FRACTIONAL OUTPUT CONVERGENCE, WITH AN APPLICATION TO NINE DEVELOPED COUNTRIES

Arielle Beyaert[#]

Universidad de Murcia(Departamento de Métodos Cuantitativos para la Economía) and CentrA

March 2004

Abstract:

In this paper we argue that cross-country convergence of output per capita should be examined in a fractional-integration time-series context and we propose a new empirical strategy to test it, which is the first one able to discriminate between fractional long-run convergence and fractional catching-up. The starting point of the paper is: since there are reasons to believe that aggregate output is fractionally integrated, the usual testing strategy based on unit-root or traditional I(1)-I(0) cointegration techniques is too restrictive and may lead to spurious results. Therefore, we propose a new classification of output convergence processes which is valid when outputs are fractionally integrated and which nests the usual definitions built for an I(1)-versus-I(0) world. The new empirical testing strategy that we propose can distinguish between these new types of convergence. It is based on the combined use of new inferential techniques developed in the fractional integration literature; compared with more traditional fractional-integration procedures, they offer the great advantage of being robust to both deterministic and stochastic non-stationarity (i.e. robust to the presence of a trend and/or to a memory parameter d above 0.5). We fully explain in the paper the importance of this advantage for testing convergence. Applying this strategy, we obtain that a group of developed countries (G-7, Australia and New Zealand) converged in the last century; we identify the type of convergence for each one. The main result is that per-capita-output differentials are typically *mean-reverting fractionally I(d), with d significantly above 0 but below 1*. This contrasts with the results of divergence obtained with six unit-root tests and also by other authors with I(1)-I(0) (co)integration techniques. The paper therefore contributes to solve the puzzling negative or inconclusive results about convergence usually obtained with I(1)-versus-I(0) tests; our results also prove that the proposed widening of the statistical definition of output convergence is necessary and that convergence does take place, but more slowly than traditionally expected.

Keywords: real convergence, fractional integration, unit roots, non-stationarity

JEL classification: C32, F43, O41

[#] email: arielle@um.es. This research has been partly financed by grants from Fundación Seneca, Project PL/8/FS/00 and from the Spanish Ministry of Science and Technology Project SEC2001/0855. All the programming and inference has been carried out in Ox, partially extending the Arfima Package of J.A. Doornik and M.Ooms (http://www.nuff.ox.ac.uk/users/doornik/)

1. Introduction

There is a vast empirical literature testing whether per capita output of different economies tend to converge over time, as predicted by the neoclassical growth model (Solow, 1956), or whether the economies do not contain any automatic mechanism that prevents them from divergent steady states, as argued by the more recent endogenous growth models (Lucas, 1988, Romer, 1986 and the posterior literature).

Discriminating between these two theoretical alternatives is obviously an important economic issue for policy makers at the national and international level. This explains why many studies have examined this question.

Output convergence between countries or regions has been tested with different types of samples, and different econometric procedures. A well-known methodology is the so-called "β-convergence" regression, associated to the works of Barro and Sala-i-Martín (1991,1992). It consists of a cross-sectional regression of the average growth rate of per-capita output over some long enough time-period on a constant, on the level of per-capita output at the beginning of the period and, if necessary, on a set of country-specific additional variables. For convergence, the coefficient of the output level at the beginning of the period should be negative, because this would reflect that the poorer the country (or region) in the initial year the faster it grew to catch up with the richer. The collected evidence with this approach is generally interpreted as favourable to some type of convergence (Baumol, 1986, Barro and Sala-i-Martin 1991, 1992, Barro, 1991, Mankiw, Romer and Weil, 1992). However, this cross-section approach has received many criticisms. Quah (1993,1996) considers that what is needed is a dynamic analysis of the per-capita income distributions of the different regions or countries over time. Evans and Karras (1996) argue in favour of a panel data approach. Bernard and Durlauf (1996) argue that within a time-series framework a more refined analysis of convergence is possible. In particular, in a time series context it becomes possible to detect which countries of a given group are converging and which are not and, within the first group, to separate the economies that are around or close to the common stationary state (what is known as zero-mean long-run convergence) from those who are still on the transition path towards this state (what is known as catching-up convergence). However, the empirical results in this context, applying unit-root and cointegration tests, are not very favourable to convergence. In most cases, some type of structural change has to be assumed in order to obtain better, although not overwhelming, evidence (see, for instance, Cellini and Scorcu, 2000). In a recent paper, Ericsson and Halket (2002) attribute the difficulty of detecting convergence with time series analysis to the use of univariate or single-equation testing procedures. They advocate the use of multivariate Johansen cointegration tests between the per capita outputs of the countries of interest. When working with the per-capita outputs of the G-7 countries, they reject convergence in the cases where unit-root tests on the output differentials (i.e. on the distance of the output-per-capita of one country from the output-per-capita of a benchmark country, which they choose to be the US) are applied; however bivariate and trivariate Johansen cointegration test applied on two or three series of outputs are more favourable to convergence (or at least to the existence of a cointegration relation between the countries per-capita outputs). But even in this multivariate framework, the evidence is not overwhelming.

However, an alternative explanation of these failures might be that the per-capita output series are not I(1), but rather *fractionally* integrated. A series is said to be fractionally integrated, or *FI(d)*, if it is integrated of order *d*, with d not necessarily integer (Granger and Joyeux, 1980, Hosking, 1981). When $d=0$, any shock that affects the series has only a short-term effect, which completely disappears in the long-run. In this case, we say that the series is "short-memory". On the contrary, when $d>0$, the value of the series will somehow be influenced by shocks that took place in the very remote past. In this case, the series exhibit "long-memory" or "persistence". The intensity of memory of the series will depend on the value of *d*: the smaller *d*, the less persistent will be the shocks. If $d<0.5$, the series is however stationary. A very interesting case for many economic issues is when $0.5<d<1$. In this situation, the series is long-memory and non-stationary *although mean-reverting*: in spite of the fact that remote shocks affect the present value of the series, this will tend to its mean value in the long-run; in other words, the series has long but *transitory memory* This is quite distinct from the case $d \geq 1$ in which the mean of the series has no influence whatsoever on the long-run evolution of it, because the series is dominated by all the remote and recent shocks. In this case, the series has *permanent or infinite memory*.

So, provided $d<1$, some sort of a long-run equilibrium level of the series exists which is represented by the mean. Obviously, when $0<d<1$, traditional unit root tests (testing $d=1$ under the null) or stationarity tests (testing $d=0$ under the null), developed after Dickey-Fuller and Phillips-Perron works, may fail to detect mean-reversion in the series and reach the wrong conclusion of infinite memory. Diebold and Rudebush (1991) provide theoretical arguments for this phenomenon and Gonzalo and Lee (2000) provide recent Monte Carlo simulations illustrating this problem.

Other problems occur when the series are FI(d) but are treated as if they were I(1): traditional I(1)/I(0) cointegration tests between two or more series are flawed not only because *d* differs from 1, but also because traditional cointegration is based on the assumption that all the possibly cointegrated series have the same order of integration, which is very difficult to guarantee when *d* may take non-integer values.

Even when all series have the same order of fractional integration, Gonzalo and Lee (2000) show that traditional cointegration tests and treating the series as if they were I(1) tend to find too much spurious cointegration, especially if the multivariate Johansen likelihood ratio tests are used. Note that the risk of spurious cointegration results might, at least in part, the somewhat better results in favour of convergence obtained by Ericsson and Halket(2002) when they use multivariate Johansen cointegration tests: their results might be spurious if the level of per-capita outputs are fractionally integrated instead of being I(1).

Michelacci and Zaffaroni (2000) is the first published paper to consider the possibility of fractional integration in per-capita outputs in an analysis of convergence. These authors point out possible contradictions between the traditional β-convergence concept and per-capita outputs being I(1) ; on the other hand they present theoretical reasons why outputs should be fractionally integrated. They then redefine β-convergence in terms of mean-reversion of *detrended levels of per-capita outputs* and obtain empirical results which are favourable to their new definition. Silverberg and Verspagen (1999) and Cunado et al. (2002) also examine convergence from a fractional-integration point of view, although they search for mean-reversion of the *output differentials*.

This paper also analyses convergence within a fractional-integration setting. It however differs in several aspects from the aforementioned papers carried out in this framework.

First, in those papers, the techniques used for estimating *d* require stationarity of the series on which they are applied: $d < 0.5$ and no deterministic trend in the series. Therefore, the authors work with detrended and/or differenced series. But the stationarizing transformations may distort the estimation of the integration order and eliminates from the data information that is relevant to the question under study; for instance, a significant trend in the output differential indicates that the countries stand in a catching-up convergence rather than in long run convergence. In this paper, we use instead recently developed estimators which are robust to nonstationarity in the original series: the possible presence of a trend or a value of *d* possibly above *0.5* do not require any detrending nor differencing of the series before estimating *d*.

Secondly, we also offer new and more precise definitions of convergence when fractional integration of outputs is allowed; they extend the traditional time-series concepts of convergence built in the I(1) paradigm to the context of fractional integration.

Thirdly, we propose and apply a strategy that allows discriminating between long-run convergence and catching-up in a fractional integration setting. This is a novel feature of convergence testing in this setting. Our strategy combines the robust estimators of *d* with new techniques (Marmol and Velasco, 2002) aimed at estimating and testing a trend in non stationary FI(d>0.5) series. Our paper therefore relies on techniques that are robust to the relevant characteristics of the series that we have to analyse and provides a more complete examination of the convergence processes than previous papers.

One of the main conclusions of the paper is that convergence has taken place between the G-7 countries, Australia and New Zealand over the last century. However, the convergence processes did not produce stationary output differentials: most differentials are FI($0.5<d<1$); this characteristic would explain why unit-root and cointegration tests applied in the past often have failed to detect convergence between output series.

The structure of the paper is as follows. Section 2 is dedicated to a brief description of the concept of fractional process and of the advantages of recently developed techniques for testing and estimating the memory parameter and the trend of fractionally-integrated non-stationary series (the Appendix offers to the interested reader a more technical discussion of these aspects). In Section 3, we first review the concept of growth convergence and we examine some possible limitations of the results obtained in the literature; we then offer new definitions of convergence in terms of the output differentials which are compatible with fractional integration of the output series and which generalises the definitions made in the I(1) - I(0) paradigm; finally, we propose an empirical strategy for detecting convergence based on the new inferential tools described in Section 2. We apply this strategy in Section 4 on the per capita outputs of G-7, Australia and new Zealand, published in Maddison (1995, 2001). Finally, Section 5 concludes.

2. Fractional integration:

general definitions, problems for estimating and testing the memory parameter and the deterministic component

The aim of this section is to offer an overview of the concept of fractionally integrated process, and of the problems generally encountered in estimating the memory parameter and the trend of a fractional process, together with possible solutions that have been developed recently. This overview is intentionally as little technical as possible. A more technical presentation is offered in the Appendix.

Consider the process X_t generated by the model $(1-L)^d (X_t - X_0) = u_t$ where X_0 is a random variable with a certain fixed distribution, *d* is not necessarily integer and u_t is a zero-mean stationary process. The process X_t is said to be fractionally integrated of order *d* or FI(*d)*. It is *stationary* and invertible if $|d| < 0.5$ and is non-stationary if $d \geq 0.5$. However as long as $d < 1$, the process is mean reverting, which means than any shock that affects the process at some point in time has a non-permanent impact on the value of the series in the future. If *d*=0, the unique source of dynamic correlation stems from the stationary dynamics of u_t; the process is then said to exhibit short memory. As soon as *d*>0, the dynamics of the process comes both from the short-run dynamics included in u_t and from the long-range dependence implied by the positive value of *d.* The process is then said to exhibit long-memory. However, this long memory is transitory if *d*<1 and becomes permanent if $d \geq 1$. The value of *d* is therefore an indication of the persistence of the shocks: the smaller *d* the less persistent will be the shocks. The value of *d* by itself is therefore of prime interest in many economic problems and not only in convergence studies. Of particular interest is to determine whether d<0.5, $0.5 \leq d<1$ or $d \geq 1$.

Different methods have been developed to estimate this long-range parameter independently from the short-run parameters belonging to the dynamics of u_t.

One of the most widely used methods is due to Geweke and Porter-Hudak (1983) and Robinson (1995a) and is based on a regression of the logarithm of the periodogram of the series on the logarithm of the Fourier frequencies. It is usually identified as the "GPH regression" or the "LP regression". Another frequently used estimator is the "local Whittle estimator" suggested by Künsch (1987) and further studied by Robinson (1995b). An important characteristic of both approaches is that they require the data to trend-free. Robinson (1995a) and Robinson (1995b) demonstrate the consistency and asymptotic normality of the LP estimator and of the local Whittle estimator, respectively. However, these results are limited to the interval $-0.5 < d < 0.5$. This explains why it has been a usual practice to difference the series previous to the estimation of *d* as soon as it is suspected that the order of integration of the original series might be above 0.5. Recently, these asymptotic properties have been partially extended. Velasco (1999a) establishes the consistency of the LP estimator for $-0.5 < d < 1$. Velasco (1999b) does the same for the local Whittle estimator. As far as asymptotic normality is concerned, he extends the result to $d<0.75$ for both estimators. In spite of these slightly more general results, differencing is still needed with these techniques if $d>0.75$ is suspected.

The practice of detrending and differencing the data before estimating *d* is however not innocuous. Sun and Phillips (1999) point out the importance of efficient detrending when estimating *d* with methods that require the series to be stationarized. They show that the usual OLS estimate of the trend is highly inefficient. As far as differencing is concerned, Agiakloglou et al.(1993) and Hurvich and Ray (1995) illustrate how the LP estimator of *d* is not invariant to differencing, so that a biased estimation may result from overdifferencing the series (i.e. from estimating *d* from a stationary series that has been unnecessarily differenced). Recently, new methods have been developed that allow estimating *d* without differencing nor detrending the original series.

To avoid differencing and detrending, Velasco (1999a) advocates the application of a weighting scheme on the original data - which is known as "tapering the data". It consists of applying a symmetric weighting scheme on the original data, such that the observations located in the central dates of the original sample receive more weight than the observations located at the beginning and at the end of the sample. The periodogram is then computed on these weighted data. Together with the tapering, Velasco(1999a) also proposes a modification of the LP regression method. The new procedure is valid for estimating *d* up to 3.75 and does not require detrending the data as long as the trend is a polynomial of degree 3 or lower.

Kim and Phillips (2000) develop an alternative modification of the LP estimator, called the "modified LP estimator", which is valid for $d<2$ and is invariant to the presence of a linear trend. It is asymptotically more efficient than the modification proposed by Velasco (1999a).

Velasco (1999b) also proposes tapering the original data to allow for non-stationary values of d and the presence of a polynomial trend in the data when using the local Whittle estimator. Shimotsu and Phillips (2000) develop a modified local Whittle estimator, which is invariant to the presence of a linear trend and is valid for d up to 2. Moreover, it is asymptotically more efficient that the LP-type estimators. However, for empirical purposes it has to be borne in mind that the original, unmodified, local Whittle estimator has better properties than this modified version if $d<0.5$ (see Shimotsu and Phillips, 2000).

These new methods therefore permit direct estimation of the memory parameter of series as well as testing whether $d<0.5$, $0.5 \leq d<1$ or $d \geq 1$ without previous transformation.

However, it is often important to determine whether a fractional integrated series contains a deterministic trend or not, and which is the sign of this trend. As we will see, this is an important element in testing convergence that has however been ignored so far in the studies of convergence carried out in a fractional-integration setting. Results from Marmol and Velasco (2002) are very useful for this objective. For $0.5<d<1.5$, these authors demonstrate that the OLS estimation of the trend is invalid and propose an alternative estimation of the trend coefficient in the frequency domain. They also indicate how to compute the variance of this estimation, so that a t-ratio can be calculated for testing the significance of the trend. They provide formulas for computing the critical value of this t-ratio, as a function of d [1]. When d is unknown, as is usually the case, an estimated value can be used. The authors suggest the estimators proposed by Velasco (1999a, 1999b). However the estimator proposed by Shimotsu and Phillips (2000) is also robust to the presence of a linear trend and to d up to 2, with the additional advantage of a lower asymptotic variance. Moreover this estimator does not require previously tapering the data, which is an advantage for our sample, as we will see below.

3. Growth convergence

3.1. Growth convergence in the existing literature: definitions, results and limitations

In the economic growth literature, there exist basically two opposed strands of thought. On the one hand, the neoclassical growth model (Solow, 1956) predicts that each country per capita real income converges to a steady state, regardless of the initial level of per capita income. If the steady state is

[1] These formulas are reproduced in the Appendix, together with a description of how to estimate the trend and how to compute the t-ratio.

common for all countries, they will all converge to the same level of per capita income. This is known as the “absolute convergence” hypothesis. If the steady states are country-specific, they are nonetheless parallel and the per capita incomes will end up differing between countries by a constant quantity. This situation is known as “conditional convergence”. On the other hand, in the more recent endogenous growth models (Romer, 1986, Lucas, 1988 and the subsequent literature) there is no force that pushes the per capita income to a specific level: per capita incomes of different economies may diverge and nothing guarantees that the income difference between poor and rich countries will not become unbounded.

Whether per capita income converge or not is therefore an important economic issue with substantial policy implications. If there is no automatic mechanism that ensures the convergence of economies over time, it is not only justified but also probably ethically necessary and "politically correct" to implement public policies aimed at helping the poorer (poorer regions or poorer countries) to catch-up as fast as possible with the richer. Knowing whether it is justified to dedicate public funds to these policies is therefore an important matter both for national and supra-national governments. This explains why there have been many attempts in the literature to develop methods of testing economic convergence.

One of the most well-known convergence test is based on the so called “β-convergence regression” associated to the works of Barro and Sala-i-Martín (1991, 1992). It consists of a cross-country regression of the average growth rate over some long enough time period of per-capita output (in log) on the level of the log of per-capita output at the beginning of the period and on a set of country-specific additional variables [2]. The typical β-convergence regression (without country-specific regressors) is as follows:

$$g_{i,T} = \alpha + \beta y_{i,0} + \varepsilon_{i,T} \qquad i = 1, \cdots, N$$

where $y_{i,t}$ is the log per capita output of economy i in year t, $g_{i,T} = T^{-1}(y_{i,T} - y_{i,0})$ is the average growth rate of economy i during the time span from year 0 to year T and T is a fixed horizon. For convergence, the coefficient of per-capita output in the initial year (β) should be negative, since it would reflect that the poorer the country in the initial year, the faster it grew over the period in order to catch up with the rich. The empirical tests made in cross-section samples generally delivered results favourable to convergence, at least when country-specific regressors are added (which corresponds to "conditional convergence").

This cross-section approach however has been criticized for various reasons. To cite only a few, Quah (1993) shows how it is possible for a negative β to be compatible with a non-decreasing

[2] The *logarithm* of output per capita is usually the variable on which the empirical work is carried out. So, unless otherwise said, in what follows , the expressions “per capita output” , “per capita income” , as well as “output” and “income” all refer to the logarithm of the per capita variable.

cross-section variance in output levels. Evans (1996,1997,1998) shows that the β-convergence regression provides invalid inference and conclusions on β, unless per capita output of all the countries of the sample follow the same AR(1) process. Obviously this is a very improbable condition. Evans and Karras (1996) propose an alternative methodology derived from panel data unit root tests, which they show to be valid under milder conditions. Bernard and Durlauf (1996) show that a negative $\hat{\beta}$ is compatible with the endogenous growth model; they also argue that the β-convergence test is ill-designed to analyse data where some countries are converging and other are not. In their opinion, convergence tests based on time series samples are more useful than cross-section, as they contain the right information to distinguish between the countries who converge to each other from those who diverge. These data can also distinguish between different types of convergence.

As a result, in the time series context, various authors have proposed and applied different operational definitions of convergence based on the per capita output differential. Let us represent the output differential of country i with respect to country j in period t as:

$$dif_{ij,t} = y_{it} - y_{jt} \;, i = 1,2,...N,\; i \neq j$$

More or less explicitly, all the existing definitions take for granted that y_{kt} is I(1) .The definitions that have been most used are [3]:

- "<u>Stochastic convergence</u>" (Carlino and Mills, 1993) and "<u>catching-up</u>"(Bernard and Durlauf, 1996):

  $$dif_{ij,t} = \alpha + \delta t + u_{ij,t} \;, i = 1,2,...N,\; i \neq j \text{ with } u_{ij,t} \text{ zero-mean I(0)}$$

  two countries exhibit *stochastic convergence* if their per-capita output differential is trend stationary. By itself, stochastic convergence is not a sufficient condition for real convergence. To make sense, stochastic convergence requires that the trend moves in the right direction, that is, that the absolute value of the differential decreases over time. This reduction of the absolute differential is what Bernard and Durlauf (1996) call *catching up convergence.*

- "Deterministic convergence" (Li and Papell, 1999):

  $$dif_{ij,t} = \alpha + u_{ij,t} \;, i = 1,2,...N,\; i \neq j \text{ with } u_{ij,t} \text{ zero-mean I(0)}$$

[3] Various names have been used in the literature to differentiate the definitions. Here we mainly use the nomenclature of Ericcson and Halket (2002), as we find it clearer than other ones.

two countries exhibit *deterministic convergence* if their per-capita output differential is stationary around a constant level. This definition is compatible with absolute convergence if this level is zero, whereas it corresponds to the concept of conditional convergence when this level differs from zero.

- "zero-mean convergence" (Bernard and Durlauf, 1996):

$$\text{dif}_{ij,t} = u_{ij,t} \;, i = 1,2,...N,\; i \neq j \text{ with } u_{ij,t} \text{ zero-mean } I(0)$$

two countries exhibit *zero-mean convergence* if their per-capita output differential is stationary around zero. This definition coincides with the concept of absolute convergence.

So, these definitions require the output differential to be I(0), either around a trend or a level. The convergence test would then consist of applying a unit root test on this differential, or checking that the output levels are cointegrated with the correct sign and values for the cointegration coefficients, and finally checking that deterministic components have the correct sign.

However, although the definition of convergence is in terms of the output *differential*, many empirical studies have worked on the *deviation* of the output per capita *from the group average*. In other words, according to the definition of convergence, the stochastic and deterministic characteristics to be studied are those of $\text{dif}_{ij,t}$ but instead of that, many studies focus rather on

$$\text{dev}_{it} = y_{it} - \bar{y}_t \;, i = 1,2,...N$$

where $\bar{y}_t$ is the cross-section arithmetic average of the output per capita of all the N countries in the group of interest, or some other measure of averaged output (see a.o. Carlino and Mills, 1993, Loewy and Papell, 1996, Strazicich, Lee and Day, 2001). However, it has to be noted that it is not innocuous to work with dev_{it} instead of dif_{it}. Let us imagine that among the N countries, country k diverges and all the other countries converge to each other (either with catching up or with deterministic or zero-mean convergence). Then $\text{dif}_{ij,t}$ should reflect convergence for all i and j *not equal to k*; by the same token, $\text{dif}_{ik,t}$ should reflect divergence for all i. So the "dif" variable contains the relevant information about the convergence characteristics of the group. On the other hand, the data of country k , which does not cointegrate with the other countries of the group, are included in all the "dev" series through the group mean value, so that *all* the dev_{it} for any i will contain the divergence characteristic of country k, and so *all* will be divergent. In other words, the inclusion in the computation of the common average of a non-converging country will mask the convergence process linking the other countries outputs. So, theoretically, either all the "dev" variables are I(0) if all countries converge, or none of them is I(0) if only one country diverges. As a first consequence, results of unit root tests on "dev" series in which convergence is detected for only a subgroup of the series should be interpreted with caution. As a second consequence, convergence tested on output differentials is at least more informative, if not more reliable.

The studies which centre the analysis on the output differentials normally apply unit root tests on the differentials or single-equation cointegration tests on the log level of outputs, seeking for a (1,-1) cointegration vector. Their results do not support convergence, unless some type of structural change in the deterministic component of the series is introduced in the testing procedure. In particular, Bernard and Durlauf (1995) obtain very little evidence of convergence between 15 OECD countries over the period 1900-1987. Oxley and Greasley (1995) analyse only three countries (Australia, USA and UK) and obtain catching-up convergence if they allow for a segmented trend in the catching-up process. More recently, Cellini and Scorcu (2000) study deterministic convergence [4] between the G-7 countries over the period 1900-1989, pair by pair; they therefore analyse 21 differentials. They detect only 6 cases of convergence over the whole period, provided a change in the intercept term is allowed. They express their surprise for not obtaining convergence in the second half of the time sample, in spite of the increased economic integration among the G-7 countries.

So, in general, the results in the time-series context using unit-root and cointegration tests do not constitute a strong support of convergence. This is however puzzling because a simple graphical observation of long term or even medium term per capita output statistics seem to indicate that most countries (at least the developed ones) have converged over time (for an illustration, see for instance Figure 1 in Section 4 of this paper).

Michelacci and Zaffaroni (2000) (MZ in the following) proposed an alternative approach in the time series context. These authors detect possible theoretical contradictions in the coexistence of I(1) outputs and β-convergence. They then show that there are theoretical reasons why the output levels series should be fractionally integrated and that mean-reverting fractional integration of output levels - i.e ouptut levels being FI(0.5<d<1)- is compatible with the empirical results of the β-convergence regression reported in the literature. They then redefine the β- convergence concept in terms of mean reversion of the deviations of output levels around a country-specific or a common trend, instead of defining it in terms of trend stationarity of output differentials. Specifically, they propose the following definition of (absence of) β- convergence: " an economy has no tendency to converge either towards its own or the common steady state if, after fitting either a country specific or a common (linear) trend respectively, the parameter of fractional integration d of the residuals is greater than or equal to 1 ($d \geq 1$). In the former case we say that there is no conditional convergence and in the latter that there is no unconditional convergence"[5]. They then test the integration order of the residuals of their definition, applying Geweke-Porter-Hudak (1983) Log Periodogram estimator on

[4] Although they call it "stochastic convergence"
[5] They also define "uniform convergence" when all the residuals have the same order of integration.

differences of previously detrended per capita output log levels and conclude that these residuals are indeed FID(0<*d*<*1),* which agrees with their new definition of convergence.

By offering theoretical and empirical evidence of fractional integration of output levels, MZ made an important contribution towards the solving of the puzzling results of no convergence usually obtained in the time series context. In particular, their results indicate that all the tests based on the assumption of the output being I(1) and searching for an I(0) differential or for traditional C(1,1) cointegration between output levels are invalid.

However, there are possible operational problems with their new definition of convergence. On the first place, if the series are actually fractionally integrated, it is not an easy task to guarantee an efficient estimation of the trend in order to obtain residuals on which the memory parameter can be reliably estimated (see Section 2, Sun and Phillips, 1999 and Marmol and Velasco, 2002 on this topic). This efficiency consideration is not addressed in MZ; they use for instance OLS estimation for the country specific trend, which is shown to be inefficient for $\mathrm{d} > 0.5$.

On the other hand, in order to estimate the memory parameter *d*, MZ use the GPH or LP estimator (see Section 2) which requires the data to be not only trend-free but also preferably integrated of order $d < 0.5$ and in any case $\mathrm{d} < 1$. To ensure this, the usual practice in the literature consists of first-differencing the series in order to obtain an estimation of (*d*-1) from data that are stationary (i.e. with a memory parameter lower than 0.5 and without trend). This is for instance what Cunado et al. (2002) and Silverberg and Verspagen (1999) do, when working with output levels and output differentials. However, Agiakloglou, Newbold and Wohar (1993) and Hurvich and Ray (1995) show that the estimation of *d* is not invariant to differencing, so that there might be bias due to overdifferencing. Moreover, MZ apply instead *fractional* differences: they estimate the memory parameter of $(1-L)^{1/2} z_t$ (where z_t are the residuals around a country-specific or common linear trend). This fractional-difference transformation necessarily implies a loss of data. Additionally, in the case where the common trend is used, this strategy does not guarantee that the common-trend fractionally differenced residuals are trend-free.

Finally, let us now suppose that the *common* linear trend has been adequately estimated, as well as the memory parameter *d* of the derived residual corresponding to, say, country k. And let us suppose that this estimated *d* falls significantly below 1, so that the deviation of output from the *common* trend is mean reverting. This implies that the deviation of country k from the *common* linear trend will return to its *own* mean in the long run. This mean may still be a trend, and there is a non-zero probability that this trend evolves in the wrong direction (from a graphical inspection of the data, this appears to be the case of Norway for example; see Figure 1 of their paper). Some additional test on the trend of the residual is therefore necessary to ensure that the mean-reversion of residuals is an indication of actual convergence (and not of divergence). In other words, obtaining mean-reversion for the residuals from a common trend is a necessary although not sufficient condition for convergence.

The tests carried out by Cunado et al. (2002) and Silverberg and Verspagen (1999) are also affected by the risk of overdifferencing and by testing a necessary but not sufficient condition for convergence: they work on the first differences of output and/or differentials and do not run any test on the possible trend of the differentials.

All these limitations could be avoided if fractional convergence is tested without detrending or differencing the output data. This is possible if the three following conditions are met:

a) convergence is defined as (mean-reverting, i.e. $d<1$,or stationary, i.e. . $d<0.5$) fractionally integrated output *differentials*, instead of as mean-reverting deviations from a trend.

b) more robust estimation techniques of the memory parameter are used. These techniques must be robust to the presence of both deterministic and stochastic non-stationarity (deterministic trend and /or $d \geq 0.5$), possibly present in the output differentials.

c) testing the significance of the trend coefficient of the output differential is carried out, with techniques that are robust to a non-stationary memory parameter ($d \geq 0.5$)

The robust estimation and testing procedures required for b) and c) are now available; their main characteristics have been described in Section 2 and they are explained with some more details in the Appendix. Extending the definition of convergence in a time-series context, in terms of the fractional-integration order and the deterministic characteristic of the output differential, as suggested in a), is therefore not only theoretically desirable but also empirically operational. We carry out his extension in the following subsection.

3.2. A new classification of convergence processes

In Table 1, we propose a new classification of the distinct types of convergence process through which different economies may be linked. This classification is based on the stochastic and deterministic characteristics of the output differentials, for reasons explained in Subsection 3.1. The type of convergence depends both on the value of the integration parameter *d* of the output differential and on the characteristic of the possible trend of this differential.

The strongest case of convergence takes place when the output differential is I(0) with zero-mean. This coincides with the "zero-mean convergence" of Bernard and Durlauf (1996). However, as long as the output differential has zero mean and is FI(*d*) with $d<0.5$, it is stationary around zero. In this case, convergence takes place, although at a lower speed than when $d=0$. An even slower case of convergence is when $0.5 \leq d<1$: the output differential is non-stationary although mean-reverting. When the mean of the output differential is a (not necessarily zero) constant, we stand in "deterministic convergence" as long as d<1; again, the lower d, the faster the convergence. "Catching-up" takes place if d<1 and the trend of the output differential is such that this differential tends to zero; here also, the lower d, the earlier convergence will be reached. If d<1 but the trend in the output differential pushes its value to ∞, the outputs will diverge. We call this "deterministic divergence". Finally, whatever the deterministic component of the output differential, divergence takes place as soon as $d \geq 1$. We call this "stochastic divergence".

Table 1: A new classification of convergence processes as a function of the behaviour of the output differential:

mean / *d*	*Zero mean*	*Constant mean*	*"decreasing"(a) trend*	"increasing"(a) trend
0	Strict *zero-mean convergence*	Strict *deterministic convergence*	Strict ***Catching-up***	Deterministic Divergence
0<d<0.5	Long-memory stationary *zero-mean convergence*	Long-memory stationary *deterministic convergence*	Long-memory stationary ***Catching-up***	
0.5 ≤ d<1	Long-memory mean-reverting *zero-mean convergence*	Long-memory mean-reverting *deterministic convergence*	Long-memory mean-reverting ***Catching-up***	
d≥1	Stochastic divergence			

(a) "decreasing" ("increasing") refers to the fact the trend coefficients have the right (wrong) signs in order to push the differential from its initial level to 0.

3.3. A new empirical strategy

The empirical strategy that we propose emerges naturally from our previous comments on how to measure growth convergence (subsection 3.1 and 3.2, and especially Table 1) and from the new techniques described in Section 2 for the estimation and testing of the fractional order of integration and trend of possibly non-stationary processes.

This strategy consists of:

1) Choose a benchmark country towards which convergence of per-capita output of the countries of the sample will be tested.

2) Compute the differential of each country in the sample with respect to this benchmark country and apply a semi-parametric method of estimation of d that must be robust to stochastic nonstationarity $(d \geq 0.5)$ and to deterministic trend. See section 2 for a description of such estimators.

3) Use the distribution results available for these estimators in order to perform the following tests:

(a) $H_0^{(1)} : d = 1$ against $H_A^{(1)} : d < 1$

(b) $H_0^{(0)} : d = 0$ against $H_A^{(0)} : d > 0$

(c) $H_0^{(1/2)} : d = 0.5$ against $H_A^{(1/2)} : d \neq 0.5$

Tests (b) and (c) have to be carried out on the level of the differential, whereas (a) can be performed both on levels and on first differences (in the latter case, the hypotheses to be tested are in fact $H_0^{(1)} : (d-1) = 0$ against $H_A^{(1)} : (d-1) < 0$, where *(d*-1*)* is the memory parameter of the first difference of the differential).

Depending on the results of the tests, different conclusions emerge. In the following, the notation "$RH_o^{(k)}$" means that the hypothesis "$H_o^{(k)} : d = \mathrm{k}$ " is rejected in favour of the alternative specified in the hypothesis testing (a), (b) or (c) where "$H_o^{(k)}$" appears as the null. On the opposite, the notation "$notRH_o^{(k)}$" means that the null "$H_o^{(k)}$" is not rejected in that test.

- If $RH_0^{(1)}$ and $notRH_0^{(0)} \Rightarrow$ *d<1* and *d* not >0:

strict zero-mean convergence is present if the mean of the differential is not significantly different from 0; strict catching-up is taking place if the differential is trending in the right direction; finally, strict deterministic convergence is in place if the differential is not trending but its mean differs significantly from zero. Traditional t-tests may be used to test these different possibilities for the trend.

- If $RH_0^{(1)}$ and $RH_0^{(0)} \Rightarrow$ *0<d<1*

the output differential is fractional and some type of convergence may be taking place. In this case, additional testing that accounts for the possibly non-stationary characteristic of the differential is necessary to nuance this conclusion:

- If $\hat{d} < 0.5$ and $RH_0^{(1/2)}$, this fractional differential is stationary and fractional stationary deterministic convergence exists if the mean of the differential is not trending.

- If $\hat{d} > 0.5$, the fractional differential is not stationary though mean reverting. Then we can estimate and test the significance of the trend coefficient of the differential, using Marmol and Velasco (2002) results, and in particular the t-ratio on the trend coefficient that they propose and their formulas for computing the one-sided critical values (see Section 2 and formulas (10), (11) and (12) of the Appendix). Fractional mean-reverting deterministic convergence is taking place if the trend of the differential is not significant. If this coefficient is significant and has the correct sign according to the level of the differential at the beginning of the period, fractional mean-reverting catching up is in place. If it has the wrong sign, deterministic divergence is operating.

➢ If $notRH_0^{(1)}$ and $RH_0^{(0)}$ $\Rightarrow$ *d* not <1

the output differential is I(1) or FID(*d*>1). These are cases of stochastic divergence. If $\hat{d} > 1$, it may be interesting to check whether it is significantly higher than 1, since this would reflect a higher grade of divergence than if *d*=1.

➢ Finally, if $notRH_0^{(1)}$ and $notRH_0^{(0)}$, it is impossible to discriminate between the two extreme cases of *d=0* and *d=1*, so that more sample information is required.

4. Empirical results

We have applied our empirical strategy on internationally comparable per-capita GDP data of different countries, provided by Madison (1995) and Madison (2001). Put together, these two sources provide annual data expressed in 1990 international (PPP) Geary-Khamis dollars, over a period that extends from 1870 to 1998. Various empirical studies of convergence have been carried out on this source or an earlier version of it (see a.o. Oxley and Greasley, 1995, Li and Papell, 1999, Michelacci and Zaffaroni, 2000, Cunado et al., 2001, Cellini and Scorcu, 2000)

The group of countries on which we apply this strategy is the G-7 group (USA, Canada, France, Germany, Italy, Japan and United Kingdom) plus Australia and New Zealand. These G-7 data coincide with one of the sample used by Ericsson and Halket (2002) who obtained rejection of convergence when applying unit root tests on output differentials and results that are somewhat more favourable to convergence when multivariate Johansen cointegration tests are applied, although these last results might be spuriously favourable in the light of the paper of Gonzalo and Lee (2000), if the output per capita are actually fractionally integrated instead of I(1). The G-7 is also the group of countries analysed by Cellini and Scorcu (2000) who obtain timid results of convergence with traditional I(1)/I(0) cointegration tests when endogenous structural breaks are allowed in the data. However, even with structural breaks, they do not detect convergence at all in the post-WWII period, which they consider as "astonishing" given "the increased economic integration among the G-7 countries (with the associated evidence of β-convergence)" (Cellini and Scorcu, 2000, p. 464).

In this section, we show among other things that the output differentials are in fact fractionally integrated with $d>0$, instead of I(0). This should shed some light on the strange results of Cellini and Scorcu, should explain why Ericsson and Halket(2002) do not obtain I(0) differentials or residuals when applying unit root tests or single-equation traditional cointegration tests, and put in doubt the statistical validity of their multivariate Johansen cointegration results, which might be spurious.

4.1. Results

4.1.1. The data

Figure 1 [6] presents the logarithm of real per capita output for the G-7 countries extracted from Maddison(1995, 2001). It this graph, the level of the US log ouptut is represented by the upper edge of the red-shaded area. The other countries' level of output can then be visually compared with this benchmark country[7]. At first glance, it is obvious that these 7 countries somehow converged over the last century. It is also worthwhile noticing the huge and longer-lasting difference in output that took place during the Second World War between the US and Japan and Italy, as well as the very sudden and huge increase in the differential with Germany by the end of this war. As we will see, this will have an impact on the results.

4.1.2. Results from unit- root tests

Before applying and commenting the results obtained by the new empirical strategy that we propose, we first reproduce in Table 2 the results obtained from the application of unit-root tests on the output differentials of each country with respect to the US. We apply six different tests, aimed at determining whether the differentials are compatible with convergence (in case they are not statistically different from an I(0) series) or indicate divergence (if the test results is that they are I(1) series). The tests that we apply are the following: the traditional ADF test, the Phillips-Perron test, the KPSS stationarity test, and the more recent DF-GLS and ERS test due to Elliott et al. (1996) as well as the Ng-Perron modified tests. In all the cases, a constant and a trend are included in the testing equation. The results of all these tests are far from supportive of convergence: for most countries, the result is that the series of differential is an I(1) series.

[6] The Figures and Tables of this section 4 are reproduced at the end of the paper, after the conclusions and before the Appendix.

[7] Unless otherwise said, the US will serve as the benchmark country both for economic reasons and because this country is traditionally used as a benchmark in most convergence studies.

4.1.3. Estimation and testing of the memory parameter of the output differentials

In order to determine the order of integration of the output differentials of the nine countries, we have to choose an estimation method that is robust to values of d >0.5 and to deterministic trends (see Sections 2 and 3 for a justification). One of the possibilities described in Section 2 consists of using an estimator based on tapered data as proposed by Velasco (1999a, 1999b). Another possibility consists of using one of the Modified estimators due to Kim and Phillips (2000) or to Shimotsu and Phillips (2000).

Tapering the data consists of weighting the original data before estimating d , according to a bell-shaped weighting scheme which takes maximum values in the central dates of the sample and symmetrically decays to zero from the central dates towards the years at the beginning and at the end of the sample. In other words, the tapered estimator is based on data in which the central dates play a role in the estimation that is much more important than the observations at the beginning or at the end of the observation period.

However, the countries have been involved in the Second World War, with impacts of different nature and intensity on their economies during the war years and just after it. For the USA, and to a lower extent for some other countries such as Canada or the United Kingdom, the war led to a major boom, whereas it led to an important reduction in GDP for the European countries directly implicated in the conflict. As a result, the differential of any country of the boom group with respect to any country of the reduction group takes abnormally high (absolute) values at the very end of the 30's and during the first half of the 40's. These years stand in the middle of our sample. This prevents us from applying the robust tapering methods advocated by Velasco (1999 a, b). In our sample, this tapering would indeed give maximum weight to the abnormal WW2 data and ever decreasing weight to more normal data further away from this major conflict. We checked by simulation that, as expected, this produces an overestimation of the memory parameter d that can be substantial, especially when the true value of d is low.

For this reason, we opted for not tapering the data. In order to use a semi-parametric estimator robust to non-stationarity and deterministic trend, we use here the Shimotsu and Phillips (2000) Modified Local Whittle estimator (denoted $\hat{d}_{m,MLW}$ in the following and in the Appendix), which is asymptotically more efficient than the also robust Modified Log Periodogram estimator developed by Kim and Phillips (2000) (see again Section 2 and the Appendix for more details)

Table 3 reproduces the results of estimating and testing the order of integration of the output differential of each of the remaining countries with respect to the US. In this table, four rows of results are presented for each country. The first two rows, identified by "(d1)" and "(d2)" refer to the results obtained when d is estimated on the level of the differential, using the Modified Local Whittle estimator of Shimotsu and Phillips (2000). The only difference between the two rows refers to the number m of frequencies that are used in the computations of the estimator. In the row identified as "(d1)", m=int($n^{0.5}$) frequencies are used, where n is the sample size and int(x) stands for the integer part of x. In the second row, identified as "(d2)", m=int($n^{0.6}$) is used [8]. According to our empirical strategy, for these estimations, all the tests described in columns (3) , (4) and (5) are carried out. The third and fourth rows, identified as "(Δ1)" and "(Δ2)" correspond to the estimation of *d* as one plus the estimated memory parameter of the first difference of the differential. This first difference is computed to bring the series to the stationarity zone, with the only purpose of obtaining an alternative test of $H_0^{(1)}$. The estimator in this case is the *unmodified* Local Whittle estimator, because the first differences of differentials are very likely to have a low value of d and no trend. The unmodified estimator indeed dominates the Modified for stationary series (see Shimotsu and Phillips, 2000). The only difference between the third and the fourth rows is the value of m (m = int($n^{0.5}$) and m = int($n^{0.6}$), respectively).

In this table we have highlighted in bold red the rejections of d=0 against d> 0 and of d =1 against d<1. The first important result of this table is the very strong rejection of d=0 in favour of positive *d.* for all countries and all estimated values of this parameter. The differentials are therefore unambiguously not I(0), which may greatly explain the non-convergence results obtained in the framework of the I(1)-I(0) dichotomy. In particular and as mentioned above, this would explain the no-convergence results of Ericsson and Halket (2002) and of Cellini and Scorcu(2000); also, in combination with Gonzalo and Lee(2000) results, they cast some doubt on the convergence results obtained by Ericsson and Halket (2002) with trivariate Johansen cointegration tests.

As far as the opposite hypothesis is concerned, namely d =1 against d<1, a unit root is unambiguously rejected in the cases of France and the United Kingdom, as well as in the case of Canada when *d* is estimated from the first differences of the differential. The results for Germany and Italy are more in favour of the existence of a unit root, although mean reversion is accepted for one of the four alternative estimations, and is rejected on the boundary of the 10% level of probability in another case. Japan is much more extreme: values of *d* larger than one cannot be excluded. On the other hand, it is also noticeable that for none country, the hypothesis of $d<0.5$ has been accepted, and for most of those with *d* significantly below 1, *d* is also significantly above 0.5. So, if convergence takes place, it is mean reverting but not stationary. This implies that convergence is slow.

[8] The number "m" of frequencies is the number of points, or "observations", in the frequency domain that are used to run the estimation of *d* (see the appendix for more details). The value m=$n^{0.5}$ has been traditionally used, as a result of the recommendation of Geweke and Porter-Hudak(1983) for the Log Periodogram regression. However, according to Hurvich and Deo(1998), it seems to be that the optimal

Given the particular situation of Germany, Japan and Italy during and after the Second World War, and in particular given their relative output position with respect to the US (remember Figure 1 and comments thereon), it seems reasonable to check whether the choice of US as the benchmark country may influence the results. In Table 4, we repeat the analysis using Germany as the benchmark country, as justified by its economic position in Europe.

In Table 4, Italy and the UK offer unambiguous symptoms of convergence with Germany. The first two converge with mean-reversion (i.e. $d \geq 0.5$) whereas the French differential is I(0) or stationary, depending on the value of m .[9] Canada stands again in an intermediate position: convergence takes place for half of the estimations of *d,* whereas *d*=1 according to the other half.

Japan is the only country that still exhibits symptoms of divergence, in spite of what could be inferred from a visual inspection of the data. In our opinion, the memory parameter for this country is overestimated. For a better understanding of the argument that we are going to develop, we reproduce in Figure 2 the differential of Japan with respect to the USA and with respect to Germany.

As can be seen in this graph, these differentials are decreasing and progressively reaching a near-zero value, which is visually compatible with catching-up. However, these trends are not linear: there are economic and graphical reasons to believe that the trend changed at least in the late forties-early fifties. The Modified local Whittle estimator that we have been using so far is robust to a linear trend, but not to a polynomial trend of degree higher than 1, or to structural changes in an otherwise linear trend. The tapering methods of Velasco (1999 a) are able to take care of a polynomial trend that involves term up to t^3 . However, as we mentioned above, they are extremely sensitive to abnormal data in the middle of the sample. Such abnormal data are present in the case of the output of Japan with respect to the US; However, since the Germany-Japan differential does not suffer from abnormal values, and it seems that a third degree trend would be able to capture adequately the downward trending evolution of this differential, we have applied a Parzen-taper to this differential with the following results (see Velasco, 1999,a for details). With 11 periodogram points in the regression, $\hat{d}$ = 0.3415 (significance value: 0.07); with 16 periodogram points, $\hat{d}$ = 0.4685 (significance value:0.035). These estimations are much lower than the estimators used in Table 3 and 4 and therefore they may indicate a positive bias in the estimation of *d* for Japan in Table 3 and 4 due to a non-linear trend. In other words, it is probable that Japan is converging more than what could be apparent from the results of Tables 3 and 4.

choice for m is in general slightly higher for these regressions. In absence of other explicit information, we have used these indications for fixing the values of m in our estimates.

[9] Given the very low value of $\hat{d}$ for the French differential with respect to Germany, there is neither need nor justification for getting the estimated value of this parameter from the first difference of the differential. This is why the third and fourth rows of Table 2 for this country are empty.

4.1.4. Deterministic convergence or catching up?

In the preceding section, we have obtained that most differentials are FI(*d*) with $0.5<d<1$. So, since $d>0.5$, it is not surprising that the simple I(0)- versus- I(1) tests yield results that mainly favour the I(1). However, the conclusion of divergence is erroneous, since in fact $d<1$; so, most economies converge, although at a slower rate that under traditional I(0) convergence. The next question is now to test the significance of the trend coefficient, in order to determine whether the data present deterministic convergence or catching-up (or even deterministic divergence if the trend evolves in the wrong direction). For that purpose, the estimation and testing procedures of Marmol and Velasco (2002) - which are valid for $d>0.5$ - are applied[10] (see Section 3 and formulas (10), (11) and (12) in the Appendix). For each case, from the alternative estimates of *d*, we have chosen that one closest to 0.5. The results, collected in Table 5, indicate on the first place that there is no case of divergence caused by a wrongly directed trend: all the signs of the significant trends are correct. On the other hand, these results indicate that what tends to dominate is catching-up. Only France as well as Australia and the US on one part, and Canada and Germany on the other exhibit deterministic convergence.

5. Summary, conclusions and possible extensions

In this paper, we have examined output convergence is in the framework of fractional integration. We have suggested new definitions of convergence in terms of the output differential; these new definitions distinguish between the case where this output differential is I(0), is fractionally integrated of order $0<d<0.5$, or is fractionally integrated of order $0.5\le d<1$. They also differentiate catching-up convergence, when convergence is in fact under way, from the cases of "zero-mean" or deterministic convergence. In order to test which type of convergence has characterised the G-7, Australia and New Zealand over the last century, we have used recently developed techniques which offer the great advantage of being robust to both deterministic and stochastic non-stationarity (linear trend and/or memory parameter $d\ge 0.5$) and which are therefore particularly adequate for our problem. Based on these new techniques, a new testing strategy is proposed which is able to differentiate catching-up form deterministic convergence, and which is operational even if $d > 0.5$. This strategy avoids some limitations of other convergence studies carried out in a fractional setting and completes the characterization of the convergence or divergence process.

Applying this new strategy, we obtain, that, except for Japan, all the other G-7 countries, Australia and New Zealand show rather strong signs of either catching-up or deterministic convergence. Except in one particular case, convergence is not I(0): the memory parameter is generally significantly higher than 0.5 but lower than 1. This means that convergence did take place although the convergence process is slower than what would be implied by I(0) differentials.

[10] Except in the case of the differential of France against Germany, for which I(0) convergence was detected. In this case, standard OLS trend estimation and t-tests may be used.

This paper contributes therefore to solve the puzzling results of rather timid convergence usually obtained in the time series context even when structural change is introduced in the model, which are probably due to limiting the study to the extreme cases of I(0) convergence versus I(1) divergence.

Possible extensions of this study, partly under progress, consists of applying the same empirical strategy to other developed and less developed countries, and of testing whether the results are sensible to the estimator that we have used (applying for instance the LP and MLP estimators mentioned in Section 3).

On the other hand, the univariate approach developed here for output convergence can also be applied on other types of convergence analysis, such as price convergence between regions, interest rate convergence between countries of an economic integration zone or for Purchasing Power Parity.

Finally, an interesting complementary analysis would consist of testing whether some type of fractional cointegration links the per-capita output series of the different countries. This analysis could be carried out along the lines of Davidson (2002, 2003) bootstrap techniques for fractional cointegration.

Tables and figures of Section 4:

Table 2: I(1)-I(0) tests on the output differentials

Benchmark country: USA - **Output variable**: log of per capita GDP in 1990 International Geary-Khamis dollars

test / country	ADF	PP	ADF-GLS	KPSS	ERS	Ng-P	conclusion
Canada	-3.51 I(1) at 1%	-3.49 I(1) at 1%	-3.35 I(1) at 1%	0.07 I(0)	4.97 I(1) at 1%	I(1) at 1%	**I(1) at 1% I(0) at 5%**
France	-3.79 I(1) at 1%	-2.43 I(1)	-3.33 I(1) at 1%	0.179 I(1) at 5%	4.21 I(1) at 1%	I(1)	**I(1) at 1%**
Germany	-2.27 I(1)	-2.79 I(1)	-2.20 I(1)	0.15 I(1) at 5%	10.04 I(1)	I(1)	**I(1)**
Italy	-2.95 I(1)	-2.52 I(1)	-2.43 I(1)	0.19 I(1) at 5%	8.45 I(1)	I(1)	**I(1)**
Japon	-1.97 I(1)	-1.93 I(1)	-1.59 I(1)	0.22 I(1)	19.63 I(1)	I(1)	**I(1)**
Reino Unido	-2.82 I(1)	-3.15 I(1)	-2.72 I(1)	0.18 I(1) at 5%	7.08 I(1)	I(1)	**I(1)**
Australia	-2.64 I(1)	-2.64 I(1)	-2.43 I(1)	0.28 I(1)	8.79 I(1)	I(1)	**I(1)**
New-Zealand	-3. I(0)	-4.10 I(0)	-3.93 I(0)	0.096 I(0)	3.74 I(0)	I(0)	**I(0)**

Table 3: Estimation and testing of the order of integration of output differentials

Benchmark country: USA - **Output variable**: log of per capita GDP in 1990 International Geary-Khamis dollars

Country	$\hat{d}$	(3) $H_0^{(0)}: d=0$ $H_A^{(0)}: d>0$	(4) $H_0^{(1)}: d=1$ $H_A^{(1)}: d<1$	(5) $H_0^{(1/2)}: d=0.5$ $H_{A,1}^{(1/2)}: d<0.5$ $H_{A,2}^{(1/2)}: d>0.5$	CONCLUSION	
		p-values				
Canada (d1)	1.054	0.0000	.6391	.0002 ⇒ d>0.5	d=1, not >1	
(d2)	.9305	0.0000	.2777	.0003 ⇒ d>0.5	d=1, not >1	
(Δ1)	.705	-	.0250	-	Mean reversion	
(Δ2)	.766	-	.0191	-	Mean reversion	
France (d1)	.554	0.0001	.0016	.3596 ⇒ d≅0.5	Mean reversion	
(d2)	.669	0.0000	.0025	.0075⇒ d>0.5 at 10%	Mean reversion	
(Δ1)	.480	-	.003	-	Mean reversion	
(Δ2)	.640	-	.0011	-	Mean reversion	
Germany (d1)	0.814	0.0000	.10913	.0185⇒ d>0.5	*d=1, not >1*	
(d2)	1.037	0.0000	.6257	.0000⇒ d>0.5	d=1, not >1	
(Δ1)	.82	-	.0317		Mean reversion	
(Δ2)	1.060	-	.6948		d=1, not >1	
Italy (d1)	.949	0.0000	.3684	.0029⇒ d>0.5	d=1, not >1	
(d2)	.849	0.0000	.1007	.0030⇒ d>0.5	*d=1, not >1*	
(Δ1)	.828	-	.1270	-	*d=1, not >1*	
(Δ2)	.783	-	.0269	-	Mean reversion	
Japan (d1)	1.186	0.0000	.8916	.0000 ⇒ d> 0.5	d=1, not >1	
(d2)	1.106	0.0000	.8154	.00000⇒ d> 0.5	d=1, not >1	
(Δ1)	1.332	-	.9862		d >1	
(Δ2)	1.189	-	.9443		d >1	
UK (d1)	.615	0.0000	.0053	.4456 ⇒ d≅0.5	Mean reversion	
(d2)	.744	0.0000	.0148	.0387⇒ d> 0.5	Mean reversion	
(Δ1)	.390	-	.0000		Mean reversion	
(Δ2)	.580	-	.002		Mean reversion	
Australia (d1)	.771	0.0000	.0645	.0720 ⇒ d≅0.5	Mean reversion	
(d2)	.910	0.0000	.2225	.0005⇒ d> 0.5	d=1, not >1	
(Δ1)	.582	-	.0028		Mean reversion	
(Δ2)	.678	-	.0031		Mean reversion	
NZealand d1)	.702	0.0000	.0239	.1806 ⇒ d≅0.5	Mean reversion	
(d2)	.901	0.0000	.1999	.0007⇒ d> 0.5	d=1, not >1	
(Δ1)	.2474	-	.0000		Mean reversion	
(Δ2)	.3310	-	.0000		Mean reversion	

Table 4: Estimation and testing of the order of integration of output differentials

Benchmark country: Germany - **Output variable**: log of per capita GDP in 1990 International Geary-Khamis dollars

Country		$\hat{d}$	(3) $H_0^{(0)}: d = 0$ $H_A^{(0)}: d > 0$	(4) $H_0^{(1)}: d = 1$ $H_A^{(1)}: d < 1$	(5) $H_0^{(1/2)}: d = 0.5$ $H_{A,1}^{(1/2)}: d < 0.5$ $H_{A,2}^{(1/2)}: d > 0.5$	CONCLUSION	
			p-values				
Canada	(d1)	.751	**0.0000**	**0.0489**	.0002 ⇒ d>0.5	**Mean reversion**	
	(d2)	.955	**0.0000**	.3503	.0003 ⇒ d>0.5	d=1, not >1	
	(Δ1)	.725	-	**.0336**	-	**Mean reversion**	
	(Δ2)	.962	-	.3727	-	d=1, not >1	
France	(d1)	.012	**0.4678**	**.0000**	.0000 ⇒ d< 0.5	**I(0)**	
	(d2)	.285	**0.0079**	**.0000**	.034⇒ d<0.5	**Stationarity**	
	(Δ1)	-	-	-	-	-	
	(Δ2)	-	-	-		-	
Italy	(d1)	.659	**0.0000**	**.0118**	.1462⇒ d≅0.5	**Mean reversion**	
	(d2)	.618	**0.0000**	**.0000**	.1590⇒ d≅0.5	**Mean reversion**	
	(Δ1)	.499	-	**.0004**	-	**Mean reversion**	
	(Δ2)	.511	-	**.0000**	-	**Mean reversion**	
Japan	(d1)	.9803	**0.0000**	.4479	.00007⇒ d> 0.5	d=1, not >1	
	(d2)	.9247	**0.0000**	.2615	.00002⇒ d> 0.5	d=1, not >1	
	(Δ1)	.9163	-	.2895		d=1, not >1	
	(Δ2)	.8533	-	.1065		Mean revers ~10%	
UK	(d1)	.7954	**0.0000**	**.0873**	.0250 ⇒ d>0.5	**Mean revers. 10%**	
	(d2)	.8501	**0.0000**	.1016	.0030⇒ d> 0.5	Mean revers ~10%	
	(Δ1)	.433	-	**.0000**		**Mean reversion**	
	(Δ2)	.641	-	**.0010**		**Mean reversion**	

Table 5: Catching-up or deterministic convergence?

Country	Benchmark country: USA				Benchmark country: Germany			
	Trend coefficient	t_m test	1-sided 5% c.v.	CONCLUSION	Trend coefficient	t_m test	2-sided 5% c.v.	CONCLUSION
Canada	-0.0025 correct sign	-6.2344	-4.9422	Mean-reverting catching up	-.0034 correct sign	-2.1452	-5.09432	Mean-reverting deterministic cvgce
France	-0.0014 correct sign	-2.3064	-4.0234	Mean-reverting deterministic cvgce	-.0013 correct sign	-3.0349	-1.645	*If I(0) convergence, strict catching-up*
Germany	0.0009	0.5210	5.9138	*If convergence, deterministic convergence*	-	-	-	-
Italy	-0.0028	-1.1623	-5.5756	*If convergence, deterministic convergence*	-0.0036 correct sign	-4.2781	-3.8355	Mean-reverting catching up
United Kingdom	.0055 correct sign	6.4042	4.1524	Mean-reverting catching up	0.0046 correct sign	4.9931	4.5032	Mean-reverting catching up
Australia	.0061	3.8558	4.1630	Mean-reverting deterministic cvgce	n.a.	n.a.	n.a.	n.a.
New-Zealand	.0059	10.791	4.920	Mean-reverting catching up	n.a.	n.a.	n.a.	n.a.

Figure1: real per capita output (log) in nine developed countries

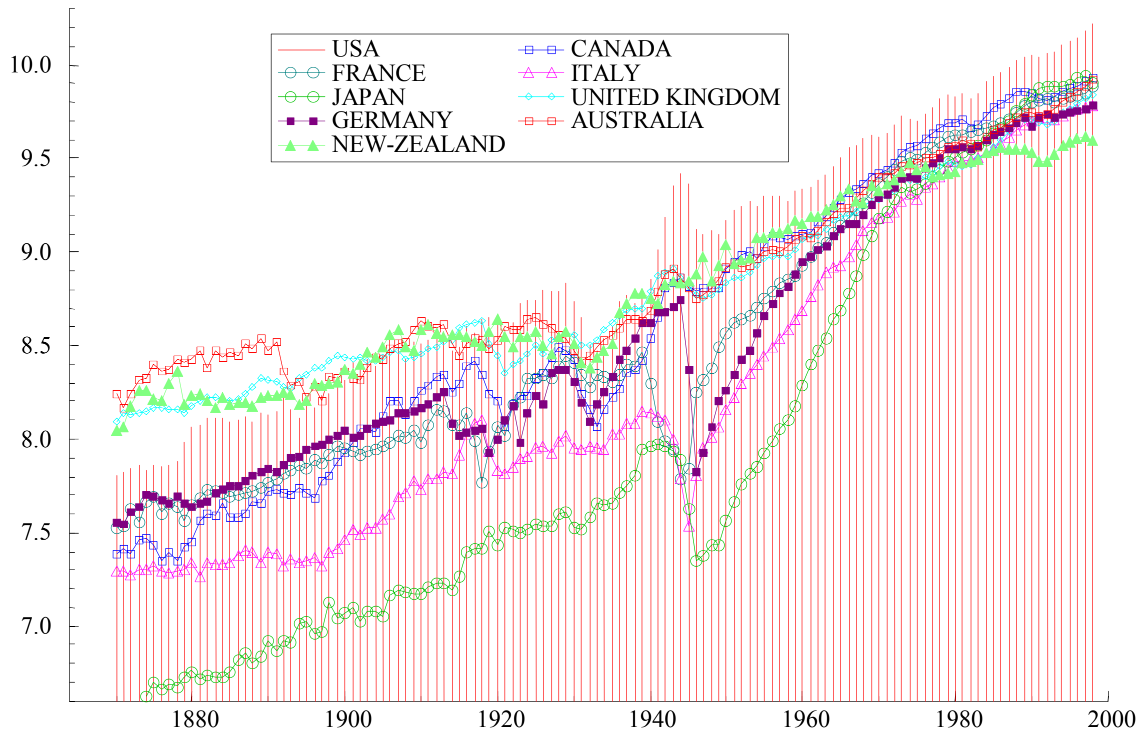
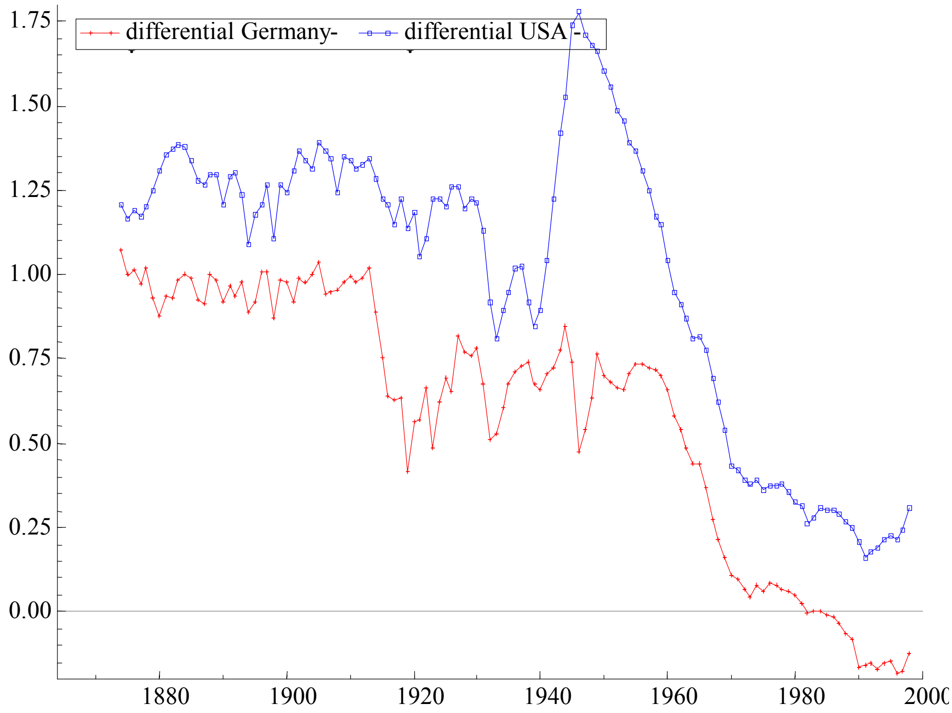


Figure 2: Differential of Japan

log(per capita GDP of USA)-log (per-capita GDP of Japan)

log(per capita GDP of Germany)-log (per-capita GDP of Japan)



Appendix:

In this Appendix, after a definition of fractional integration valid for both the stationary and the non-stationary intervals, we focus on the technical aspects of the recently developed estimation and testing methods that we consider are more suitable than previous techniques for their application on convergence problems, and that are especially useful for the type of convergence that we have defined in this paper.

Following Shimotsu and Phillips (2000), who developed the estimator of the memory parameter used in this paper, let us consider the fractional process X_t generated by the model

$$(1-L)^d (X_t - X_0) = u_t \quad t = 0,1,2,\cdots \qquad (1)$$

where X_0 is a random variable with a certain fixed distribution, u_t is a zero-mean stationary process with continuous spectrum $f_u(\lambda) > 0$, and $u_t = 0$, $t \le 0$, d is not necessarily integer and the fractional difference operator is defined as:

$$(1-L)^d a_t = \sum_{k=0}^{t} \frac{\Gamma(-d+k)}{\Gamma(-d)} a_{t-k}$$

where $\Gamma(\cdot)$ is the gamma function.

Depending on the value of *d*, the series exhibits the following properties. If $|d| < 0.5$, X_t is stationary and invertible. For $d \ge 0.5$, X_t is non stationary. If d <1, X_t is mean-reverting, exhibiting transitory memory. For $d \ge 1$, the memory is permanent and there is no mean-reversion.

Inverting (1), an alternate form for X_t is:

$$X_t = (1-L)^{-d} u_t + X_0 = \sum_{k=0}^{t-1} \frac{\Gamma(d+k)}{\Gamma(d)} u_{t-k}$$

Note that for d integer and positive, model (1) expresses the behaviour of X_t in terms of first of higher order differences. Note also that u_t is not necessarily serially uncorrelated, so that short-run dynamics -additional to long-range dependence- is not excluded in X_t. Finally, note that when u_t follows and ARMA(p,q) process, X_t is called an ARFIMA(p,d,q) process.

Mainly during the last decade, different methods have been developed to estimate the so-called "memory parameter" *d*. Very often, the main empirical interest lies exclusively on the value of this parameter. So instead of developing and using parametric methods tending at estimating all the parameters of the process, various authors focused on a semi-parametric estimation of *d* alone, ignoring the short-run parameters. Until very recently, these semi-parametric methods were restricted to the stationarity cases ($|d| < 0.5$), which required to previously difference the series when non-stationarity was suspected[11] .

An important property of stationary fractional series on which some of these semi-parametric methods are based is:

$$f(\lambda) \approx G\lambda^{-2d} \text{ as } \lambda \to 0^+ \qquad (2)$$

where $f(\lambda)$ is the spectral density of the series and $0 < G < \infty$.

so that

$$log(f(\lambda)) \approx k + d(-2\,log\,\lambda) \qquad (3)$$

for small frequencies.

Therefore, defining the Discrete Fourier Transform of a series a_t for t=1,2,...,n as

$$w_a(\lambda_j) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} a_t e^{it\lambda_j}$$

where $\lambda_j = \frac{2\pi j}{n}$ are the Fourier frequencies, we may compute the periodogram of the series X_t as

$$I_x(\lambda_j) = w_x(\lambda_j) w_x(\lambda_j)^*$$

Then property (3) suggests to regress $log\lfloor I_x(\lambda_j) \rfloor$ on a constant and $-2\,log\,\lambda_j$, for $j = l, l+1, \cdots, m$, with $l \geq 1$ and $m < n$. This is the basis for the so-called "log-periodogram (LP) regression" or "GPH regression" (Geweke and Porter-Hudak ,1983 , Robinson,1995a). It is an appealing method because it is rather simple to implement. This explains why it has been widely used.

Robinson (1995a) showed the consistency and asymptotic normality of a version of this estimate for stationary and invertible processes, i.e. for -0.5<d<0. Velasco (1999a) studied the asymptotic properties of the LP estimator for $d \geq 0.5$, showing that it is asymptotically Normal for

---

[11] Note that if X_t is FI(d_0+θ), with d_0 integer and $|\theta| < 0.5$, $(1-L)^{d_0} X_t$ is FI(θ) and therfore stationary.

$d \in [0.5, 0.75)$ and consistent for $d < 1$. This author also advocates data tapering together with a slight modification of the LP estimator which guarantee consistency and asymptotic normality of this log periodogram regression for any value of *d*; moreover, with adequate tapering, the estimate would be invariant to the presence of a deterministic polynomial trend, without any need of previous estimation of this trend. Tapering the data consists of applying weights to them before estimating the periodogram; the weights are ≤ 1, and decay symmetrically from the centre observation of the sample towards the extreme. The tapered estimator, however, has a higher variance than the non-tapered periodogram Velasco (1999a) also proposes different tapering schemes that are suitable for a value of *d* up to 3.75 and a polynomial trend up to third degree. On the other hand, Kim and Phillips (2000) develop a slight modification of the LP regression derived from adding a term in the Discrete Fourier Transform of X_t, which therefore modifies the values of the periodogram. They show that the new estimator, based on this modified periodogram and called "Modified LP estimator" is consistent and asymptotically normal for d <2, as well as invariant to the presence of a linear trend. The asymptotic distribution of the LP estimator ($\hat{d}_{m,LP}$) and the asymptotic distribution of this Modified LP estimator ($\hat{d}_{m,MLP}$) are identical:

$$\sqrt{m}\left(\hat{d}_{m,LP} - d\right) \xrightarrow[n\to\infty]{} N\left(0, \frac{\pi^2}{24}\right)$$

and (4)

$$\sqrt{m}\left(\hat{d}_{m,MLP} - d\right) \xrightarrow[n\to\infty]{} N\left(0, \frac{\pi^2}{24}\right)$$

Another frequently used estimator is the local Whittle estimator suggested by Künsch (1987) and further studied and justified by Robinson (1995b). It is obtained by minimising the following objective function with respect to G and *d* (see Künsch, 1987, Robinson, 1995b) over a closed interval of admissible values for *d* [12]:

$$Q(G,d) = \frac{1}{m}\sum_{j=1}^{m}\left\{ log(G\lambda_j^{-2d}) + \frac{I_x(\lambda_j)}{G\lambda_j^{-2d}} \right\} \qquad (5)$$

Robinson(1995b) established the conditions under which this estimator of *d*, that we will denote $\hat{d}_{m,LW}$ is consistent and asymptotically normal, with asymptotic distribution given by:

$$\sqrt{m}\left(\hat{d}_{m,LW} - d\right) \xrightarrow[n\to\infty]{} N\left(0, \frac{1}{4}\right) \qquad (6)$$

[12] i.e. $G \in (0, \infty)$ and $d \in [\Delta_1, \Delta_2]$ with $-0.5 < \Delta_1 < \Delta_2 < 0.5$. In practice, Δ_1 and Δ_1 can be chosen arbitrarily close to their limit values.

Therefore, for the same m sequence, it is asymptotically more efficient than $\hat{d}_{m,LP}$. Velasco (1999b) extends these properties of $\hat{d}_{m,LW}$ to part of the non-stationarity zone. In particular, he shows that this estimate is consistent for $d \in (-0.5,1)$ and asymptotically Normal for $d \in (-0.5,0.75)$. He also shows that adequately tapering the observations allows estimating any degree of non-stationarity, even in the presence of a deterministic polynomial trend without the need of previously estimating this trend. More recently, Shimotsu and Phillips (2000) used the same modification of the periodogram as Kim and Phillips (2000) to develop a new local Whittle estimator, called the Modified Local Whittle estimator, which we will denote $\hat{d}_{m,MLW}$. It results from minimising a modification of the objective function given in (5):

$$Q(G,d) = \frac{1}{m}\sum_{j=1}^{m}\left\{log(G\lambda_j^{-2d}) + \frac{\tilde{I}_x(\lambda_j)}{G\lambda_j^{-2d}}\right\} \tag{7}$$

where $\tilde{I}_x(\lambda_j) = \left(w_x(\lambda_j) + \frac{e^{i\lambda_j}}{1-e^{i_j}} \cdot \frac{X_n - X_0}{\sqrt{2\pi n}}\right)\left(w_x(\lambda_j) + \frac{e^{i\lambda_j}}{1-e^{i_j}} \cdot \frac{X_n - X_0}{\sqrt{2\pi n}}\right)^*$. As can be seen, the modified periodogram $\tilde{I}_x(\lambda_j)$ differs from the periodogram $I_x(\lambda_j)$ as a result of adding the term $\frac{e^{i\lambda_j}}{1-e^{i_j}} \cdot \frac{X_n - X_0}{\sqrt{2\pi n}}$ in the Discrete Fourier transform.

These authors demonstrate that this estimator is invariant to the presence of a deterministic linear trend in the series. They also show that it is consistent for $d \in (0\ ,\ 2)$. Its asymptotic distribution is given by

$$\sqrt{m}\left(\hat{d}_{m,MLW} - d\right) \xrightarrow[n\to\infty]{} N(0,\frac{1}{4}) \tag{8}$$

for $d \in (0.5,1.75)$, which is the same as (6) This makes this estimator asymptotically more efficient than those of the LP-type.

Finally, an interesting paper by Marmol and Velasco (2002) examines the problem of discriminating between fractional (non-stationary) integration and trend-stationarity around a linear trend, and the related problem of correctly estimating and testing the significance of the trend of an FID(d) series when $d \in (0.5,1.5)$. For this range of *d,* the authors extend the results of Durlauf and Phillips (1988) obtained for *d=1*. They demonstrate that:

- the OLS estimation of the constant term diverges in distribution at a rate $n^{d-1/2}$

- the OLS estimation of the trend coefficient has a well defined limiting distribution, *upon suitable standardization given by* $n^{-d+3/2}$.
- as far as hypothesis testing on the constant term and the slope is concerned, the distributions of traditional t-students tests for both parameters diverge at a rate $n^{1/2}$

Then they propose an alternative estimate for the trend coefficient based on a local version in the frequency domain of least squares. For a process such as

$$X_t = \alpha + \beta t + X_t^o \qquad (9)$$

where X_t^0 is a zero-mean FI(*d*) process with $d \in (0.5, 1.5)$, and with α and β possibly zero, they propose the following estimates:

$$\hat{\beta}_m = \left(\sum_{j=1}^{m} I_{tt}(\lambda_j) \right)^{-1} \sum_{j=1}^{m} \Re I_{tX}(\lambda_j) \qquad (10)$$

where $1 \leq m \leq n/2$, $\Re$ stands for the real part and the (cross) periodogram is defined as

$$I_{ab}(\lambda_j) = w_a(\lambda_j) w_b(-\lambda_j)$$

Estimating the intercept term as

$$\hat{\alpha}_m = \overline{X} - \hat{\beta}_m t$$

the residual is then

$$\hat{u}_t = X_t - \hat{\alpha}_m - \hat{\beta}_m t$$

For the estimator of the variance of $\hat{\beta}_m$ in the frequency domain, they propose:

$$\hat{V}_m = \frac{1}{2} \left(\sum_{j=1}^{m} I_{tt}(\lambda_j \right)^{-2} \sum_{j=1}^{m} I_{tt}(\lambda_j) I_{\hat{u}\hat{u}}(\lambda_j)$$

The corresponding t-ratio would be:

$$\hat{t}_m = \frac{\hat{\beta}_m}{\hat{V}_m^{1/2}} \qquad (11)$$

which has a non-degenerated although non-standard symmetric limiting distribution. Their Monte-Carlo simulations provide useful information about how to obtain the approximated critical values. They can be generated according to the following formulas, where c_α stands for the critical values for a two-sided test at α probability level:

$$\begin{aligned} c_{1\%} &= 5.409 - 7.819d + 19.896d^2 \\ c_{5\%} &= 4.836 - 6.263d + 11.958d^2 \\ c_{10\%} &= 4.119 - 5.092d + 8.879d^2 \end{aligned} \qquad (12)$$

By symmetry, for one-sided tests, formulas (12) can be used with the corresponding size correction. For instance, the third equation of (12) would provide the 5% critical level for a right-tailed test.

An empirical strategy for testing the significance of the trend in a non-stationary series with $d \in (0.5, 1.5)$ unknown would then be:

- Obtain a consistent estimate $\hat{d}$ of d with methods that are robust to linear trends and to stochastic non-stationarity.
- Compute the critical values at the desired probability levels introducing $\hat{d}$ in (12)
- Estimate $\hat{t}_m$ as defined in (11) and compare it with the critical value(s) obtained in the previous step

For obtaining $\hat{d}$, the authors suggest the semi-parametric tapered estimators of Velasco (1999a, 1999b). However, it has to be noted that Kim and Phillips (2000) MLP and Shimotsu and Phillips (2000) MLW estimators share the same desirable properties, with the advantage of a lower asymptotic variance, especially in the case of the MLW estimator.

References

Agiakloglou, C. Newbold, P. and Wohar, M. (1993): “Bias in an estimator of the fractional difference parameter”, *Journal of Time Series Analysis*, 14, 235-246.

Barro, R. J., and Sala-i-Martin, X. (1991): “Convergence across states and regions”, *Brookings Papers of Economic Activity*, 1, 107-182.

Barro, R. J., and Sala-i-Martin, X. (1992): “Convergence”, *Journal of Political Economy*, 100, April, 223-251.

Bernard, A. and Durlauf, F. (1995): “Convergence in International Output”, *Journal of Econometrics*, 10, 97-108.

Bernard, A. and Durlauf, F. (1996): “Interpreting tests of the convergence hypothesis”, *Journal of Econometrics*, 71, 161-173.

Carlino, G. A. and Mills, G. O. (1993): “Are U. S. regional incomes converging?. A Time Series Analysis”, *Journal of Monetary Economics*, 32, 335-346.

Cellini,R. and Scorcu, A. (2000), "segmented stochastic convergence across the G-7 countries", *Empirical Economics*, 25, 463-474.

Cunado, J., Gil-Alana, L. A., and Pérez de Gracia, F. (2002): “Additional Empirical Evidence On Real Convergence: A Fractionally Integrated Approach”, Working Paper, Universidad de Navarra.

Davidson, J. (2002), "A model of fractional cointegration, and tests for cointegration using the bootstrap", *Journal of Econometrics,* 110(2), pp.187-212

Davidson, J. (2003), "Testing for fractional cointegration: the relationship between government popularity and economic performance in the UK" , Electronic Discussion Paper, Cardiff Business School, Revised March 2003, htttp://www.cf.ac.uk/carbs/econ/davidsonje/research.html.

Diebold, F.X., Rudebush, G.FD. (1991): "On the power of Dickey-Fuller tests against fractional alternatives", *Economics Letters,* 35, 155-160.

Ericsson, N. R. and Halket, J. R. (2002): “Convergence of Output in the G-7 Countries”, mimeo, Division of International Finance, Federal Reserve Board.

Evans, P. (1996): “Using corss-country variances to evaluate growth theories”, *Journal of Economic Dynamics and Control,* 20, 1027-1049.

Evans, P. (1997): “How fast do economies converge?”, *Review of Economics and Statistics*, vol. 79, May, 2, 219-225.

Evans, P. (1998): “Using panel data to evaluate growth theories”, *International Economic Review*, vol. 39, n° 2, May, 295-306.

Evans P. and Karras, G. (1996): “Convergence revisited”, *Journal of Monetary Economics*, 37, 249-2265.

Geweke, J.and Porter-Hudak, S. (1983): “The Estimation and Application of Long Memory Time Series Models”, *Journal of Time Series Analysis*, Vol. 4, n° 4, 221-238.

Gonzalo,J. Lee,T.H. (2000), "On the robustness of cointegration tests when series are fractionally integrated", *Journal of Applied Statistics,* vol.27, nº 7, 821-827.

Hurvich, C. M. and Deo, R. S. (1998): “Plug-in Selection of the Number of Frequencies in Regression Estimates of the Memory Parameter of a long-Memory Time Series”, *Journal of the Time Series Analysis*, Vol. 20, N° 3, 331-341.

Hurvich, C. M. and Ray, B. K. (1995): “Estimation of the Memory Parameter for Nonstationary or Noninvertible Fractionally Integrated Processes”, *Journal of the Time Series Analysis*, Vol. 16, 17-42.

Kim, C. S. and Phillips, P. C. B. (2000): “Modified Log Periodogram Regression”, mimeo, Yale University.

Künsch, H. S. (1987): “Statistical Aspects of Self-Similar Processes”, *Proceedings of the First World Congress of the Bernouilly Society*, Vol. 1, 67-74, VNU Science Press.

Li, Q. and Papell, D. H. (1999): “Convergence of International Output Time Series Evidence for 16 OECD Countries”, *International Review of Economics and Finance*, 8, 3, 267-280.

Loewy, M.B. and Papell, D., 1996, "Are Us regional incomes converging? Some further evidence", *Journal of Monetary Economics,* 38,587-598.

Maddison, A. (1995): *Moniotoring the World Economy 1820-1992,* , Development Centre Studies, OECD.

Maddison, A. (2001): *The Word Economy, a Millenial Perspective,* Development Centre Studies, OECD.

Marmol, F. and Velasco, C. (2002): “Trend stationarity versus long-range dependence in time series analysis”, *Journal of Econometrics*, 108, 25-42.

Michelacci, C. and Zaffaroni, P. (2000): “(Fractional) Beta Convergence”, *Journal of Monetary Economics,* Vol. 45.

Ooms, M. and Doornik, J. (1999): “Inference and Forecasting for Fractional Autoregressive Integrated Moving Average Models, with an application to US and UK inflation”, Econometric Institute Report 9947/A. Econometric Institute. Erasmus University Rotterdam.

Oxley,L., Greasley,D. (1995), "A time-series perspective on convergence: Australis,Uk and UDSA since 1870", *The Economic Record*, 71, 259-270.

Quah,D. (1993), " Galton's fallacy and test of the convergence hypothesis", *The Scandinavian Journal of Economics,* 95, 427-443.

Quah, D. (1996): “Convergence Empirics Across Countries with (Some) Capital Mobility”, *Journal of Economic Growth* 1 (1), 95-124.

Robinson, P. M. (1995a): “Gaussian Semiparametric Estimation of Long Range Depaendence”, *The Annals of Statistics*, Vol. 23, N° 5, 1663-1661.

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

Romer. P. (1986): “Increaing returns and long-run growth”, *Journal of Political Economy*, 94, 1002-1037.

Shimotsu, K. and Phillips, P. C. B. (2000): “Modified Local Whittle Estimation of the Memory Parameter in the Nonstationary Case”, Cowles Foundation Discussion Paper N° 1265. Yale University.

Silverbeg, G. and Verspagen, B. (1999): "Long memory in time series of economic growth and convergence”, paper prepared for Session II-3: Applied Dynamic modelling, of the Austrian Focus Sessions of the 10th SASE Conference, Vienna, Austria, July 13-16, 1998.

Solow, R. M. (1956): “A Contriburtion to the Theory of Economic Growth”, *Quarterly Journal of Economics*, 70, 65-94.

Strazicich,M.C., Lee,J., Day,E. (2001), "Are incomes converging among OECD countries? Times series evidence with two structural breaks", Working Paper.

Sun, Y. , Phillips, P. C. B. and Lee, C. (1999): “Efficient Detrending in the Presence of Fractional Errors”, mimeo, Cowles Foundation, Yale University.

Velasco, C. (1999a): “Non-stationary log-periodogram regrssion”, *Journal of Econometrics,* 91, 325-371.

Velasco, C. (1999b): “Gaussian Semiparametric Estimation of Non-Stationary Time Series”, *Journal of Time Series Analysis*, Vol. 20, N° 1, 87-127.