$$|s_n - s| = |\hat{s}_n - 2s + 2\Re\check{s}_n| = |\hat{s}_n + 2\Re\tilde{s}_n| \leq \hat{s}_n + |2\Re\tilde{s}_n|.$$

The integral over the interval $[-\pi,\ \pi]$ of $\hat{s}_n$, being the variance of $z_n - z$, tends to zero. Moreover

$$\left|\int_{-\pi}^{\pi} \Re\tilde{s}_n(\lambda)d\lambda\right| \leq \sqrt{\int_{-\pi}^{\pi} \hat{s}_n(\lambda)d\lambda}\sqrt{\int_{-\pi}^{\pi} s(\lambda)d\lambda}.$$

QED

Proof of Theorem 8. (i) To prove (ii) and (iii) we observe that (see (12):

$$\sum_{i=1}^{n} \frac{1}{1-\alpha_i^2} \leq \sum_{i,j=1}^{n} \frac{1}{1-\alpha_i\alpha_j} \leq \left[\sum_{i=1}^{n} \frac{1}{(1-\alpha_i^2)^{1/2}}\right]^2$$

and apply Corollary 1 for $k = 1/2$ and $k = 1$. QED

Proof of Theorem 9. For $b > 0$ the random variable $1/(1-\alpha)$ has finite expectation

$$\int_G \frac{1}{1-\alpha} B(\alpha; b)d\alpha.$$

We have:

$$\left|\int_G \frac{1}{1-\alpha} B(\alpha; b)d\alpha - \sum_{s=0}^{k} \mu_{b,s}\right| = \int_G \frac{\alpha^{k+1}}{1-\alpha} B(\alpha; b)d\alpha \leq M\mu_{b,k+1}.$$

For (ii) and (iii) apply Corollary 1 with $k = 1$. QED

Proof of Theorem 10. For $\lambda \neq 0$ the modulus of $(1-\alpha e^{-i\lambda})^{-1}$ is bounded, so that the integral

$$\int_G \frac{1}{1-\alpha e^{-i\lambda}} B(\alpha; b)d\alpha \tag{13}$$

is finite. On the other hand, like in the proof of Theorem 9:

$$\left|\int_G \frac{1}{1-\alpha e^{-i\lambda}} B(\alpha;b)d\alpha - \sum_{s=1}^{k} \mu_{b,s} e^{-i\lambda s}\right| \leq \left|\int_G \frac{\alpha^{k+1} e^{-i\lambda(k+1)}}{1-\alpha e^{-i\lambda}} B(\alpha;b)d\alpha\right| \leq M_\lambda \mu_{b,k+1}.$$

Thus, for $\lambda \neq 0$ and $b > -1$ we have that $s_b^U(\lambda)$ is equal to the squared modulus of (13), this proving (i). From (i) one obtains immediately that given $\lambda \neq 0$, for A a.e. in $\mathcal{G}$, $s_{A_n}^U(\lambda)$ converges to $s_b^U(\lambda)$. As an easy consequence we have that for A a.e. in $\mathcal{G}$, $s_{A_n}^U(\lambda)$ converges to $s_b^U(\lambda)$ for λ belonging to a countable dense subset of $[0\ \pi]$, the rational frequencies for example. We only have to show that (13) is continuous for $\lambda \neq 0$. But, for $\lambda_j \to \lambda$,

$$\left|\int_G \left[\frac{1}{1-\alpha e^{-i\lambda}} - \frac{1}{1-\alpha e^{-i\lambda_j}}\right] B(\alpha;b)d\alpha\right|$$
$$= \left|\int_G \frac{\alpha(e^{-i\lambda} - e^{-i\lambda_j})}{(1-\alpha e^{-i\lambda})(1-\alpha e^{-i\lambda_j})} B(\alpha;b)d\alpha\right| \leq M_\lambda |e^{-i\lambda} - e^{-i\lambda_j}| \mu_{b,1}.$$

QED

Proof of Theorem 11. (*i*) By construction the sequence $\{\mu_{b,k}\}$ is monotone. Moreover $\mu_{b,k} \sim C\Gamma(b+1)\, k^{-(b+1)}$, for $C > 0$, as $k \to \infty$. The spectral density $s_b^U(\lambda)$ can be written as

$$s_b^U(\lambda) = \frac{\sigma_u^2}{2\pi}\left[\left(\sum_{k=0}^{\infty} \mu_{b,k} cos(k\,\lambda)\right)^2 + \left(\sum_{k=0}^{\infty} \mu_{b,k} sin(k\,\lambda)\right)^2\right],$$

so that no convolutions of the $\mu_{b,k}$'s, not defined when $b \leq -1/2$, are involved. The result follows by applying Yong (1974): Theorem I-9 and Theorem I-10 when $b < 0$, Theorem I-18 and Theorem I-21 when $b = 0$, Theorem I-31 and Theorem I-32 when $b > 0$.

(*ii*) Taking the first difference $(1-L)U_{A_n,t}$ and adapting Theorem 6, we have a stationary limit, call it $\tilde{U}_{b,t}$, for A a.e. in $\mathcal{G}$, whose Wold representation and spectral density are, respectively,

$$\begin{aligned} \tilde{U}_{b,t} &= u_t + (\mu_{b,1} - 1)u_{t-1} + (\mu_{b,2} - \mu_{b,1})u_{t-2} + \cdots \\ s_b^{\tilde{U}}(\lambda) &= |1 - e^{-i\lambda}|^2 s_b^U(\lambda). \end{aligned}$$

The order of integration of $\tilde{U}_{b,t}$ is $-(1+b)$, so that the non-stationary variable produced by integrating $\tilde{U}_{b,t}$ has order of integration $-b$. QED

Proof of Theorem 12. (i) Starting with

$$U_{A_n,T} = \sum_{k=0}^{T-1} \left(\frac{1}{n}\sum_{i=1}^{n}\alpha_i^k\right) u_{T-k} + \sum_{k=T}^{\infty} \left(\frac{1}{n}\sum_{i=1}^{n}\alpha_i^k\right) u_{T-k},$$

and evaluating the conditional variance, the result follows from the fact that as $n \to \infty$

$$\frac{1}{n}\sum_{i=1}^{n}\alpha_i^k - \mu_{b,k} \sim \frac{2^{1/2}\, var^{1/2}(\alpha_i^k)}{n^{1/2}} \left(\log\log(n\, var(\alpha_i^k))\right)^{1/2} a.s., \qquad (14)$$

by a version of the law of iterated logarithm (Stout 1974, Corollary 5.2.1)[4], and from the asymptotic equivalences, for $k \to \infty$:

$$\mu_{b,k} \sim C\Gamma(b+1)k^{-(b+1)}, \qquad \mathrm{var}(\alpha_i^k) \sim \mu_{b,2k}.$$

For, $E_{A_n,T}$ one obtains

$$var(E_{A_n,T} \mid \mathcal{F}_0) = \frac{\sigma_\epsilon^2}{n}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1-\alpha_i^{2T}}{1-\alpha_i^2}\right) \sim \frac{c\,\sigma_\epsilon^2}{n}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1-\alpha_i^{T}}{1-\alpha_i}\right),$$

for a constant $c > 0$, as $n \to \infty$. By the mean value theorem, for some $0 < \tilde{T} < T$

$$\frac{1-\alpha^T}{1-\alpha} = -\alpha^{\tilde{T}}\frac{log\,\alpha}{1-\alpha}T.$$

Using $1 - 1/a \leq \log\, a \leq a - 1$ for any $a > 0$ one obtains

$$\frac{T}{n}\sum_{i=1}^{n}\alpha_i^{\tilde{T}-1} \leq \frac{1}{n}\sum_{i=1}^{n}\frac{1-\alpha_i^T}{1-\alpha_i} \leq \frac{T}{n}\sum_{i=1}^{n}\alpha_i^{\tilde{T}}.$$

Then the result follows by using (14) and summing terms applying Yong (1974, Lemma I-11 (1-32') and Lemma I-16 (1-41')).

(ii) For the common component, the proof follows (i) and uses

$$\sum_{k=1}^{T}\frac{1}{k} \sim log\,(T),$$

as $T \to \infty$. For the idiosyncratic component just set $b = -1/2$ in (i). QED

[4]We can assume that $2^{-(b+1)}\, n\,\Gamma(b+1)\,T^{-(b+1)} > 1$, so that the double logarithm is always well defined asymptotically. This condition for instance is satisfied for $n \sim c\,T^{b+1}$, provided that $c > 2^{b+1}/\Gamma(b+1)$.

References

AL-NAJJAR, N. (1995): "Decomposition and characterization of risk with a continuum of random variables," *Econometrica*, 63/5, 1195–1224.

BAK, P., K. CHEN, J. SCHEINKMAN, AND M. WOODFORD (1993): "Aggregate fluctuations from independent sectoral shocks: self organized criticality in a model of production and inventory dynamics," *Ricerche Economiche*, 47/1, 3–30.

BERTOLA, G., AND R. J. CABALLERO (1990): "Kinked Adjustment Costs and Aggregate Dynamics," in *NBER Macroeconomics Annual*, ed. by O. J. Blanchard and S. Fisher, pp. 237–95. Cambridge, MA: MIT Press.

BHATTACHARYA, R., V. GUPTA, AND E. WAYMIRE (1983): "The Hurst effect under trends," *Journal of Applied Probability*, 20, 649–662.

BROCKWELL, P., AND R. DAVIS (1987): *Time series: theory and methods.* Springer Verlag.

DIEBOLD, F., AND G. RUDEBUSCH (1989): "Long Memory and Persistence in aggregate Output," *Journal of Monetary Economics*, 24, 189–209.

FORNI, M., AND M. LIPPI (1997): *Aggregation and the microfoundations of dynamic macroeconomics.* Oxford: Oxford University Press.

GEWEKE, J., AND S. PORTER HUDAK (1983): "The estimation and application of Long Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221–238.

GIL-ALAÑA, L., AND P. ROBINSON (1997): "Testing of unit root and other nonstationary hypothesis in macroeconomic time series," Econometrics Discussion Paper 317, STICERD-LSE.

GONÇALVES, E., AND C. GOURIEROUX (1988): "Agregation de processus autoregressifs d'ordre 1," *Annales d'Economie et Statistique*, 12, 127–149.

GOODFRIEND, M. (1992): "Information-aggregation bias," *American Economic Review*, 82, 508–19.

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

HALL, R. E. (1992): "Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence," *Journal of Political Economy*, 96, 971–87.

HANNAN, E. (1980): "The estimation of the order of an ARMA process," *The Annals of Statistics*, 8-5, 1071–1081.

HANSEN, L. P., AND T. J. SARGENT (1980): "Formulating and estimating dynamic linear rational expectations models," *Journal of Economic Dynamics and Control*, 2, 7–46.

HANSEN, L. P., AND T. J. SARGENT (1991): *Rational expectations econometrics*. London: Westview.

HAUBRICH, J., AND A. LO (1989): "The sources and nature of long-term memory in the business cycle," discussion paper 2951, NBER.

HURVICH, C., AND B. RAY (1995): "Estimation of the memory parameter for non-stationary or non-invertible fractionally integrated processes," *Journal of Time Series Analysis*, 16/1, 17–41.

KUNSCH, H. (1986): "Discrimination between monotonic trends and long-range dependence," *Journal of Applied Probability*, 23, 1025–1030.

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

LOÈVE, M. (1977): *Probability Theory I*. Springer Verlag - Berlin.

MICHELACCI, C., AND P. ZAFFARONI (1997): "Beta convergence," Econometrics Discussion Paper 332, STICERD-LSE.

NICKELL, S. (1985): "Error correction, partial adjustment and all that: an expository note," *Oxford Bulletin of Economics and Statistics*, 47, 119–29.

PISCHKE, J.-S. (1995): "Individual Income, Incomplete Information, and Aggregate Consumption," *Econometrica*, 63, 805–40.

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

ROBINSON, P. M. (1995a): "Gaussian semiparametric estimation of long range dependence," *Annals of Statistics*, 23, 1630–1661.

ROBINSON, P. M. (1995b): "Log-periodogram regression of time series with long-range dependence," *Annals of Statistics*, 23, 1048–1072.

ROZANOV, Y. (1963): *Stationary random processes*. Holden-Day, San Francisco.

SARGENT, T. J. (1978): "Estimation of dynamic labor demand under rational expectations," *Journal of Political Economy*, 86, 1009–44.

SOWELL, F. (1992): "Modeling long-run behavior with the fractional ARIMA model," *Journal of Monetary Economics*, 29, 277–302.

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

UHLIG, H. (1996): "A law of large numbers for large economies," *Economic Theory*, 8, 41–50.

YONG, C. H. (1974): *Asymptotic Behaviour of Trigonometric Series*. Chinese University of Hong Kong.