Journal of Econometrics 199 (2017) 1–11

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Long memory, fractional integration, and cross-sectional aggregation

Niels Haldrup *, J. Eduardo Vera Valdés

Department of Economics and Business Economics, Aarhus University and CREATES, Fuglesangs Allé 4, 8210 Aarhus V, Denmark

ARTICLE INFO

Article history:
Received 30 January 2016
Received in revised form
23 December 2016
Accepted 6 March 2017
Available online 23 March 2017

JEL classification:
C2
C22

Keywords:
Long memory
Fractional integration
Aggregation

ABSTRACT

It is commonly argued that observed long memory in time series variables can result from cross-sectional aggregation of dynamic heterogeneous micro units. In this paper we demonstrate that the aggregation argument is consistent with a range of different long memory definitions. A simulation study shows that the cross-section dimension needs to be rather large to reflect the theoretical memory when using commonly used methods to estimate the memory parameter, especially when the theoretical memory is not too high. We show that the aggregated process will converge to a generalized fractional process in the limit. The coefficients of the moving average representation of the series decay hyperbolically but they differ from the coefficients arising from inversion of the fractional difference filter. It appears that the fractionally differenced series will have an autocorrelation function that still exhibits hyperbolic decay, but at a rate that ensures summability. The fractionally differenced series is thus $I(0)$ but standard *ARFIMA* modeling is invalid when the long memory is caused by aggregation. It is shown that standard methods for estimating and selecting *ARFIMA* specifications fail in properly fitting the dynamics of the series.

1. Introduction

Without specifically discussing long memory, the study of this concept in econometrics goes back to Granger (1966) in his article about the spectral shape near the origin for economic time series variables. He found that *long-term fluctuations, if decomposed into frequency components, are such that the amplitudes of the components decrease smoothly with decreasing period* (Granger, 1966, p. 155). This certainly applies for non-stationary $I(1)$ processes and more generally for the class of fractionally integrated processes as demonstrated by Granger and Joyeux (1980). Such processes have long lasting autocorrelations that decay hyperbolically instead of the standard geometric decay characterizing *ARMA* processes.

This kind of behavior has led to several definitions of long memory. In this paper we consider five definitions of long memory.

Definition. Let x_t be a stationary time series with autocovariance function $\gamma_x(k)$ and spectral density function $f_x(\lambda)$, and let $d \in (0, 1/2)$, then x_t has long memory

(i) in the **covariance sense** if $\gamma_x(k) \approx C_x k^{2d-1}$ as $k \to \infty$ with C_x a constant,

(ii) in the **spectral sense** if $f_x(\lambda) \approx C_f \lambda^{-2d}$ as $\lambda \to 0$ with C_f a constant,

(iii) in the **rate of the partial sum sense** if $\mathrm{Var}(\sum_{t=1}^{T} x_t) = O_p(T^{1+2d})$,

(iv) in the **self-similar sense** if $m^{1-2d}\mathrm{Cov}(x_t^{(m)}, x_{t+k}^{(m)}) \approx C_m k^{2d-1}$ as $k, m \to \infty$ where $x_t^{(m)} = \frac{1}{m}(x_{tm-m+1} + \cdots + x_{tm})$ with $m \in \mathbb{N}$, $m/k \to 0$, and C_m is a constant,

(v) in the **weak convergence sense** if $X_n(\xi) = \sigma_n^{-1} \sum_{t=1}^{\lfloor n\xi \rfloor} x_t \Rightarrow B_H(\xi)$, where $\sigma_n^2 = \mathrm{E}[(\sum_{t=1}^{n} x_t)^2]$, $\xi \in [0, 1]$, $B_H(\xi)$ is a fractional Brownian motion, $H = d + 1/2$, and $\Rightarrow$ denotes weak convergence on $D[0, 1]$, the space of real-valued functions that are continuous from the right with finite left limits.

Above, $g(x) \approx h(x)$ as $x \to x_0$ means that $g(x)/h(x)$ converges to 1 as x tends to x_0, $O_p(\cdot)$ denotes order in probability, and $\lfloor \cdot \rfloor$ denotes the integer value of its argument.

Definition (i) is concerned with the behavior of the autocorrelation function for long lags and was one of the motivations behind the *ARFIMA* class of models due to Adenstedt (1974), Granger and Joyeux (1980), and Hosking (1981). Basically, they extended the *ARMA* model to account for fractional differencing. That is, for a stationary fractional process

$$A(L)(1-L)^d x_t = B(L)\epsilon_t, \qquad (1)$$

where ϵ_t is a white noise process, $d \in (-1/2, 1/2)$, and $A(L)$, $B(L)$ are polynomials in the lag operator with no common roots, all

* Corresponding author.
E-mail addresses: nhaldrup@creates.au.dk (N. Haldrup), evera@creates.au.dk (J.E. Vera Valdés).

http://dx.doi.org/10.1016/j.jeconom.2017.03.001
0304-4076/

outside the unit circle. They used the standard binomial expansion to decompose $(1-L)^d$ in a series with coefficients $\pi_j = \Gamma(j+d)/(\Gamma(d)\Gamma(j+1))$ for $j \in \mathbb{N}$. Using Stirling's approximation it can be shown that these coefficients decay at a hyperbolic rate ($\pi_j \approx j^{d-1}$ as $j \to \infty$), which translates to slowly decaying autocorrelations.

Definition (ii) is the feature considered by Granger (1966) in his study of the typical spectral shape for economic variables. The behavior of the spectrum near the origin is also used in the construction of one of the most popular estimators for long memory due to Geweke and Porter Hudak (1983) who proposed an estimation procedure based on semiparametric log periodogram regression near the zero frequency.

Diebold and Inoue (2001) based their work on spurious long memory on definition (iii). They showed that structural breaks or regime switching schemes can be confused with long memory of the fractional type by focusing on the stochastic order of the variance of partial sums. Their paper demonstrates that certain stochastic processes are long memory by one definition but not necessarily by other definitions.

Definitions (iv) and (v) are largely based on the work of Mandelbrot and Van Ness (1968) for fractals. They defined the self-similarity condition and showed that the fractional Brownian motion in particular has this property. Basically, self-similarity implies that the degree of memory is constant for different levels of temporal aggregation. Weak convergence to a fractional Brownian motion of an appropriately scaled partial sum is important for many parametric long memory models, but the class of processes is broader than often being considered as we shall later see.

It is well known that *ARFIMA* processes are long memory by definitions (i) through (ii), and an analogous derivation as in the proof of Theorem 1 shows that it is also long memory in the self-similar sense, definition (iv). Moreover, a scaled partial sum of an *ARFIMA* process converges to fractional Brownian motion, see for instance Davydov (1970) and Davidson and de Jong (2000). Thus, in the time series literature the *ARFIMA* model has become the canonical specification for modeling long memory.

Even though the *ARFIMA* model seems to be an appropriate specification to study long memory, the source underlying its dynamic features is still not clear. Physical (turbulence, see for instance Kolmogorov (1941)), as well as psychological reasons (Pearson (1902) *personal equation*), have been used to explain the presence of long memory. More recently, Parke (1999) proposed the error-duration model which relies on a decomposition of the time series into the sum of a sequence of shocks of stochastic magnitude and duration. He shows that if only a small proportion of the errors survive for large periods of time then the resulting series shows long memory in the covariance sense, definition (i). Nonetheless, given the nature in which the data is collected, one of the main arguments often given in economics to why time series data seems to have long memory features is due to cross-sectional aggregation. It is also commonplace to see arguments for cross-sectional aggregation motivating the presence of fractional long memory in real data.

Granger (1980), in line with the results of Robinson (1978) on random $AR(1)$ models, showed that cross-sectional aggregation of $AR(1)$ processes with random coefficients could produce long memory. Assuming a Beta distribution for the generation of cross-sectional $AR(1)$ coefficients, he showed that, as the cross-sectional dimension goes to infinity, the autocovariance function exhibits a slow hyperbolic decay, rather than the standard geometric decay characterizing *ARMA* processes. Thus, cross-sectional aggregation of dynamic micro units can produce long memory in the covariance sense under certain conditions.

In this paper we focus on some features of the aggregation argument leading to long memory. We address the particular specification considered by Granger because the Beta distribution is a rather flexible specification that allows closed-form solutions but the analysis can be extended to other aggregation schemes as well. Zaffaroni (2004) shows that Granger's result applies to a broader class of distributions to which the Beta distribution belongs. We demonstrate that this aggregation scheme implies that the aggregated series is long memory using all the definitions considered in this paper.

Since the aggregation result is an asymptotic property we conduct a Monte Carlo simulation study to quantify how aggregation can lead to long memory in finite samples. The theoretical degree of memory of the aggregated series is tied to a particular parameter of the Beta distribution which affects the density mass around one. The simulations show that the cross-sectional dimension has to be rather large for the theoretical degree of memory to apply, while the time series dimension needs to be large to obtain a precise estimator. Finite samples of the series will still exhibit long memory but the estimated memory parameter can be rather large compared to its theoretical value, especially when the memory is only of moderate degree.

In the third part of the paper, we focus on the extent to which the memory implied by aggregation can be removed by fractional differencing. In particular, we are interested in how *ARFIMA* type of long memory models can be useful for practical model building for the class of processes considered. It occurs that fractionally differencing the series, using the theoretical degree of memory, does remove the long memory of the process. The resulting series has absolutely summable autocorrelations and thus it is $I(0)$ by the definition of Davidson (2009). However, the fractionally differenced series will still have autocorrelations that decay hyperbolically and hence will decay slower than what an *ARMA* specification will be able to fit. This feature is most dominant when the degree of memory is moderate as opposed to being close to non-stationarity. Our findings have implications for the argument that is often given for estimating *ARFIMA* models in applications, namely that the observed long memory of time series can occur due to cross-sectional aggregation. A simulation study shows that fitted *ARFIMA* models will generally be inadequate to fit the dynamics of the underlying process.

The paper is structured as follows. In Section 2, the Granger aggregation scheme is presented and the features of the aggregated series are examined using the different long memory definitions that we consider. Section 3 presents the simulation study, and Section 4 derives the features of fractional differencing of cross-sectionally aggregated long memory processes. Finally, Section 5 concludes.

2. Long memory and cross-sectional aggregation

Consider the random $AR(1)$ process given by:

$$x_{i,t} = \alpha_i x_{i,t-1} + \varepsilon_{i,t}, \tag{2}$$

where $\varepsilon_{i,t}$ is a white noise process independent of α_i with $\mathrm{E}[\varepsilon_{i,t}^2] = \sigma_\varepsilon^2$, $\forall t \in \mathbb{Z}$ and $\alpha_i^2 \sim \mathcal{B}(\alpha; p, q)$ with $p, q > 1$ and $\mathcal{B}(\alpha; p, q)$ is the Beta distribution with density:

$$\mathcal{B}(\alpha; p, q) = \frac{1}{B(p,q)}\alpha^{p-1}(1-\alpha)^{q-1} \quad \text{for} \quad \alpha \in (0,1), \tag{3}$$

where $B(\cdot,\cdot)$ is the Beta function.

Robinson (1978) showed that the process given by (2) admits a variance–covariance stationary solution. Furthermore, the unconditional autocorrelation function of this process shows hyperbolic decay. However, the process is not ergodic in the sense that random samples will depend on the realization of α_i.