LONG MEMORY IN TIME SERIES OF ECONOMIC GROWTH AND CONVERGENCE

by

Gerald Silverberg
and
Bart Verspagen (MERIT & ECIS)

Paper prepared for Session II-3: Applied Dynamic Modeling, of the Austrian Focus Sessions of the 10th SASE Conference, Vienna, Austria, July 13-16, 1998.

April 1999

MERIT, Maastricht University,
PO Box 616
NL-6200 MD Maastricht
The Netherlands

ECIS, Eindhoven University of Technology
Faculty of Technology Management
PO Box 513
NL-5600 MB Eindhoven
The Netherlands

Email:
gerald.silverberg@merit.unimaas.nl
bart.verspagen@merit.unimaas.nl

1. Introduction

One of the most hotly debated topics in macroeconomics in recent years has been the nature of fluctuations in the growth process of aggregate output in the longer term. Traditionally, economists have conceived of the growth process as consisting of a deterministic trend (such as exponential growth) on which is superimposed either deterministic cyclical fluctuations (a decidedly minority view in the meantime) or a stable, mean-reverting stochastic process of ARMA type. Since Nelson and Plosser (1982) this view has been challenged by purported evidence of a *unit root* in univariate time series such as GDP, i.e., essentially a random walk component, which puts into question the notion of any underlying deterministic trend. However, this knife-edge distinction between a unit root and a near unit root but stable process has proven difficult to maintain empirically, especially since the main test employed, the Dickey-Fuller test, has been shown to have limited power against a range of alternative hypotheses.

The main empirical feature about which most authors agree consists of the observation of *persistence* in aggregate time series, meaning concretely the slower decay of the autocorrelation function than would be expected of a pure ARMA process (which is geometric in the lag). An equivalent way of looking at this is in the frequency domain, where it was already observed by Granger (1966) that most macroeconomic variables display a typical spectral shape, apparently diverging to infinity at zero frequency and declining uniformly at high frequencies. Time series with this property show periods of all lengths and have no characteristic time scale, and are often referred to as *aperiodic cycles*. The random walk itself is one example of this class of processes, but in fact it is only a special case of a family of such processes which can be parameterized by their degree of *fractional integration*. Given that agreement holds about persistence, one can enlarge the test hypothesis from unit root or not to the question of the degree of fractional integration of the process.

A second question naturally arises as to what the underlying mechanism in the economy could be which would produce this typical pattern. One path is via dynamic aggregation of an economy with cross-sectional heterogeneity (Granger 1980, Lo and Haubich 1989, Michelacci 1997). Another, somewhat related approach, invokes a dynamic, evolutionary framework of technology diffusion which naturally results in cross-sectional heterogeneity (Silverberg and Lehnert 1996, Silverberg and Verspagen 1995, 1996). We will describe this approach in more detail in section 4.

Why is persistence and long memory of interest in the analysis of long-term patterns of growth and fluctuations? For one thing, it is intermediate between a relatively unstructured stochastic world in which the present is just the summation of unrelated random events in the past (a random walk), and a rigidly predictable deterministic cycle or trend with relatively negligible, mean-reverting stochastic disturbances. It preserves the notion of even the distant past continuing to influence the present in a somewhat law like fashion, allowing the future to be structured while remaining shrouded in a haze of uncertainty. It also allows for developments at a range of time

scales, with what appear to be trends at one time scale being revealed to be parts of irregular cycles at longer time scales. Moreover, many natural phenomena display long memory, such as the flows of rivers, which originally led Hurst (1951) to first investigate the matter statistically.

Economists have been interested in long memory since the original work of Mandelbrot in the 1960s on price series. Since then attention has concentrated on long memory in asset prices and foreign exchange rates, and within the last ten years, on real national output measures as well. In the present paper we propose to use recent refinements in technique due to Beran (1994) and Sowell (1992a), and the more extensive national income data set of Maddison (1995) to examine both long memory in national times series and measures of international convergence. The paper is organized as follows. Section 2 reviews definitions of long memory and various statistical tests which allow hypothesis testing. Section 3 briefly reviews the literature on long memory in real and monetary economic variables. Section 4 discusses possible modeling approaches yielding long memory in output variables, including some of our own work on evolutionary selforganization in an endogenous growth setting. Our main empirical results are presented in section 5. Section 6 concludes with a discussion of possible alternative hypotheses and directions for further work.

2. Long Memory: Definition and Statistical Tests

A good survey of the basic mathematical statistics can be found in Beran (1994), while the econometric literature on long memory was recently reviewed by Baillie (1996). Here we will only recapitulate some of the salient points relevant to our work. There are several related ways of defining long memory of a discrete, real-valued time series y_i formulated either in the time or the frequency domains:

$$\gamma_k \sim const \cdot k^{2d-1}, \ as \ k \to \infty,$$

where γ_k is the autocovariance function at lag k;

$$f(\omega) \sim const \cdot \omega^{-2d}, \ as \ \omega \to 0^+,$$

where $f(\omega)$ is the spectral density of the process. A short-memory process such as an ARMA process will have exponentially decaying autocovariances and 0 spectral density at the origin. Two well-known cases can be singled out: Gaussian white noise, for which $d = 0$, and a random walk, where $d = 1$. Another definition involves the sum of the autocorrelations:

$$\lim_{n \to \infty} \sum_{j=-n}^{n} |\rho_j| = \infty.$$

Long memory was first related to *fractional integration* by Granger and Joyeux (1980) and Hosking (1981) using the following discrete time stochastic process:

$$(1-L)^d y_t = \epsilon_t$$

where L is the lag operator, d is the fractional integration parameter, which need not be an integer, y_t is the stochastic process, and ϵ_t is iid white noise.

Hurst (1951) developed the first estimation procedure for long memory processes, defining the Hurst parameter based on *rescaled range (R/S) analysis*, a natural quantity derived from his field of hydrology for determining the appropriate size of a reservoir. While R/S analysis is relatively easy to implement, it suffers from two disadvantages. First, little is known about the statistical properties of the estimator, so that it does not lend itself to hypothesis testing. And second, the results may be biased by the presence of short memory effects in addition to long ones. Lo (1991) proposed a modified rescaled range estimator to overcome both of these difficulties, but simulation results have not borne out its robustness.

Another approach is based on the characteristic pattern of the periodogram at low frequencies. Geweke and Porter-Hudak (1983) proposed an estimator based on regressing the log of the periodogram on the log of a trigonometric function of frequency. A difficulty with this approach concerns the number of values of the periodogram to use in the regression. They proposed a heuristic based on the length of the time series, but this estimator suffers as well from serious bias in the event of short memory (Baillie 1996, p.33). In a series of papers Robinson has further refined this approach.

An alternative is to use a full or approximate maximum likelihood estimator. One such approximation is due to Whittle (1951) in the frequency domain. This is the basis for an estimator described in detail in Beran (1994). The exact maximum estimator has been derived by Sowell (1992a). One way of dealing with short memory contamination in the context of maximum likelihood estimators is to simultaneously estimate a full ARFIMA(p,d,q) model of the form

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

where *A(L)* and *B(L)* are polynomials with roots outside the unit circle of order p and q, respectively. The basic Fractional Gaussian Noise model is reobtained by setting *p* and *q* equal to 0. Computer programs for the Whittle and exact likelihood estimators have been developed by Beran, and Doornik and Ooms, respectively. With the availability of fast computers, the additional computational complexity of maximum likelihood methods is no longer a problem on PCs.

While asymptotic results have been derived for the maximum likelihood estimators, their small-sample properties can differ considerably. Recently, Hauser (1998) has systematically compared the biases, mean square errors, and other measures of the various maximum likelihood estimators on samples of size 100 using Monte Carlo methods. He concludes that the Whittle estimator with

tapered data is most reliable. The exact maximum likelihood estimator for demeaned data, in contrast, "should only be used with care for fractionally integrated models due to its potential large negative bias of the fractional integration parameter."

The tapering used with the Whittle estimator transforms the original time series y_t to reduce the leakage effect of the periodogram:

$$y_t' = h_t y_t / H,$$

$$H = \sum_{t=1}^{n} h_t^2$$

where

$$h_t = \begin{cases} 1/2[1 - \cos\{\pi(t-1/2)/l\}] & t = 1,...,l \\ 1 & t = l+1,...,n-l \\ 1/2[1 - \cos\{\pi(n-t+1/2)/l\}] & t = n-l+1,...,n \end{cases},$$

with a taper chosen such as $l = \sqrt{n}$.

Hauser *et al.* (1999) are critical of ARFIMA models for estimating persistence in aggregate output. They show that overparametrization of an ARMA model may bias the estimates of persistence downwards, an effect which may be reflected in some of our results reported later in this paper.

3. Studies of Long Memory in Real and Financial Economic Variables

One of the main motivations for searching for long memory in economic time series is an observation by Granger (1966) about the `typical shape' of the spectra of economic variables. Considerable interest has focused on long memory in asset price series, a literature which is reviewed by Brock and De Lima (1996). The evidence is still rather ambiguous, although long memory seems much more likely in asset volatility than in asset returns themselves. Studies of real output measures commenced with Diebold and Rudebusch (1989) and Haubrich and Lo (1989). The former authors examined a number of US output time series using a two-stage process based on the GPH estimator to determine the degree of fractional integration, and an approximate maximum likelihood ARMA on the filtered residuals to handle the short memory aspect. They found significant evidence of long memory in these series. In contrast, Haubich and Lo, using Lo's modified rescaled range estimator, found no evidence for long memory in a quarterly time series

4

of postwar US data, as opposed to using the unmodified R/S statistic. Both of these studies suffer, however, from the estimation problems raised in the previous section.

These estimation problems can be circumvented in some extent by employing a maximum likelihood estimator of a full ARFIMA model. Sowell (1992b) reexamined postwar US quarterly real GNP with the exact maximum likelihood estimator, nesting deterministic trend and unit root models within the ARFIMA model. He reports a significantly smaller estimate of the fractional parameter than Diebold and Rudebusch (1989) due to the bias of the latter method. He concludes that the data are consistent with both trend models. In fact, the best model according to the AIC from a set of estimations with p and q ranging from 0 to 3 is ARFIMA(3, -0.59, 2).

Recently, tests have been conducted by Michelacci and Zaffaroni (1997) for a sample of countries. They also find a significant number of positive results and relate them to the issue of beta convergence. In the present study we will make use of the availability of more extensive data sets (both in terms of length and number of countries) and maximum likelihood estimators to reexamine both long memory in individual country growth series and in some intercountry convergence indicators.

4. Theoretical Basis for Long Memory in Aggregate Variables

Although long memory has been established in numerous domains of the natural sciences and is at least an open question in social sciences, until now there has been a dearth of theoretical explanations for its occurrence. Of course one might argue that climatological long memory might spill over into the economic domain without the need for a specifically economic mechanism, much as sunspots had been proposed as the driving fource of business cycles by Jevons. While this is not entirely implausible, it seems more reasonable to look for an explanation for long memory intrinsically rooted in the economic process.

An observation due to Granger (1980) is useful in this regard. He remarked that if the economy is composed of a large number of individual units each following an AR(1) process, and whose coefficients are drawn at random from a distribution (such as a beta distribution on (0,1)), then the additive aggregate of these variables will display long memory. Thus cross-sectional heterogeneity can be a source. However, this cross-sectional heterogeneity is simply assumed to be an exogenous and invariant given, and does not itself emerge from a more fundamental economic mechanism. Haubich and Lo (1989) also invoke an argument in this vein. Michelacci (1997) take this argument a step further towards endogenization by relating it to a technological vintage model in which small and large firms position themselves at different relative distances to a moving technological frontier, leading to a modified Gibrat-like stochastic process.

Endogenous cross-sectional heterogeneity, however, is one of the hallmarks of evolutionary models.[1] In these models firms are usually characterized by bounded rationality and rule-of-thumb search behavior, possibly with some form of adaptive learning, basic uncertainty, disequilibrium interactions, and system-emergent selection along the lines of Schumpeterian competition. Cross-sectional heterogeneity will be a natural outcome of such a model, as numerous authors have remarked, but heterogeneity that may be time-varying, and that involves constant reordering of the units in the distribution as a result of their relative performance.

A sequence of related growth models developed by Silverberg and Lehnert (1993), Silverberg and Verspagen (1994, 1995, 1996) will furnish the basis for the following discussion. Consider a `technology space' consisting of a line of discrete technologies, with technologies to the right being improvements on their immediate predecessors by a fixed factor.[2] At any point in time the firms in this economy will be utilizing some (finite) subset of the nodes in this space. The economy is assumed to progress by a) discovering new nodes to the right, one step at a time, and b) diffusing the known superior technologies into the capital stock by investment[3] and eliminating obsolete technologies by scrapping.[4] The discovery of new technologies will be assumed to be governed by a stochastic process (such as Poisson) whose rate will either be exogenously given or endogenously determined by the profit-seeking activities of the agents themselves.

A common feature of all versions of this underlying structure is that time series of aggregate variables display Granger's `typical shape' in the frequency domain. At the same time technologies diffuse into and out of the economy according to the well-known logistic replacement laws, firms grow and contract, and structural breaks can occur in such system-level characteristics as concentration indexes. In other words, a complicated, endogenously generated process of structural change is induced which displays many of the well-known empirical micro and meso-regularities as well as an aggregate growth process.

Silverberg and Verspagen (1995) take the model a step further into the international domain by allowing for a world of different national economies structured in this way with limited possibilities of technological and behavioral transfer between countries. The main result is that such a world economy itself will be characterized by aperiodic cycles of convergence and divergence of the long-memory type. Measures of inequality such as the coefficient of variation

---

[1] Some classical references in this regard are Nelson and Winter (1982), Iwai (1984a,b) and Silverberg, Dosi and Orsenigo (1988).

[2] For example, having a 10% higher labor productivity. Silverberg (1984) demonstrates that improvements in labor productivity are the predominant direction of technical change in evolutionary settings.

[3] i.e., innovations are assumed to be *capital-embodied* and thus require investment both for their discovery and their exploitation. In this latter respect these models differ from Nelson and Winter's growth model and endogenous growth models à la Aghion and Howitt (1992).

[4] In this respect these models resemble traditional vintage models except that investment and scrapping take place on all vintages simultaneously and not only on the marginal ones. For this reason, we denote them as *quasi-vintage* models.

of GDP per capita between countries (standard deviation divided by mean) show an irregular pattern of fluctuation at all time scales. What may look like a deterministic pattern of convergence over, say, 100 years, may later be reversed by an equally long period of divergence. Thus over finite windows of observation it becomes impossible to differentiate between apparent trends and segments of superlong `cycles.'

We believe the exclusive focus on apparent convergence in OECD countries in the postwar period obscures the more complex overall picture, as several authors have also pointed out. First, there is a problem of sample bias due to testing convergence only on a sample of countries with high per capita income in the final period. Looking at a more inclusive sample clearly demonstrates that there is no unconditional tendency to convergence (in the sense, e.g., that countries with lower per capita incomes in the initial period have proportionally higher growth rates over the period—one need only think of the African countries). Second, there is also a temptation on the basis of the Solow growth model to expect universal convergence to a common world frontier without ever asking what were the sources of backwardness in the more distant past (or else there never would have been a problem of convergence in the first place) nor of divergence in the `present' (as the Twentieth Century history of some of the more advanced Latin American countries such as Argentina, or the most recent experience of the Asian tigers demonstrate). Prima facie one could argue that precisely such an irregularly fluctuating pattern accords better with the historical record than a basic underlying tendency to convergence and that it is more reasonable to posit two opposing forces at work: catchup due to technological and behavioral imitation on the one hand, and the operation of virtuous and vicious cycles in the technological and institutional domains allowing the rich to grow richer and the poor poorer for extended periods of time. Combined with the operations of chance and threshold effects, these same positive feedbacks can allow some countries to take off and even overtake the leaders while others may fall behind and forfeit long-dominant positions. While at first glance such patterns can appear to be devoid of interpretable regularity and of the simplicity of a convergence hypothesis, the net result may precisely be long-memory time series properties of the type we will investigate in the next section.

5. Estimation results

The data set we employ is taken from Maddison (1995). It covers 29 countries for which a time series on GDP (and population) of 70 years or more is available. The longest time span for which data is available is 124 years: 1870-1994. This is the case for 15 countries, which are mostly highly developed countries in the postwar period. Outside the present OECD area, seven Latin American countries are included, as well as four Asian countries in addition to Japan.

The data are all expressed in US$, 1990 international prices (Geary Khamis PPP). More details about the construction of the data can be found in Maddison (1995). For the present purposes,

it should be kept in mind that in many cases, data had to be estimated, especially for the early years and the less developed countries.

Our first test of long memory is one proposed by Beran (1994), based on the Whittle approximate maximum likelihood method. This test models the time series under consideration as a fractional Gaussian noise process, and estimates the Hurst parameter h $(= d+\frac{1}{2})$. A value of h equal to 0.5 indicates no persistence (i.e., Gaussian white noise), while a value larger (smaller) than 0.5 indicates (anti) persistence. We use Beran's software, which is coded in S+, and is available from his Web page or from the appendix of his book.

Our first set of results, documented in Tables 1 to 5, focuses on time series for output at the individual country level. We consider both GDP, and GDP per capita, both series being first differenced in logs to remove the obvious nonstationarity. We used the standard deviation of the estimated h parameter provided by Beran's program to calculate a t-statistic for the null hypothesis $h = 0.5$, and the standard deviation of the d parameter in the ARFIMA case to calculate a t-statistic for the hypothesis $d = 0$.[5]

[insert Table 1 about here]

Table 1 gives the results for GDP per capita, using the complete sample period. For the 29 countries, we find 8 cases for which $h > 0.5$ at the 1% confidence level, another 5 cases for which this is the case at the 5% level, and 3 cases with 10% significance. There are two cases (India, New Zealand) for which we find antipersistence ($h < 0.5$) significant at the 10% level. Thus, in 16 out of 29 cases we find $h > 0.5$, so that approximately 60% of all countries display long memory for GDP per capita, with two displaying antipersistence.

Visual inspection of the time series reveals that in many cases, the Second World War had a major impact on (per capita) GDP. In the US, as well as in some other countries, the war led to a major boom in GDP, while especially in the European countries output collapsed in the 1940 - 1945 period due to war destruction. Also, the periods immediately before and after the war are very different, with the Great Depression sweeping through the world in the 1930s, while the 1950 - 1970 period was generally one of high growth and significant catching up to the United States, the only participant in the war to emerge physically and institutionally unscathed (in Maddison's terms, this is the 'golden age' of capitalism).

This led us to investigate our hypothesis for the pre- and postwar period separately, as well as for the complete period, excluding the war. For this purpose, we took 1937 and 1950 as demarcation years, i.e., we defined the prewar period as the period up to (and including) 1937,

---

[5] Beran shows that his test statistic is asymptotically $N(0,1)$. We use a two-sided test of $h = 0.5$ to detect departures in the direction of both persistence and antipersistence.

and defined the postwar period as the period from 1950 (inclusive) onwards.[6] We thus obtain three new time series: one for the prewar period, one for the postwar period, and one in which we concatenate these two periods. Note that the latter procedure is somewhat artificial, as it involves the 'glueing together' of the years 1937 and 1950 (which in many countries represented similar levels of income). While this procedure may appear a bit unorthodox, it does provide a handy device for testing to what extent the war years alone influence the detection of long memory in the series.

[Insert Table 2 about here]

The results for these three additional cases are documented in Tables 2 - 5. For the prewar period, documented in Table 2, we find very little evidence for long memory in GDP per capita. For the FGN test, of the 11 cases which show results that are significant at the 10% level or better, five cases point to antipersistence ($h < 0.5$): Denmark, France, India, Japan, and the Netherlands. The cases for which we find significant long memory in the prewar period are Austria (1% level), Finland, the UK and Indonesia (5% level) and Switzerland and Venezuela (10% level).

[Insert Table 3 about here]

The postwar period is a different story. We find 21 cases (i.e., the large majority) with significant long memory, and only one case of significant antipersistence. Most of the results are significant at the 1% (10 cases) or 5% (7 cases) levels.

[Insert Table 4 about here]

Table 4 documents the results for the time series where we leave out the war years. This yields 14 cases with significant (8 cases at the 1% level, two at the 5% level, four at the 10% level) long memory, and one case with significant antipersistence (5%, India).

While the results for the different subsamples are often consistent for many countries, there are a number of interesting exceptions. Denmark is significantly antipersistent prewar, but essentially Gaussian for all other samples. In contrast, Spain is highly significantly persistent except for the prewar period. Japan and the Netherlands are antipersistent prewar, persistent postwar, but neither over the larger samples. On the basis of the FGN test alone, one might venture the hypothesis that as countries moved from a peripheral, semi-industrial status to an industrial one

---

[6] The years 1937 and 1950 are 'inclusive', which, in terms of first differences, means that the last datapoint in the prewar period is the difference between GDP in 1937 and 1936, while the first datapoint of the postwar period is the difference between 1950 and 1949.

in the world economy their growth dynamics shifted from antipersistent to persistent, or at least that the value of their Hurst parameter increased.

The effects of including or excluding the Second World War years in the full sample period are also consistent, with the major exception of the United States, which loses its highly significant long memory result (1% level) by excluding them, even though long memory is also confirmed for the postwar years (but only at the 10% level).

[Insert Table 5 about here]

Although we feel that GDP per capita is a better indicator of economic growth, we also report results based on GDP without correcting for population growth. This follows more closely the approach taken by previous studies of long memory in aggregate output as well as the literature on unit roots in GDP data. Table 5 reports these results, which we only document for the full period (including the war). The results are significant in 16 cases (12 at the 1%, four at the 5% level). In general, the results are consistent with those for GDP per capita over the whole period. The four exceptions to this are India and Korea, for which we found (weakly) significant persistence in GDP per capita but no significant result for GDP; Mexico, for which we found no significant long memory in GDP per capita but a significant value for GDP; and New Zealand, which showed antipersistence in the case of GDP per capita, but no significant result in the case of GDP. New Zealand is also a curious case because antipersistence is only detected in the postwar subperiod.

Overall, these results seem to indicate that long memory models may be useful for describing fluctuations in aggregate output. However, the presence of long memory in both test models seems to depend to a considerable extent on its occurrence in the postwar rather than in the prewar period, as well as the effect of the war years. Partly, this may be due to the quality of the data, which is generally better for the postwar period. However, one should not exclude the possibility that a structural break internal to some of these countries, or a more general change in the postwar economic regime, is responsible for this result. There are certainly enough institutional and other circumstances that changed to account for such a break (e.g., the Bretton Woods system, the Cold War and the 'Pax Americana'; see Maddison 1991 for a more detailed account of the changing economic environment over the time span for which we have data).

We also investigated the hypothesis of long memory in indicators of convergence. To this end we defined two indicators, one specific to each country and the other based on a sample of countries. The first one measures the GDP per capita gap relative to the US as the log of the ratio of the two values. The US can be considered as the technology (or productivity) leader over most of the period under consideration. (Note that during the earlier years—late 19th century—the UK can be seen as the technology leader, and there is even a brief time span at the end of the 19th century when Australia has the highest value of GDP per capita in the sample).

[Insert Table 6 about here]

We apply the same test as before to the time series of GDP per capita relative to the US. Again, we take first differences of the log of this indicator. The results are presented in Table 6. There are 19 (out of 28) cases for which significant long memory is found, and one case (Canada), for which significant antipersistence is found. Of the significant cases with $h > 0.5$, 15 are significant at the 1% level, and the other 4 cases are significant at the 5% level.

It is relatively well known that there is a difference in terms of convergence between country groups. In general terms, the OECD countries have converged in terms of per capita GDP levels in the postwar period, while countries outside the OECD group have a mixed experience of either convergence (e.g., some of the Asian 'tigers') or divergence (e.g., many Latin American countries). See Baumol (1986) and Verspagen (1993) for evidence and interpretations of these phenomena. In the group of countries in which we find significant long memory, one finds both countries which have converged (to the US), and countries which have fallen behind. Examples of the latter are Argentina and Venezuela.

Convergence can also be measured within a group of countries, rather than a binary comparison of individual countries to the US. For this purpose, we use the coefficient of variation of GDP per capita (standard deviation divided by the mean) for various country subsamples. This is generally used in the literature as an indicator of so-called σ-convergence (e.g., Dollar and Wolff, 1992). We construct time series for the first difference of the coefficient of variation for four different samples of countries, to which we apply the long memory test.

[Insert Table 7 about here]

Table 7 documents the results of this procedure. Our base sample is labelled 'intermediate sample' in the table. This is the group of 14 OECD countries for which the maximum time span is available. For this sample, we find highly significant long memory for the complete time period. This result differs from that of a second subsample ('small sample') of eight countries consisting of non-European OECD countries, or European countries which are somewhat peripheral (Northern Europe), plus the US as reference economy. Some of these countries did not show very significant results in Tables 1 - 6. In this smaller sample, no significant long memory is found.

Japan is absent from the two samples considered thus far, because there are no Japanese data before 1886. The third sample adds Japan to the intermediate sample, thereby reducing the number of observations to 109, but enlarging the country sample to include all major OECD industrial countries in the final period. This does not change the result on long memory noticeably, with significant persistence prevailing. The final sample adds nine countries to the third sample, thereby reducing the time period further. Among the nine countries added, there are many non-OECD countries, some of which have been diverging relative to the mean value of GDP per capita. In this sample, we find again strong support for long memory.

11

Splitting up the period around the Second World War, we have three additional sets of estimations in Table 7. The differences compared to the results for the complete period are similar to the ones obtained for subperiods in Table 2 - 4. For the prewar period, as well as for the complete period excluding the war years (this time series is constructed in the same way as before), no support for the hypothesis of long memory is found. All results point to a value of *h* which is not significantly different from 0.5. In the postwar period, a value significantly larger than 0.5 is found for three of the four samples. In other words, the long memory result for the complete period seems to be driven mainly by the postwar period. Overall, we thus find strong support for long memory in time series of convergence of GDP per capita in various samples of countries for the postwar period.

We also applied a second test of long memory, which applies an ARFIMA model to take account of short memory, as discussed in Section 2 above. The ARFIMA model is implemented using both the software of Doornik and Ooms (1996), which uses a full maximum likelihood method, and Beran's ARFIMA software, which applies the Whittle maximum likelihood method. In the latter case, we apply tapering to the data as described in Section 2. These ARFIMA models were implemented for all combinations of ARMA parameters in the range 0..4.

We do not document all the ARFIMA results for reasons of space. A full set of the results is available from the authors on request. In general, however, the results from the ARFIMA approach differ substantially from those using the FGN model documented above. In fact, in the majority of cases where we find persistence using the FGN test, we do not find such a result for the ARFIMA model. The time series without the war years seem to yield the most consistent results in this respect, but even here there are large differences. Based on the two tests together, we find the strongest support for long memory in the cases of Spain (postwar, time series without war and total period), Austria, Germany, France and Japan (all postwar), and Italy, South Korea and Venezuela (time series without war years). In general, there are no large differences between the ARFIMA results using the Doornik and Ooms test and the Beran test (with tapered data).

The weaker results obtained from the ARFIMA model may well result from the problems with this model in estimating persistence as identified by Hauser *et al*. (1999), as touched upon in Section 2. Our impression from the ARFIMA estimates was that the results for different ARMA parameters vary greatly. Both the estimated values of these parameters and that of the long memory parameter jump around quite a bit for different ARMA specifications. This may indeed indicate the sensitivity to overparametrization as identified by Hauser *et al*. (1999), and therefore bias our estimates of persistence downwards.

Nevertheless, the ARFIMA results are indicative of some of the problems of separating short and long memory. This is why we document a subset of our ARFIMA results and discuss the implications. We choose the results for the time series of the coefficient of variation of GDP per capita. These are documented in Table 8, for the Beran-Whittle test with tapered data and for various ARMA parameters. As is seen in the table, almost no significant long memory is found.

The only case is for ARMA(3,3), intermediate sample, full period. Most of the other estimates point to significant anti-persistence.

Our overall conclusion is therefore that there is evidence for long memory in time series of economic growth, but that this evidence depends on the application of the FGN model. ARFIMA models, which take into account short memory as well, yield significantly less evidence of long memory.

6. Concluding Remarks and Directions for Future Research

Our statistical tests establish a strong case for pervasive but not universal long memory in time series of both individual countries and measures of convergence. The results must be relativized, however, because of their dependence to some extent on the subperiod examined and on the salient effect of the Second World War. While the effects of short memory can be separated out from those of long memory by employing maximum likelihood ARFIMA estimators, these models also yield much less favorable evidence of long memory.

In terms of establishing a theoretical basis for long memory, it would be useful to apply these tests to artificially generated data from the evolutionary models discussed in Section 4, under a range of assumptions and parameter values. Another question that should be pursued further is extending the class of alternative statistical models. For example, fractional Brownian motion and self-similar, scaling processes are related but not identical time series models which deserve attention, although they are not always amenable to the kind of hypothesis testing we have undertaken here. Levy-stable walks are also an alternative to Gaussian fractional noise where evidence exists of fat tails in the marginal distribution, as seems to be the case both in growth (major depressions and war-induced recoveries and collapses) and asset returns.[7]

In summary, improvements in both data resources and the statistical toolkit now place us at the threshold of understanding and estimating a much larger class of models than have dominated macroeconomic thinking until now. While care must be taken in applying and interpreting these tests, and the data must always be regarded with a certain healthy scepticism, we nevertheless feel that the effort will be richly rewarded in terms of the insights to be gained into the nature of the international process of economic growth.

---

[7] Cf. McCulloch (1996) for a review of the application of stable distributions to asset models.

References

Aghion, P. and Howitt, P., 1992, "A Model of Growth through Creative Destruction", *Econometrica*, **60:** 323-351.

Baillie, R., 1996, "Long Memory Processes and Fractional Integration in Econometrics", *Journal of Econometrics,* **73**: 5-59.

Baumol, W.J., 1986, "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show", *American Economic Review,* **76:** 1072-85.

Beran, J., 1994, *Statistics for Long-Memory Processes,* New York: Chapman & Hall.

Brock, W. and de Lima, P., 1996, "Nonlinear Time Series, Complexity Theory, and Finance", in Maddala, G. S. and Rao, C. R., (eds), *Handbook of Statistics 14: Finance,* Amsterdam: North Holland.

Diebold, F.X. and Rudebusch, G.D., 1989, "Long Memory and Persistence in Aggregate Output", *Journal of Monetary Economics,* **24:** 189-209.

Dollar, D. and E.N. Wolff, 1993, *Competitiveness, Convergence and International Specialization*, Cambridge MA: MIT Press.

Doornik, J.A. and M. Ooms, 1996, "A Package for Estimating, Forecasting and Simulating ARFIMA models", mimeo description of Ox ARFIMA package, Nuffield College, Oxford and Erasmus University, Rotterdam.

Geweke, J. and Porter-Hudak, S., 1983, "The Estimation and Application of Long Memory Time Series Models", *Journal of Time Series Analysis,* **4:** 221-238.

Granger, C. W. J., 1966, "The Typical Spectral Shape of an Economic Variable", *Econometrica,* **34(1):** 150-161.

Granger, C. W. J., 1980, "Long Memory Relationships and the Aggregation of Dynamic Models", *Journal of Econometrics,* **14:** 227-238.

Granger, C. W. J. and Joyeux, R., 1980, "An Introduction to Long Memory Time Series Models and Fractional Differencing", *Journal of Time Series Analysis,* **1:** 15-39.

Haubrich, J.G. and Lo, A.W., 1989, "The Sources and Nature of Long-Term Memory in the Business Cycle", Cambridge, MA: NBER Working Paper No. 2951.

Hauser, M. A., 1998, "Maximum Likelihood Estimators for ARMA and ARFIMA Models: A Monte Carlo Study", Working paper, University of Economics and Business Administration, Department of Statistics, Vienna, forthcoming in *Journal of Statistical Planning and Inference*.

Hauser, M.A., Pötscher, B.M. and E. Reschenhofer, 1999, 'Measuring Persistence in Aggregate Output: ARMA Models, Fractionally Integrated ARMA Models and Nonparametric Procedures', *Empirical Economics*, fortcoming.

Hosking, J. R. M., 1981, "Fractional Differencing", *Biometrika,* **68:** 165-176.

Hurst, H., 1951, "Long Term Storage Capacity of Reservoirs", *Transactions of the American Society of Civil Engineers,* **116:** 770-799.

Iwai, K., 1984a, "Schumpeterian Dynamics. I: An Evolutionary Model of Innovation and Imitation ", *Journal of Economic Behavior and Organization,* **5:** 159-90.

Iwai, K., 1984b, "Schumpeterian Dynamics. II: Technological Progress, Firm Growth and "Economic Selection"", *Journal of Economic Behavior and Organization,* **5:** 321-51.

Lo, A.W., 1991, "Long Term Memory in Stock Market Prices", *Econometrica,* **59(5):** 1279-1314.

Maddison, A., 1991, *Dynamics Forces in Capitalist Development*, Oxford: Oxford University Press.

Maddison, A., 1995, *Monitoring the World Economy 1820-1992,* Paris: OECD.

McCulloch, J. H., 1996, "Financial Applications of Stable Distributions", in Maddala, G. S. and Rao, C. R., (eds), *Handbook of Statistics 14: Finance,* Amsterdam: North Holland.

Michelacci, C., 1997, "Cross-Sectional Heterogeneity and the Persistence of Aggregate Fluctuations", London School of Economics: Working paper.

Michelacci, C. and Zaffaroni, P., 1997, "(Fractional) Beta Convergence", London School of Economics: working paper.

Nelson, C. R. and Plosser, C. I., 1982, "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications", *Journal of Monetary Economics,* **10:** 139-162.

Nelson, R. R. and Winter, S. G., 1982, *An Evolutionary Theory of Economic Change,* Cambridge MA: The Belknap Press of Harvard University Press.

Silverberg, G., 1984, "Embodied Technical Progress in a Dynamic Economic Model: the Self-Organization Paradigm", in Goodwin, R.,Krüger, M. and Vercelli, A., (eds), *Nonlinear Models of Fluctuating Growth,* Berlin-Heidelberg-New York: Springer Verlag.

Silverberg, G., Dosi, G. and Orsenigo, L., 1988, "Innovation, Diversity and Diffusion: A Self-Organisation Model", *Economic Journal,* **98:** 1032-54.

Silverberg, G. and Lehnert, D., 1993, "Long Waves and 'Evolutionary Chaos' in a Simple Schumpeterian Model of Embodied Technical Change", *Structural Change and Economic Dynamics,* **4:** 9-37.

Silverberg, G. and Lehnert, D., 1996, "Evolutionary Chaos: Growth Fluctuations in a Schumpeterian Model of Creative Destruction", in Barnett, W. A.,Kirman, A. and Salmon, M., (eds), *Nonlinear Dynamics in Economics,* Cambridge: Cambridge University Press.

Silverberg, G. and Verspagen, B., 1994, "Learning, Innovation and Economic Growth: A Long-Run Model of Industrial Dynamics", *Industrial and Corporate Change,* **3:** 199-223.

Silverberg, G. and Verspagen, B., 1995, "An Evolutionary Model of Long Term Cyclical Variations of Catching Up and Falling Behind", *Journal of Evolutionary Economics,* **5(3):** 209-227.

Silverberg, G. and Verspagen, B., 1996, "From the Artificial to the Endogenous: Modelling Evolutionary Adaptation and Economic Growth", in Helmstädter, E. and Perlman, M., (eds), *Behavorial Norms, Technological Progress and Economic Dynamics: Studies in Schumpeterian Economics,* Ann Arbor, MI: University of Michigan Press.

Sowell, F. B., 1992a, "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models", *Journal of Econometrics,* **53:** 165-188.

Sowell, F. B., 1992b, "Modeling Long-Run Behavior with the Fractional ARIMA Model", *Journal of Monetary Economics*, **29:** 277-302.

Verspagen, B., 1993, *Uneven Growth Between Interdependent Economies. The Evolutionary Dynamics of Growth and Technology,* Aldershot: Avebury.

Whittle, P., 1951, "Hypothesis Testing in Time Series Analysis", Uppsala: Almquist and Wiksells.

Table 1. Estimations results for long memory tests, GDP per capita, complete period, FGN model using Beran-Whittle estimation algorithm

Country	N	h	t(0.5)
Argentina	94	0.5252	0.39
Australia	124	0.5108	0.19
Austria	124	0.5296	0.52
Belgium	124	0.6357	2.33 **
Brazil	94	0.6110	1.67 *
Canada	124	0.6417	2.43 **
Switzerland	95	0.6604	2.40 **
Chile	94	0.4118	-1.42
Colombia	69	0.6158	1.49
Germany (West)	124	0.6945	3.30 ***
Denmark	124	0.4833	-0.30
Spain	94	0.7084	3.07 ***
Finland	124	0.6567	2.68 ***
France	124	0.6081	1.87 *
United Kingdom	124	0.7288	3.86 ***
India	92	0.3926	-1.73 *
Indonesia	92	0.9263	6.05 ***
Italy	124	0.6491	2.55 **
Japan	109	0.5745	1.22
South Korea	81	0.6334	1.85 *
Mexico	73	0.6174	1.55
Netherlands	124	0.5324	0.57
Norway	124	0.5447	0.79
New Zealand	124	0.3997	-1.87 *
Peru	81	0.4353	-0.96
Sweden	124	0.6416	2.43 **
Taiwan	89	0.7293	3.76 ***
Venezuela	94	0.76	3.80 ***
United States	124	0.6766	3.01 ***

One, two and three stars denote significance at the 10%, 5% and 1% level in a 2-sided *t*-test, respectively.

Table 2. Estimations results for long memory tests, GDP per capita, prewar period, FGN model using Beran-Whittle estimation algorithm

Country	N	h	t(0.5)
Argentina	37	0.6215	1.14
Australia	67	0.4377	-0.84
Austria	67	0.7911	3.58 ***
Belgium	67	0.5405	0.52
Brazil	37	0.3638	-1.41
Canada	67	0.6303	1.65
Switzerland	38	0.6859	1.75 *
Chile	37	0.3905	-1.12
Colombia	12	na	na
Germany (West)	67	0.4658	-0.45
Denmark	67	0.3539	-2.05 **
Spain	37	0.5879	0.83
Finland	67	0.6633	2.05 **
France	67	0.3668	-1.86 *
United Kingdom	67	0.6786	2.24 **
India	37	0.0858	-5.12 ***
Indonesia	37	0.726	2.08 **
Italy	67	0.5062	0.08
Japan	52	0.1789	-4.48 ***
South Korea	26	na	na
Mexico	16	na	na
Netherlands	67	0.3442	-2.20 **
Norway	67	0.3877	-1.55
New Zealand	67	0.4941	-0.08
Peru	24	na	na
Sweden	67	0.6268	1.60
Taiwan	34	na	na
Venezuela	37	0.6865	1.73 *
United States	67	0.5147	0.19

One, two and three stars denote significance at the 10%, 5% and 1% level in a 2-sided *t*-test, respectively.

Table 3. Estimations results for long memory tests, GDP per capita, postwar period, FGN model using Beran-Whittle estimation algorithm

Country	N	h	t(0.5)
Argentina	45	0.5231	0.25
Australia	45	0.6089	1.13
Austria	45	0.7875	2.90 ***
Belgium	45	0.6690	1.74 *
Brazil	45	0.8176	3.19 ***
Canada	45	0.4872	-0.14
Switzerland	45	0.7138	2.18 **
Chile	45	0.6246	1.29
Colombia	45	0.7227	2.27 **
Germany (West)	45	0.8275	3.28 ***
Denmark	45	0.5097	0.10
Spain	45	0.9065	4.04 ***
Finland	45	0.8504	3.50 ***
France	45	0.8365	3.37 ***
United Kingdom	45	0.6974	2.02 **
India	43	0.3968	-1.13
Indonesia	43	0.7243	2.23 **
Italy	45	0.8287	3.29 ***
Japan	45	0.9367	4.33 ***
South Korea	43	0.5869	0.89
Mexico	45	0.7664	2.69 **
Netherlands	45	0.6885	1.93 *
Norway	45	0.7041	2.08 **
New Zealand	45	0.2958	-2.44 **
Peru	45	0.7867	2.89 ***
Sweden	45	0.8630	3.62 ***
Taiwan	43	0.7585	2.56 **
Venezuela	45	0.6791	1.84 *
United States	45	0.6705	1.75 *

One, two and three stars denote significance at the 10%, 5% and 1% level in a 2-sided *t*-test, respectively.

Table 4. Estimations results for long memory tests, GDP per capita, war years excluded, FGN model using Beran-Whittle estimation algorithm

Country	N	h	t(0.5)
Argentina	82	0.5495	0.71
Australia	112	0.4825	-0.30
Austria	112	0.8211	5.08 ***
Belgium	112	0.6074	1.77 *
Brazil	82	0.6419	1.98 *
Canada	112	0.6016	1.67 *
Switzerland	83	0.7009	2.79 ***
Chile	82	0.4475	-0.78
Colombia	57	0.6030	1.21
Germany (West)	112	0.6090	1.79 *
Denmark	112	0.4513	-0.84
Spain	82	0.7701	3.68 ***
Finland	112	0.6939	3.13 ***
France	112	0.4738	-0.45
United Kingdom	112	0.6812	2.93 ***
India	80	0.3602	-2.14 **
Indonesia	80	0.7474	3.34 ***
Italy	112	0.6278	2.09 **
Japan	97	0.6063	1.63
South Korea	69	0.6621	2.06 **
Mexico	61	0.6074	1.30
Netherlands	112	0.4453	-0.95
Norway	112	0.5146	0.25
New Zealand	112	0.4441	-0.97
Peru	69	0.4616	-0.52
Sweden	112	0.6797	2.91 ***
Taiwan	77	0.6169	1.59
Venezuela	82	0.7028	2.79 ***
United States	112	0.5379	0.63

One, two and three stars denote significance at the 10%, 5% and 1% level in a 2-sided *t*-test, respectively.

Table 5. Estimations results for long memory tests, GDP, complete period, FGN model using Beran-Whittle estimation algorithm

Country	N	h	t(0.5)
Argentina	94	0.5605	0.92
Australia	124	0.5257	0.45
Austria	124	0.5142	0.25
Belgium	124	0.6428	2.45 **
Brazil	94	0.6385	2.07 **
Canada	124	0.6710	2.91 ***
Switzerland	95	0.7100	3.11 ***
Chile	94	0.4170	-1.33
Colombia	69	0.6189	1.53
Germany (West)	124	0.7056	3.48 ***
Denmark	124	0.4817	-0.33
Spain	94	0.7145	3.16 ***
Finland	124	0.6613	2.75 ***
France	124	0.6588	2.71 ***
United Kingdom	124	0.7325	3.92 ***
India	92	0.4953	-0.07
Indonesia	92	0.9317	6.11 ***
Italy	124	0.6575	2.69 ***
Japan	109	0.5688	1.13
South Korea	81	0.6021	1.43
Mexico	73	0.6748	2.28 **
Netherlands	124	0.5428	0.75
Norway	124	0.5447	0.79
New Zealand	124	0.4971	-0.05
Peru	81	0.4394	-0.90
Sweden	124	0.6473	2.52 **
Taiwan	89	0.7475	3.53 *
Venezuela	94	0.7561	3.74 *
United States	124	0.6798	3.06 *

One, two and three stars denote significance at the 10%, 5% and 1% level in a 2-sided *t*-test, respectively.

Table 6. Estimations results for long memory tests, GDP per capita relative to the US, complete period, FGN model using Beran-Whittle estimation algorithm

Country	N	h	t(0.5)
Argentina	82	0.6589	2.36 **
Australia	112	0.4142	-1.59
Austria	112	0.4971	-0.05
Belgium	112	0.7502	4.21 ***
Brazil	82	0.6977	2.92 ***
Canada	112	0.3701	-2.46 **
Switzerland	83	0.8851	5.57 ***
Chile	82	0.5132	0.20
Colombia	57	0.8393	4.20 ***
Germany (West)	112	0.6249	2.15 **
Denmark	112	0.6899	3.22 ***
Spain	82	0.6778	2.63 ***
Finland	112	0.6666	2.84 ***
France	112	0.7221	3.75 ***
United Kingdom	112	0.5746	1.30
India	80	0.5131	0.20
Indonesia	80	0.9043	5.74 ***
Italy	112	0.6885	3.20 ***
Japan	97	0.5560	0.92
South Korea	69	0.6023	1.43
Mexico	61	0.6847	2.41 **
Netherlands	112	0.6632	2.78 ***
Norway	112	0.7477	4.16 ***
New Zealand	112	0.4647	-0.64
Peru	69	0.6863	2.56 **
Sweden	112	0.6856	3.16 ***
Taiwan	77	0.7874	4.07 ***
Venezuela	82	0.7704	3.95 ***

One, two and three stars denote significance at the 10%, 5% and 1% level in a 2-sided t-test, respectively.

Table 7. Estimations results for long memory tests, convergence time series (coefficient of variation), FGN model using Beran-Whittle estimation algorithm

Complete period	Period	h	t (0.5)
Intermediate sample	1871 - 1994	0.7114	3.58 ***
Peripheral sample + US	1871 - 1994	0.5950	1.65
Intermediate sample + Japan	1886 - 1994	0.7272	3.59 ***
Large sample	1901 - 1994	0.7121	3.09 ***
Prewar			
Intermediate sample	1871 - 1937	0.4251	-1.01
Peripheral sample + US	1871 - 1937	0.4271	-0.99
Intermediate sample + Japan	1886 - 1937	0.4307	-0.82
Large sample	1901 - 1937	0.5691	0.66
Postwar			
Intermediate sample	1950 - 1994	0.7458	2.49 **
Peripheral sample + US	1950 - 1994	0.6220	1.27
Intermediate sample + Japan	1950 - 1994	0.8115	3.13 ***
Large sample	1950 - 1994	0.7303	2.29 **
No war			
Intermediate sample		0.4881	-0.20
Peripheral sample + US		0.4474	-0.91
Intermediate sample + Japan		0.5340	0.53
Large sample		0.5938	1.31

One, two and three stars denote significance at the 10%, 5% and 1% level
in a 2-sided *t*-test, respectively.

Table 8. Estimations results for long memory tests, convergence time series (coefficient of variation), ARFIMA model using Beran-Whittle estimator with tapered data

		ARFIMA(1,d,1)		ARFIMA(2,d,2)		ARFIMA(3,d,3)	
Complete period	Period	h	t(0.5)	h	t(0.5)	h	t(0.5)
Intermediate sample	1871 - 1994	-0.1344	-1.90 *	-0.5422	-3.91 ***	0.6606	1.60
Peripheral sample + US	1871 - 1994	-0.2966	-2.45 **	-0.9704	-5.85 ***	-1.3954	-7.76 ***
Intermediate sample + Japan	1886 - 1994	-0.1131	-1.68 *	-0.3740	-2.56 **	0.1942	-2.00 *
Large sample	1901 - 1994	-0.0878	-1.45	-0.3809	-3.29 ***	0.1880	-2.08 *
Prewar							
Intermediate sample	1871 - 1937	-0.5768	-2.18 **	-1.2561	-2.74 ***	-1.3977	-4.92 ***
Peripheral sample + US	1871 - 1937	0.4159	-0.71	-0.9660	-2.83 ***	-1.1606	-2.43 **
Intermediate sample + Japan	1886 - 1937	-0.0201	-3.51 ***	1.0062	1.28	0.6849	0.34
Large sample	1901 - 1937	-0.7009	-2.74 ***	-0.3522	-1.47	-0.4454	-0.74
Postwar							
Intermediate sample	1950 - 1994	-0.4036	-1.97 *	-0.3304	-1.12	-0.4635	-1.70 *
Peripheral sample + US	1950 - 1994	-0.5867	-2.31 **	-0.8160	-1.77 *	-1.6632	-7.13 ***
Intermediate sample + Japan	1950 - 1994	0.7324	1.55	-0.1067	-0.72	-0.2664	-1.41
Large sample	1950 - 1994	0.9332	1.80 *	-0.2498	-0.82	-0.2905	-2.68 ***
No war							
Intermediate sample		0.2867	-0.88	0.3098	-1.40	0.5775	0.60
Peripheral sample + US		0.5310	0.29	-0.4406	-1.92 *	-0.6219	-1.91 *
Intermediate sample + Japan		0.5858	0.66	0.6522	0.78	0.5158	0.08
Large sample		0.4793	-0.10	1.0462	1.08	0.5214	0.11

One, two and three stars denote significance at the 10%, 5% and 1% level in a 2-sided *t*-test, respectively.