(Fractional) Beta Convergence

Claudio Michelacci
CEMFI

Paolo Zaffaroni
Banca d Italia

Working Paper No. 9803
September 1998

We would like to thank Simon Burgess, Marco Lippi, Domenico Marinucci, Renzo Orsi, Peter M. Robinson, and Danny Quah for very useful comments and thorough readings. Financial support is gratefully acknowledged: to Claudio Michelacci from CEP and Universita Bocconi; to Paolo Zaffaroni from the EC-HCM grant n. ERBCHBITC941742 and ESRC grant n. R000235892. The usual disclaimer applies. (E-mail address: c.michelacci@cem .es).

CEMFI, Casado del Alisal 5, 28014 Madrid, Spain.
Tel: 34 91 4290551, fax: 34 91 4291056, www.cem .es.

Abstract

Unit roots in output, an exponential 2% rate of convergence and no change in the underlying dynamics of output seem to be three stylized facts that can not go together. This paper extends the Solow-Swan growth model allowing for cross-sectional heterogeneity. In this framework, aggregate shocks might vanish at an hyperbolic rather than at an exponential rate. This implies that the level of output can exhibit long memory and that standard tests fail to reject the null of a unit root despite mean reversion. Exploiting secular time series properties of GDP, we conclude that traditional approaches to test for uniform (conditional and unconditional) convergence suit rst step approximation. We show both theoretically and empirically how the uniform 2 % rate of convergence repeatedly found in the empirical literature is the outcome of an underlying parameter of fractional integration strictly between 0.5 and 1. This is consistent with both time series and cross-sectional evidence recently produced.

1 Introduction

The debate on unit roots and stochastic trends has dominated macroeconometrics over the eighties. Since the seminal work of Nelson and Plosser (1982), this literature has noted how standard unit roots tests have failed to reject the null of a unit root in output per capita. The nineties has signed the revival of the empirics on growth and convergence. Conditional uniform convergence, namely Beta convergence, means that aggregate shocks are absorbed at an uniform exponential rate. Most of empirical studies conclude that outputs per capita of very different economies converge to their long run steady state values at a uniform exponential rate of 2% for year, (see for example Barro, 1991, Barro and Sala-I-Martin, 1991, 1995, Mankiw, Romer and Weil 1992). These seem to be two of the most striking empirical regularities in modern empirical macroeconomics. More recently, Jones (1995b), has observed that in line with the standard exogenous growth Solow model, the trend of output per capita for OECD economies is pretty smooth over time and does not exhibit any persistent changes in the post World War era.

These three stylized facts seem to be inconsistent. On the one hand a unit root in output implies that shocks are permanent so that output does not exhibit mean reversion. On the other hand Beta convergence, henceforth β-convergence implies that output converges to its steady state level at a rate that even if very low it is positive and uniform across economies. The Jones invariance property implies that steady state output could well be represented by a smooth time dependent linear trend. If this is true, unit roots tests and β-convergence are testing for the same hypothesis.

This paper starts from the observation that the size of the unit root component in GDP (the long-run effect of a unit shock) is usually found to be very low, (see Cochrane 1988, Cambell and Mankiw 1987, and Lippi and Reichlin 1991) and follows Quah (1995) in noting that cross-sectional and time series analysis can not get different conclusions. In agreement with Diebold and Rudebusch (1989), and Rudebusch (1993) we propose a different explanation. Perhaps the speed with which aggregate shocks are absorbed is so low that standard unit roots tests fail to reveal it[1]. This could actually be

[1] Diebold and Senhadji (1996) show that Rudebusch (1993) approach produces evidence

the case if GDP per capita exhibits long memory (Diebold and Rudebusch 1991). If we consider the standard Solow-Swan model and we allow for cross sectional heterogeneity in the speed with which different units in the same countries adjust, we show that the dynamics of output can exhibit long memory. We can then test for both uniform conditional and unconditional convergence allowing for rate of convergence different from the exponential one. In this framework, we show how a 2% percent rate of convergence superimposed as exponential and estimated over a time span that ranges from a minimum of 20 years to a maximum of 100 years correspond to a parameter of fractional integration that ranges from 0.51 to 0.99. This process is not covariance stationary but still mean reverting, so that standard unit roots test are likely not to reject the null of non stationarity despite the fact that convergence takes place. Using GDP per capita data for OECD countries for the period 1885- 1994, we test for this hypothesis. We conclude, that it can not be rejected, so that convergence takes place at an hyperbolic very slow rate.

The contribution of this paper is to put together two different strands of research. On the one hand, time series analysis has concluded that shocks tend to have permanent effect on the level of output. On the other hand the literature on growth and convergence has concluded that countries converge to their long run steady state value at an exponential rate that is very low and uniform across countries. In this paper we note that the two literatures are inconsistent once we allow for the invariance property by Jones and we follow the standard exogenous growth Solow model in approximating the dynamics of long run GDP per capita with a linear trend. In line with Diebold and Senhadji (1996) we propose a theoretical solution and we test for it. We conclude that standard tests for convergence suit rst step approximation despite the mispeci cation of the empirical model. In doing so we show that the parameters of fractional integration of different OECD countries, though of similar magnitude and smaller than one, are signi cantly different one from the other. This delivers a possible explanation of why time series tests of convergence based on cointegration reject the null of convergence even among OECD countries (see for example Quah 1992, Bernard and Durlauf

that distinctly favors trend-stationarity using long spans of annual data.

1993 and Bernard and Durlauf 1996). As they are these tests are mispeci ed as different variables can be cointegrated only if they exhibit the same order of integration.

Section 1 reviews the Solow-Swan model. In this context we highlight further why the three stylized facts can not go together. Section2 brie y reviews the theory of long memory processes and shows how in a extension of the theoretical model, the path of adjustment of output can exhibit long memory. In this context we show why standard unit roots can not reject the null of a unit root while a uniform 2% rate of convergence can be found to be statistically signi cant. In this framework we check for uniform (conditional and unconditional) convergence. This is done in section 3. Section 4 concludes.

2 Empirics of the Solow Growth Model and Unit Roots

We begin by brie y reviewing the Solow growth model. We then focus on the time series properties of the reduced form of the model.

Solow Growth Model

Solow model takes the rates of saving and technological progress as exogenous. There are two inputs, capital and labor. We assume a Cobb-Douglas production function, so production at time t is given by

$$Y(t) = K(t)^{\alpha}(A(t)L(t))^{1-\alpha}, \qquad 0 < \alpha < 1.$$

The notation is standard:Y is output, K capital, L labor and A the level of technology. A is assumed to grow exogenously at rate g.

The model assumes that a constant fraction of output s is invested. De n-ing $\hat{k}$ and $\hat{y}$ as respectively the stock of capital and output per effective unit of labor, $\hat{k} = K/AL$ and $\hat{y} = Y/AL$, the evolution of $\hat{k}$ is governed by

$$\frac{d\,\hat{k}_t}{dt} = s\hat{k}_t^{\alpha} - (g+\delta)\hat{k}_t, \tag{1}$$

where δ is the depreciation rate. Equation (1) implies that $\hat{k}_t$ converges towards a steady state level $\hat{k}^*$ de ned by

$$\hat{k}^* = (\frac{s}{g+\delta})^{\frac{1}{1-\alpha}}. \tag{2}$$

We can then consider a log-linear approximation of equation (1) around the steady state so that

$$\begin{aligned} \frac{d[ln(\hat{y}_t)]}{dt} &= -\beta[ln(\hat{y}_t) - ln(\hat{y}_t^*)], \\ &\text{with} \\ \beta &= (1-\alpha)(g+\delta), \end{aligned} \tag{3}$$

where $\hat{y}^* = (\hat{k}^*)^{\alpha}$. Discretizing equation (3) and indicating with y_t the log of output per capita, viz. $y = ln(Y/L)$ and by y_t^* the log of the level of output per capita in steady states we get

$$y_t - y_{t-1} = g + \beta y_{t-1}^* - \beta y_{t-1}, \qquad 0 < \beta < 1, \tag{4}$$

or equivalently

$$y_t - y_t^* = (1-\beta)[y_{t-1} - y_{t-1}^*]. \tag{5}$$

We now analyze the time series properties of both equations (4) and (5).

Time Series Properties

Equation (4) is the basic equation used to test for β-convergence (see for example Barro 1991, Barro and Sala-I-Martin 1991 and Mankiw, Romer and Weil 1992). β-convergence applies if a poor economy tends to grow faster than a rich one. This arises if the coefficient β in equation (4) is found to be positive and signi cantly different from zero. If this is the case, aggregate shocks that have pushed the current level of output away from the steady state level will be absorbed at the exponential rate β so that the dynamics of output will exhibit mean reversion. The standard approach to test for this property consists of approximating $g + \beta y^*_{t-1}$ with some control or environmental variables like the investment rate, population growth, government expenditure and so on, then estimating the regression (4) and eventually testing for the signi cance of the coefficient β. In practice, empirical studies repeatedly nd a 2% coefficient, uniform across countries and signi cantly different from zero (Quah 1993) .

A test of unit root, like for example the Dickey Fuller s test (Dickey 1979), still uses an equation like (4) and tests for the coefficient β being signi cantly different from zero, where the term $g+\beta y^*_{t-1}$ is substituted by a smooth time dependent function. A value for the coefficient β not signi cantly different from zero is interpreted as an hint of the presence of a unit root in the underlying data generating process. If this is the case a temporary shock has permanent effects on the level of output and the dynamics of output does not exhibit mean reversion towards the smooth trend. Since the seminal work of Nelson and Plosser (1982) these tests have not been able to reject the null of a unit root in GDP per capita, even if their low power is well recognized (see for example Diebold and Rudebusch 1991, Rudebusch 1993 and Diebold and Senhadji 1996).

In general the existence of a unit root in output is not in contradiction with β-convergence if we allow for the steady state level of output to be cointegrated with the current level of output. In this case aggregate shocks are still absorbed at an exponential rate despite the fact that output is integrated, as implied by equation (5).

Jones (1995a, 1995b) has observed that the dynamics of aggregate output has moved smoothly and independently of most of the controlled variable used for testing β-convergence. This is in line with the standard exogenous growth Solow model where the level of long run GDP per capita, y^*_t, is represented by the linear trend, gt. If we take the data from Maddison (1995) for 16 OECD countries over the period 1885-1994 and we plot the dynamics of per capita GDP versus a common linear trend among all the countries in the sample, we note that this simple common trend ts long run per capita GDP extremely well. This is shown in Figures 1 and 2 where we plotted each series together with a country speci c linear trend and a common linear trend obtained pooling together the series of all 16 OECD countries in our sample. The former has been estimated with OLS, the latter with GLS. Particular informative is the GLS estimate of the common trend. GLS estimating procedure implies that the better the t of the speci c trend the greater is the weight of this country in the determination of the common trend. In this case the US case outperforms by far all the other countries. This shows up in the nal outcome, in fact the common GLS trend and the US OLS speci c trend are almost undistinguishable (see Figures 1 and 2). Thus we can think of the US performance as representing the long run benchmark of all the other countries performances. Nelson and Kang (1984) argue, however, that regressions of driftless integrated series against a time trend can result in the inappropriate inference that the trend is signi cant and that it is a good description of the data, as Durlauf and Phillips (1988) show. Instead Jones (1995a) notes how a time trend calculated using data only from 1880 to 1929 forecasts extremely well the current level of GDP of the US economy. Following Diebold and Senhadji (1996) this is clearly incompatible with difference stationarity in aggregate output, as new information seems to be irrelevant for forecasting on very long horizons.

This suggests that, in accordance with the standard exogenous growth Solow model where the level of long run GDP per capita, y_t^*, is represented by the linear trend, $g\,t$, the dynamics of steady states output mimics a simple trend. As a deterministic function can not be cointegrated with a variable exhibiting stochastic trends, it turns out that β-convergence and unit root tests are both checking for mean reversion towards a smooth time dependent trend. In a time series formulation we can say that β-convergence is testing for trend stationarity in output where the stationary disturbance is superimposed as an autoregressive process of order one [2]. These simple con-

[2] The nature of the problem is just further complicated by the fact that growth theorists use panel data instead of just time series. Recent results (e.g. Levin and Lin 1992) show, however, that panel data just make dramatically increase the power of a unit root test as the cross sectional dimension increases.

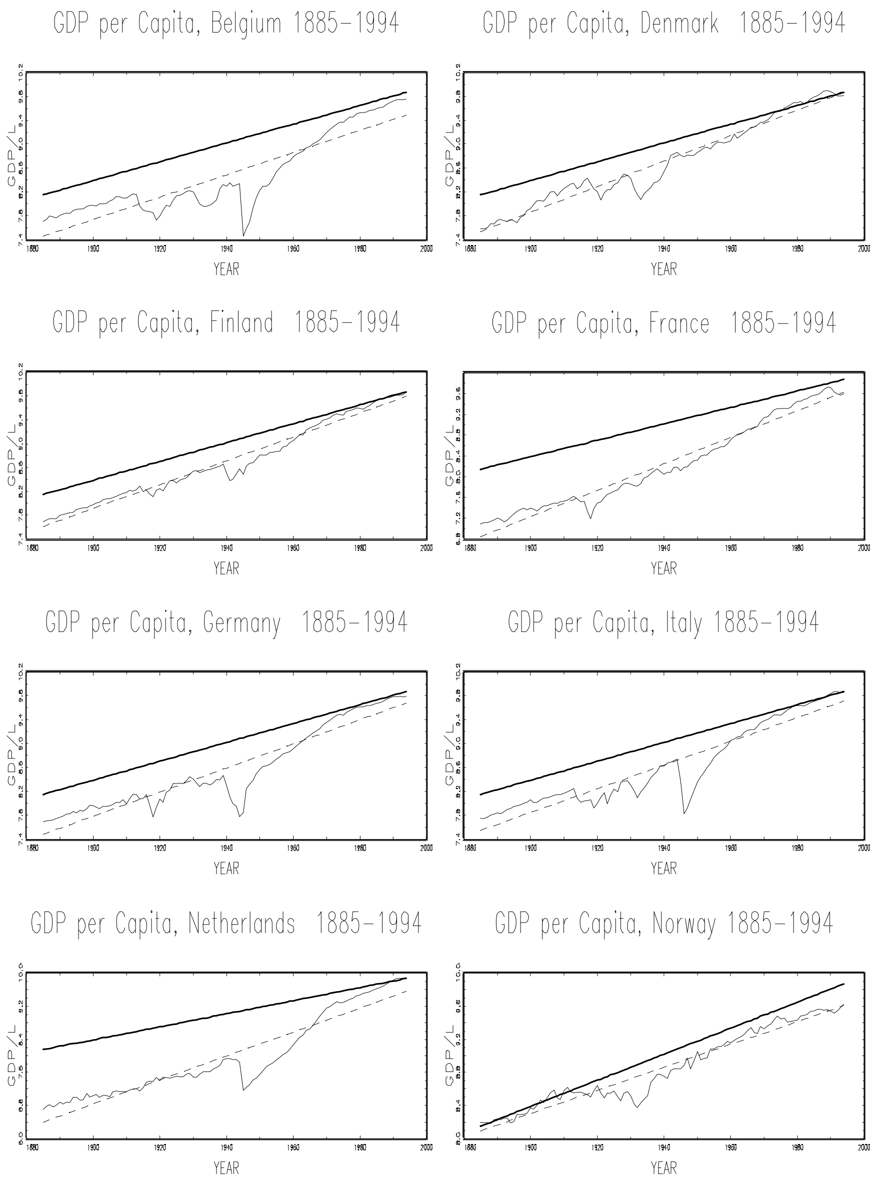

Figure 1: The dashed and bold lines represent the country-speci c (OLS) and common (GLS) trend, respectively. The solid line represents logged GDP.

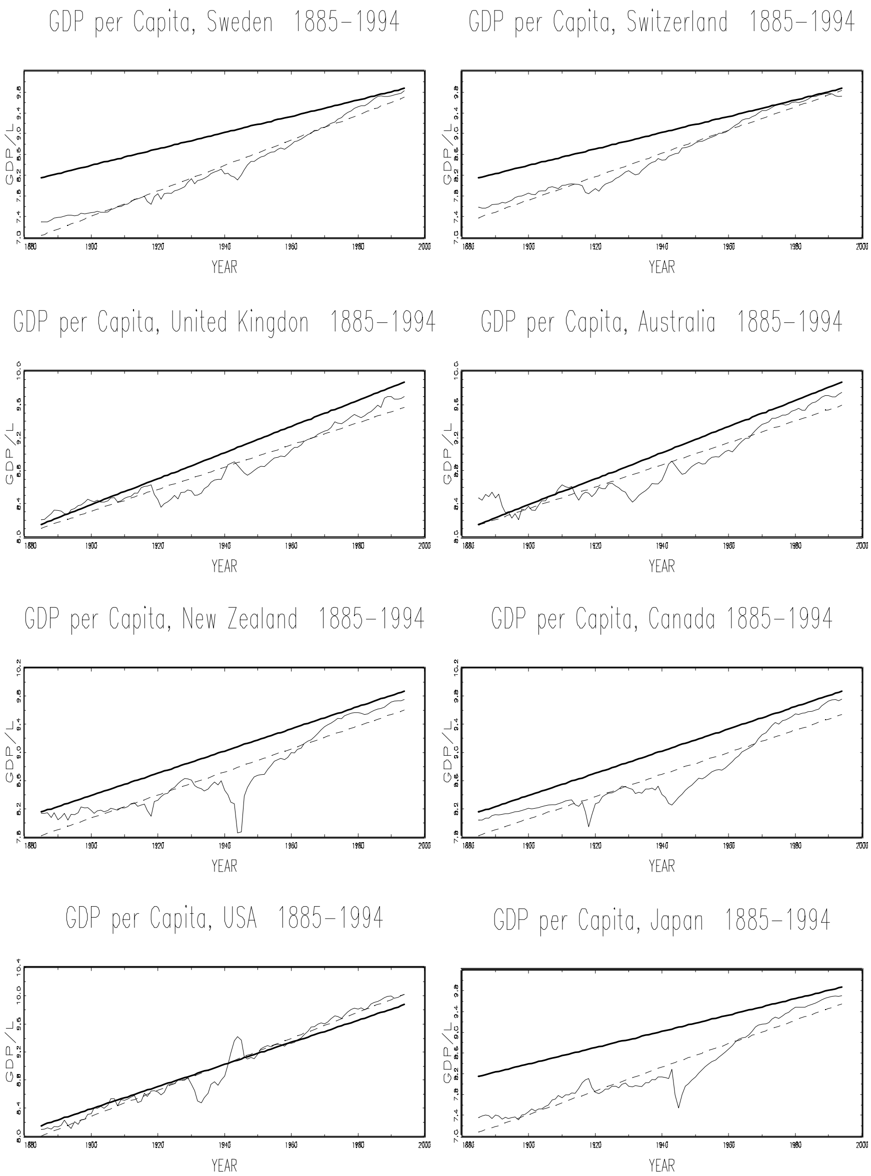

Figure 2: It continues previous gure.

siderations imply that testing for β-convergence is meaningless if we assume the Jones invariance property together with the existence of a unit root in output[3]. As they stand, these three stylized facts can not go together. Our claim is that the (two equivalent) tests are both checking for a superimposed rate of exponential mean reversion.

If the rate of mean-reversion in (logged) GDP per capita or equivalently the rate of absorption of the shocks is hyperbolic (in a sense to be de ned precisely below) instead of exponential, β-convergence would apply in the sense that poorer economy would grow faster and would converge towards their long run steady state and standard unit-root tests would fail to reject a unit-root albeit not present (see for example Diebold and Rudebusch 1991).

3 Theory of Long Memory and the Barro Regression

In this section we brie y review the theory of *long memory* processes which allows the possibility of hyperbolic mean reversion together with non-stationarity. We will then analyze why the Barro regression might be robust to rate of convergence different from the exponential one delivering the right answer to the problem of convergence.

3.1 Theory of Long Memory Processes

Unit roots describe only a small set of nonstationary processes. A class that embeds either (covariance) stationary processes and unit roots is given by *strongly dependent* processes also known as *long memory* or *long range dependent* processes (see Robinson 1994 for a survey on the topic). Usually only the second moments properties are considered in order to characterize such a behaviour in terms of either the behaviour of the autocorrelation function at the long lags or the power spectrum at the zero frequency.

[3] For example Den Haan (1995) notes that the slow speed of convergence observed in the data can be reconciled quantitatively with the neoclassical growth model assuming either a capital share equal to around 0.8 or a sufficient amount of persistence in the stochastic process driving technological progress. In either cases, the 2% rate of convergence is incompatible with aggregate output exhibiting a unit root (see his equation 3.4).

We shall assume that K denotes any positive constant (not necessarily the same) and $\sim$ asymptotic equivalence.

De nition 1
A real valued scalar discrete time process X_t is said to exhibit long memory in terms of the power spectrum with parameter $d > 0$ if

$$f(\lambda) \sim K\,\lambda^{-2d}, \qquad as \qquad \lambda \to 0^+.$$

In the nonstationary case ($d \geq 1/2$, see below) $f(\lambda)$ is not integrable and thus it is de ned as a pseudo-spectrum.

The importance of this class of processes derives from smoothly bridging the gap between standard stationary processes and unit roots in an environment that maintains a greater degree of continuity (Robinson 1994). For the purpose, let us consider a parametric example.

Let $\{y_t\}$ be a discrete time scalar time series , $t = 1, 2, \ldots$, suppose v_t is an unobservable covariance stationary sequence with spectral density that is bounded and bounded away from zero at the origin, such that

$$(1-L)^d y_t = v_t, \qquad t = 1, 2, \ldots \tag{6}$$

where L is the lag operator. If $d = 0$, then y_t is a standard or better *weak memory* (covariance) stationary process with spectral density bounded away from zero (i.e. an $ARMA$ process), whereas y_t is a random walk if $d = 1$. The parameter d however does not need to be an integer.

In what follows, we focus on the case in which y_t is a long memory process with parameter d positive, real with $0 < d < 1$. In this case, when v_t is assumed to be a white noise process, the process y_t de ned in (6) is called an $ARFIMA(0, d, 0)$ process and more in general when v_t is an (inverted) $ARMA(p, q)$ we obtain an $ARFIMA(p, d, q)$ process.

The power spectrum of the y_t process is given by

$$f_y(\lambda) = \mid 1 - e^{i\lambda} \mid^{-2d} f_v(\lambda) = (2sin(\lambda/2))^{-2d} f_v(\lambda)\,, -\pi \leq \lambda \leq +\pi\,,$$

where $f_v(.)$ denotes the power spectrum of the v_t process. Thus from $sin(\omega)/\omega \sim 1\,, \omega \to 0$, when $d > 0$ as $\lambda \to 0^+$ we get

$$f_y(\lambda) \sim 4^{-d} f_v(0) \lambda^{-2d}\,.$$

Whenever $d > 0$ the power spectrum is unbounded at the zero frequency, implying that the series y_t exhibits long memory. This class of processes have many important properties. When $0 < d < 1/2$, y_t has both nite variance and exhibits mean reversion. When $1/2 < d < 1$ the process has in nite variance but it still exhibits mean reversion. This process is not (covariance) stationary but less non stationary than an unit root process so that standard unit root tests exhibits low power with respect to this alternative despite the presence of mean reversion (Diebold and Rudebusch 1991). When $d \geq 1$ the process has in nite variance and stops exhibiting mean reversion. In particular a unit root process is obtained when $d = 1$. This represent a particular case of a long memory process: a process with an in nite memory.

If $-1/2 < d$, (6) can be inverted so that

$$y_t = \sum_{i=0}^{\infty} \gamma_i v_{t-i}, \;\; \gamma_i = \textstyle\prod_{k=1}^{i} \frac{k-1+d}{k}, \;\; i \geq 1, \;\; \gamma_0 = 1. \tag{7}$$

By use of Stirling s approximation it follows that as $i \to \infty$

$$\gamma_i \sim Ki^{d-1} . \tag{8}$$

This can be interpreted such as the effect of a shock v_{t-i} , i periods ahead, vanishes at an hyperbolic rather than exponential rate exhibiting an high level of persistence, higher the bigger the parameter d. When $d = 1$, the unit root case arises where a shock arbitrarily far away in time exhibits permanent effects on the current level of y_t.

This persistence property re ects the characterization already given in the frequency domain. We have seen that a long memory process for $0 < d$ is de ned by an unbounded spectrum at the origin. It is well accepted that the degree of persistence of a shock can be expressed by the level of the spectral density at zero frequency (Cochrane 1988) . The de nition of long memory and the previous considerations suggest to take as an exact measure of persistence the slope of the logged spectrum at the origin[4]. In fact, taking logs in both terms in *De nition 1*, we obtain as $\lambda \to 0^+$ the following representation

$$ln\left(f(\lambda)\right) \sim K - 2dln(\lambda)\,, \tag{9}$$

[4]This concepts is directly derived from a well known strand in the semi-nonparametric econometrics literature Robinson (1995), Geweke and Porter Hudak (1983).

With respect to the scatterplot of the logged spectrum and $2ln(\lambda)$,the unit root case will be represented by a line with slope minus $\pi/4$ while the case $d < 1$ is represented by a atter line. Obviously the bigger (in absolute value) the slope the greater the level of persistence. The idea expressed by (9) is at the core of the estimation procedure suggested by Geweke and Porter Hudak (1983) and formalized in Robinson (1995) that is brie y described in the Appendix 1 .

3.2 Robustness of the Barro regression

This section tries to rationalize the nding of a signi cant regression coefficient of β-convergence in (4) .

At rst let us consider some back of the envelope calculations. A 2% rate of convergence superimposed as exponential over a time span of 20 - 110 years is almost observational equivalent to a parameter of fractional integration strictly between 0.5 and 1. In fact, bearing in mind the result in (8) a parameter of fractional integration, d, that resembles the 2% exponential rate of decay after a k period ahead shock can be obtained solving the simple equation[5]

$$(0.98)^k = k^{d-1}. \tag{10}$$

In Table 1 below we report the solutions of this simple equation, for values of k that ranges from 10 to 110. As most of empirical studies have used sample that ranges from 20 to 100 years, we can consider an underlying parameter of fractional integration strictly between 0.5 and 1 as the driving force behind the 2% rate of convergence found in the empirical literature on β-convergence.

Secondly, let us consider now the following theoretical result due to Sowell (1990), theorem 4. Regressing a variable on its lagged value , the Student

[5]Of course this is just a very simple and approximate exercise yet useful in order to understand the main intuition of the paper.

$d \sim 2\%$ exp. rate	N. of Obs.
0.912	10
0.865	20
0.821	30
0.781	40
0.742	50
0.703	60
0.667	70
0.631	80
0.596	90
0.561	100
0.527	110

Table 1: Parameter of Fractional Integration Corresponding to the 2 % exponential rate

t of the coefficient behaves discontinuously when the process generating the variable is an $ARFIMA(0, d, 0)$ with $d > 0$. When $d = 1$ we obtain that the asymptotic distribution of the Student t normalized at the value one, is the well known Dickey Fuller distribution (Dickey 1979). But when $d \neq 1$ one obtains very different results that is

$$t \rightarrow_p \begin{cases} \infty & , 1 < d < 3/2 \, , \\ -\infty & , 1/2 < d < 1 \, . \end{cases}$$

Let us now start to draw our conjecture.

If per capita GDP is well represented by a long memory process with parameter d with $1/2 < d < 1$, thus displaying in nite variance together with (what is important) mean-reversion, tting the Barro regression would tend to give a signi cative negative Student t (actually converging to negative in nity in probability). Thus this simple inference gives exactly the same conclusion of the aforementioned regression (4) obtained in the literature when tting an exponential rate of convergence.

More, the back of the envelope calculations show that superimposing an exponential rate of decay over a long memory process with $1/2 < d < 1$ gives precisely the well established 2% rate of β-convergence.

Finally, the property of long memory processes to nest the unit root case in a class that maintains a greater level of continuity rather than standard weak dependent processes motivates the empirical nding of systematically non signi cant unit root tests.

If our conjecture is right, we could say that the standard approach to test for β-convergence (Barro 1991), suits rst step approximation despite the mispeci cation of the empirical model. This test tends to exhibit negative Student t in the case of mean reversion ($d < 1$), leaving nonetheless some margins of ambiguity in a particular case of lack of convergence, the unit root case ($d = 1$). On the other hand the Student t will diverge to plus in nity when $d > 1$ delivering the right answer to the issue of convergence.

At this stage, our conjecture still lacks of two elements, a purely economic one and a conclusive statistical one. We will show a possible source of the long memory feature of the data in a version of the Solow model augmented by cross sectional heterogeneity. Secondly, there is the need of a rigorous time series analysis of the data to show that the logged per capita GDP is well represented by a mean-reverting long memory process with $1/2 < d < 1$. This is done in section 6.

4 The Solow Model Augmented by Cross-sectional Heterogeneity

In this section we show how long memory could arise in the Solow growth model. Suppose that the economy is characterized by N units each behaving as in the standard Solow model outlined in section 2. That means that each of these units representing either different rms or sectors in the same economy are investing a fraction s_i of their output in the accumulation of capital [6]. If this is the case the dynamics of output, y_t^i, of each of these rm-sector, with steady state output $y_t^{*^i}$, is governed by

$$y_t^i - y_t^{*^i} = (1-\beta_i)[y_{t-1}^i - y_{t-1}^{*^i}] + \epsilon_t^i + \eta_t, \qquad i = 1...N,\ 0 < \beta_i < 1. \quad (11)$$

[6] Theoretically this structure could arise either in a world with imperfect capital markets where human capital is used as a collateral or because of adjustment costs (see Barro and Sala-I-Martin 1995). We decided not to model directly these frictions here because of the space constraint. Even the assumption that each units is evolving as an autoregressive process of order one is a simplifying assumption that it is is not needed to get the result as it will become clear thereafter.

where ϵ_t^i, η_t represent respectively idiosyncratic and aggregate shock assumed mutually uncorrelated white noise and β_i is equal to $(1-\alpha_i)(g_i+\delta_i)$. Here α_i, g_i and δ_i are respectively the unit s speci c productivity of capital, the rate of technological progress and the depreciation rate. It follows that the variable $x_t^i = y_t^i - y_t^{*,i}$ behaves like a rst-order autoregressive process.

If we indicate respectively with

$$\begin{aligned}\bar{y}_t &= \frac{1}{N}\sum_{i=1}^{N} y_t^i\,,\\ \bar{y}_t^* &= \frac{1}{N}\sum_{i=1}^{N} y_t^{*^i}\,,\end{aligned}$$

current and long run equilibrium aggregate output, we then have that the amount of disequilibrium in the economy evolves as

$$\bar{y}_t - \bar{y}_t^* = 1/N\sum_{i=1}^{N}(1-\beta_i)[y_{t-1}^i - y_{t-1}^{*,i} + \epsilon_t^i] + \eta_t\,. \tag{12}$$

Let us de ne $\bar{x}_t = \bar{y}_t - \bar{y}_t^*$.

The above equation can behave very differently from equation (5) even if all the coefficients β_i are bounded between zero and one. We will show that under certain conditions on the cross sectional distribution of the coefficients β_i, $\bar{x}_t$ exhibits long memory. In fact if we assume that the aggregate η_t and idiosyncratic ϵ_t^i shocks are uncorrelated, we get that the power spectrum $f_k(\lambda)$ of x_t^k is equal to

$$f_k(\lambda) = \frac{var(\epsilon_t^k)}{2\pi \mid 1-(1-\beta_k)e^{i\lambda}\mid^2} + \frac{var(\eta_t)}{2\pi \mid 1-(1-\beta_k)e^{i\lambda}\mid^2}. \tag{13}$$

This implies that the power spectrum $\bar{f}(\lambda)$, $-\pi \le \lambda < \pi$ of the aggregate $\bar{x}_t$ is equal to

$$\bar{f}(\lambda) = \bar{f}_1(\lambda) + \bar{f}_2(\lambda)\,, \tag{14}$$

where

$$\begin{aligned}\bar{f}_1(\lambda) &= \frac{1}{N^2}\sum_{k=1}^{N}\frac{var(\epsilon_t^k)}{2\pi \mid 1-(1-\beta_k)e^{i\lambda}\mid^2}\,,\\ \bar{f}_2(\lambda) &= \frac{var(\eta_t)}{2\pi N^2}\mid \sum_{k=1}^{N}\frac{1}{(1-(1-\beta_k)e^{i\lambda})}\mid^2\ .\end{aligned}$$

If we assume that the coefficients β_k are independent drawings from a distribution $F(\beta)$ and that the $var(\epsilon_t^u)$ are drawn from another distribution independent of the rst, we follow Robinson (1978) and Granger (1980) to obtain

$$\bar{f}(\lambda) \simeq \frac{1}{2\pi N}(E[var(\epsilon_t^k)]) \int_{\mathcal{B}} \frac{1}{\mid 1-(1-\beta)e^{i\lambda} \mid^2} dF(\beta) + \qquad (15)$$

$$\frac{var(\eta_t)}{2\pi} \mid \int_{\mathcal{B}} \frac{1}{(1-(1-\beta)e^{i\lambda})} dF(\beta) \mid^2 , \qquad (16)$$

where $\mathcal{B}$ denotes the support of the distribution $F(\beta)$ and $\simeq$ denotes that the relation holds approximately for N big but nite.

In general long memory arises if the integral in (15)

$$\int_{\mathcal{B}} \frac{1}{(1-(1-\beta))^2} dF(\beta) = E_F(1/\beta^2) , \qquad (17)$$

diverges, where $E_F(.)$ denotes the expectation over the measure $F(.)$ In fact the second integral in (16), viz. $E_F(1/\beta)$, diverges under stronger conditions which imply the divergence of the former integral in (15) but not viceversa as we will make clear in the sequel.

We can establish necessary and sufficient conditions on the distribution function $F(.)$ such that the integral (17) is unbounded. In general we know (e.g. in Rudin 1973) that the integral $\int_a^b h(t)dt$ for a continuous function $h(x)$ on an interval $[a, b)$ is unbounded, if $h(.)$ has at least the same order of in nity as $1/(b-x)^\alpha$ when x goes to b, that is

$$1/(b-x)^\alpha = O(h(x)), \qquad x \to b^- .$$

If we assume that the distribution function $F(\beta)$ is absolutely continuous having a density $f(\beta)$, the integrand function of (17) is given by $f(\beta)/\beta^2$. Thus a sufficient condition for $\bar{x}_t$ to exhibit long memory is simply given by

$$f(\beta) \geq K\beta, \quad \text{as } \beta \to 0^+ ,$$

for some positive constant K. Thus the density $f(\beta)$ might go to zero as $\beta \to 0$ but at slower rate than β [7].

[7] Instead for the integral in (16) to diverge we need the stronger condition $f(\beta) \geq K$, as $\beta \to 0$ which clearly implies the former one. Moreover, the presence of the N and N^2 terms in (15) and (16) does not affect the result as we assume that the above arguments hold for a big but nite N.

The main implication is that the aggregate process might display long memory even if the aggregating elements are stationary with probability one. Also the result is valid even if the aggregating elements are $ARMA$ processes. In this case the condition to be satis ed is that the probability of extracting a unit root in the autoregressive component dies slowly enough. Moreover it is important to stress that the result does not depend from either the nature of the idiosyncratic and common shocks given their stationarity or from the type of dependence among them. Mankiw, Romer and Weil (1992) argue that the slow speed of convergence observed empirically can be reconciled quantitatively with the the neoclassical growth model if the capital share is sufficiently high and around 0.8. This result, on the other hand, delivers a different rational for the low rate of convergence found in the empirics of the Solow growth model based on aggregation of cross-sectional heterogenous units [8].

Intuitively, long-range dependence means that shocks arbitrarily far away in time still exhibit some in uence on the future dynamics of the process. Cross-sectional aggregation kills the Markovian property implicit in standard weak memory (covariance) stationary processes provided that there are some units with a sufficiently amount of persistence. In this case, to keep track of the future dynamics of the aggregate system we must recover the past history of the units of the system if we want to know the relative distribution of disequilibria in the economy.

[8] As an example we can consider Granger (1980) formulation where the coefficients β_k are drawn (independently both of the idiosyncratic shocks, ϵ_t^i, and common shocks, η_t) from a Beta(p,q) distribution. Thus we get that the integrand function (neglecting unimportant constant terms) is given by

$$\frac{(1-\beta)^{p-1}\beta^{q-1}}{\beta^2},$$

thus yielding the condition $q < 2$ which coincides with what Granger (1980) obtained by expanding the integral in terms of autocovariances. In fact in this case the aggregate process can be shown to display long memory with parameter $d = 1 - q/2$.

5 Generalizing the concept of Beta convergence

In this version of the Solow model augmented by cross-sectional heterogeneity, it seems reasonable to propose the following de nitions of β-convergence:

(i) An economy has no tendency to converge towards either its own or the common steady state if, after tting either a country speci c or a common (linear) trend repectively, the parameter of fractional integration d of the residuals is greater or equal than one ($d \geq 1$). In the former case we say that there is no *conditional* convergence and that there is no *unconditional* convergence in the latter.

(ii) The case of the Solow model without cross sectional aggregation is represented by the absence of long memory that is d equal to zero. In this case, if we want to recover the rate of convergence of the economy, we must solve for the roots of the characteristic equation and look for the greatest solution in absolute value.

(iii) *Uniform unconditional* convergence means that if we t a common (linear) trend across all the units in the sample, then the residuals exhibit similar parameter of fractional integration d.

(iv) *Uniform conditional* convergence means that if we t a country speci c (linear) trend for all the units in the sample, then the residuals exhibit similar parameter of fractional integration d.

We consider further evidence of the exponential 2% rate of convergence, if we nd a parameter of fractional integration strictly between 0.5 and 1 (c.f. see section 3.).

In order to make inference on the parameters of long memory of the series we employ the semiparametric approach introduced by Geweke and Porter Hudak (1983) . Rigorous analysis of this estimator is given in Robinson (1995) who established consistency and asymptotic normality of the estimator. Also the result has been developed in a multivariate framework , a novel feature in this literature, which represents a crucial property in order to apply this estimator to a multicountries issue as the question of convergence. Robinson (1995) results are valid without assuming any a priori restriction on the degree of dependence in the data allowing for either antipersistence ($-1/2 < d < 0$), weak ($d = 0$) or long memory ($0 < d < 1/2$), the only restriction on the parameter space being nite variance, viz. $\mid d \mid < 1/2$. We defer to Robinson (1995) for the formal proofs of the results, describing the main features in Appendix 1.

6 Empirical Results

At rst, to motivate our conjecture that the per capita GDP is characterized by a dynamics that is well approximated by a long memory process let us consider Figures 3, 4, 5 and 6. Interpreting the result according to *De nition 1*, Figures 3, 4, 5 and 6 show how the periodogram (i.e. an estimate of the spectrum) for each of the series in our sample displays a peak at the origin. This is what Granger (1966) de nes to be the typical spectral shape of an economic variable and it is the main feature of a long memory process [9] .

Figures 3, 4, 5 and 6 plot the logged periodogram against twice the logged frequency. As shown in section 3, the slope of the interpolating line expresses approximately the parameter d. The unit root case is represented by the line with slope minus $\pi/4$. It is evident how the interpolating line is always atter than the bisector thus supporting the absence of the unit root case. Nevertheless the slope appears still positive and in particular between 1/2 and 1.

Table 2 reports the estimates based on the log-periodogram estimator [10]. Most of the parameters of fractional integration,$d's$, are less than one even if with a very high standard error. As we are interested in the OECD coun-

[9] For an analysis of the behaviour of the periodogram for non stationary processes see Hurvich and Ray (1994).

[10] Diebold and Rudebusch (1989) has used a similar estimator, valid in a univariate case only (Geweke and Porter Hudak 1983). The multivariate framework, the gains in efficiency and computability of the Robinson (1995) estimator motivates our choice of using the latter instead of the former thus explaining the difference in the estimates of the parameter d for the US case obtained by Diebold and Rudebusch (1989). The appendix reviews the main features of the estimating procedure.

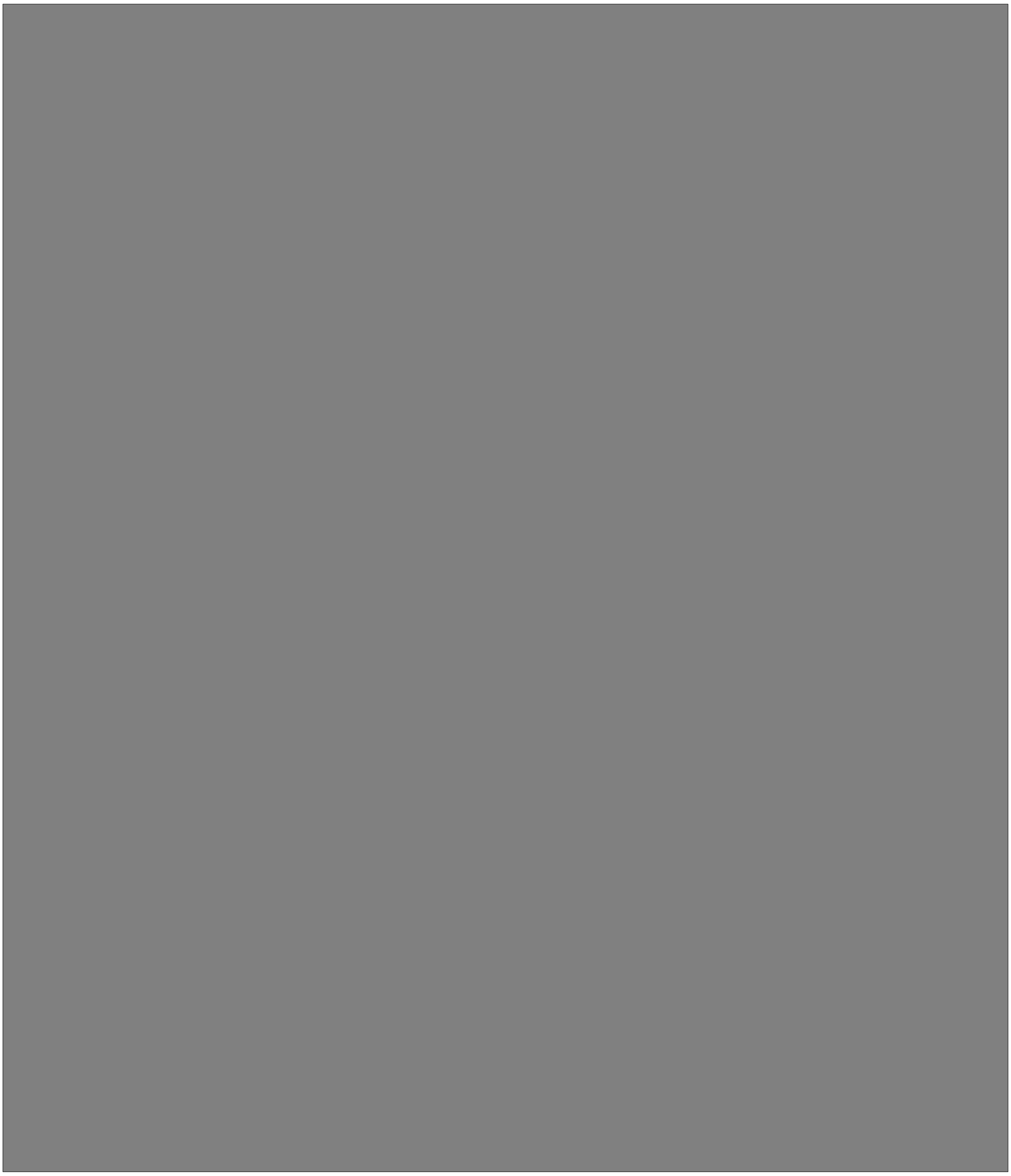

Figure 3: The left hand side column displays the periodogram of the logged GDP (1865-1994) for the 16 OECD countries here considered. The right hand side column displays three lines versus the logged frequency: the continuous line represents the logged periodogram ordinates, the bold line represents the OLS interpolating line (cf. Table 2) while the dashed line represents the unit root case (slope $\pi/4$). An interpolating line atter than the bysector corresponds to a value of the long memory parameter d smaller than one.

Figure 4: It continues previous gure.

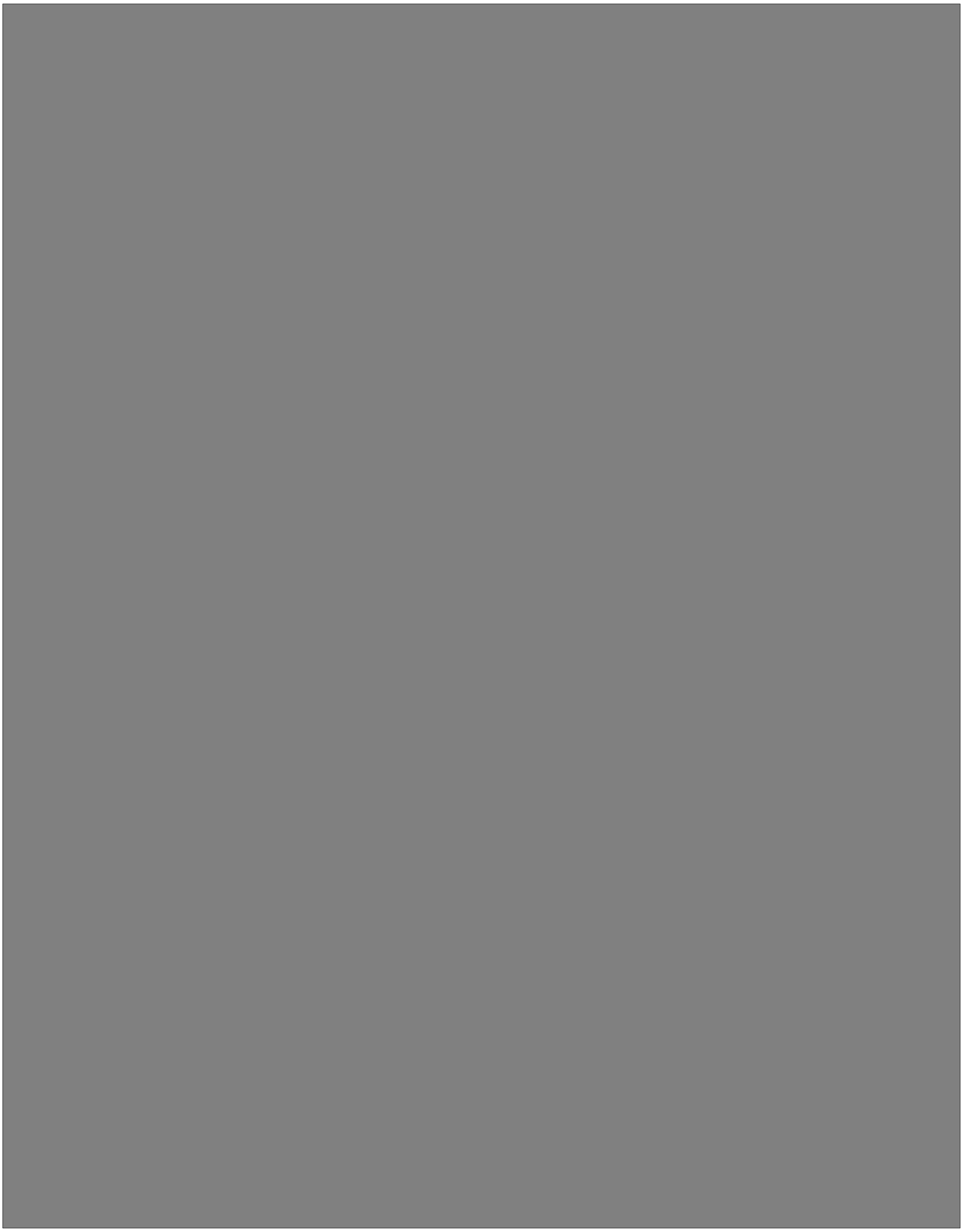

Figure 4: It continues previous gure.

Figure 5: It continues previous gure.

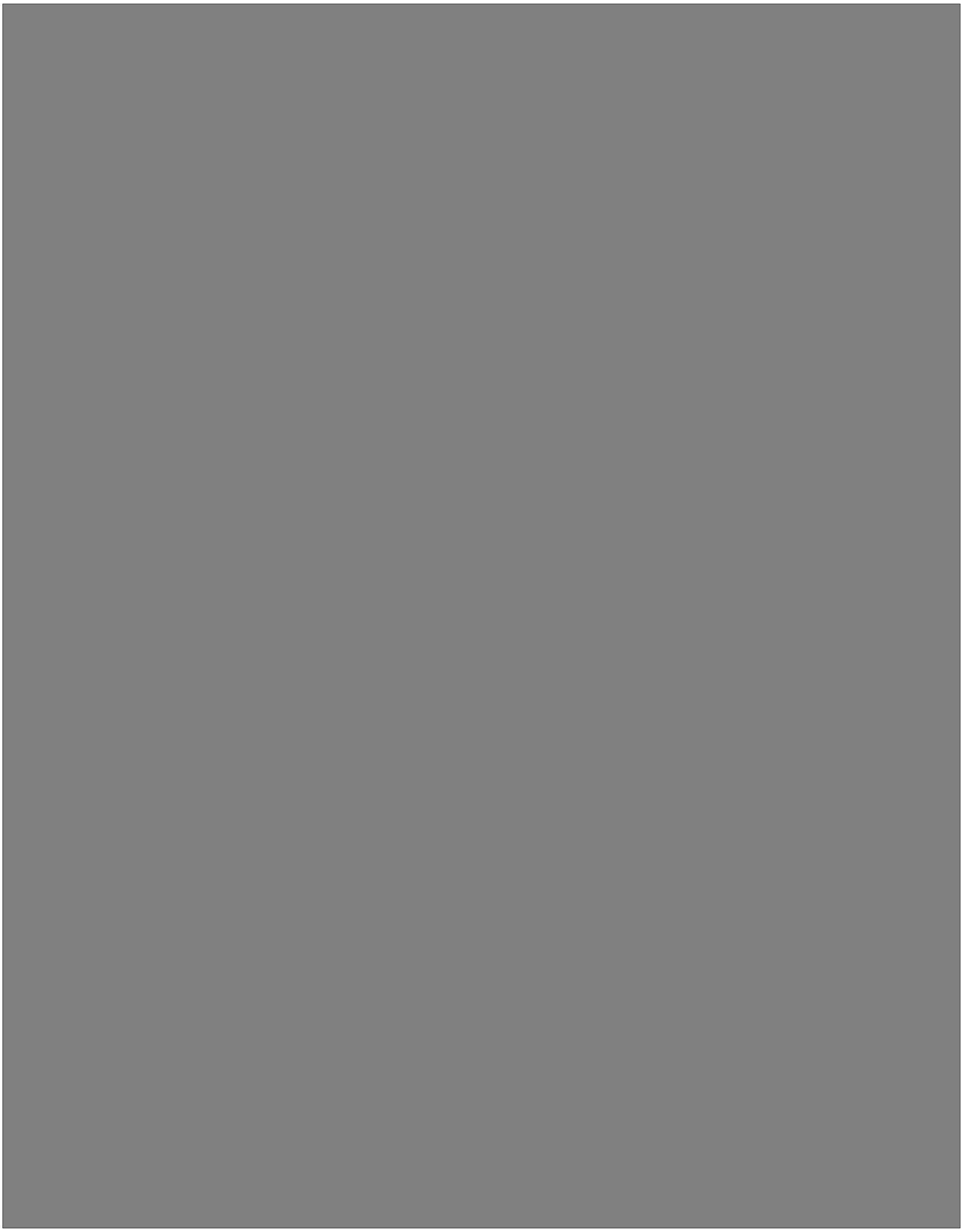

Figure 5: It continues previous gure.

Figure 6: It continues previous gure.

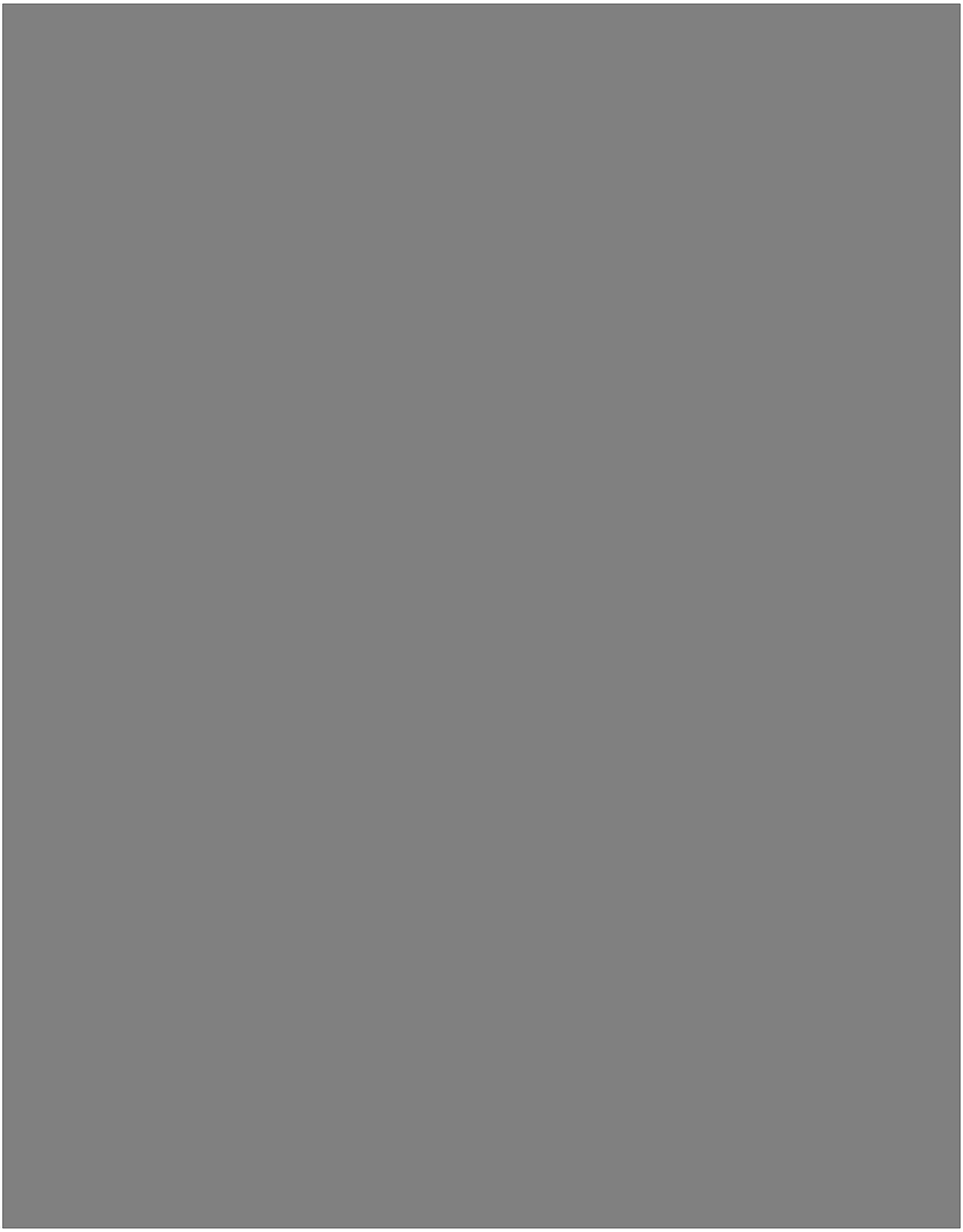

Figure 6: It continues previous gure.

Country	Conditional	Unconditional
Belgium	0.52	0.55
Denmark	0.84	0.55
Finland	0.99	0.98
France	0.56	0.94
Germany	0.83	0.83
Italy	0.56	0.65
Netherlands	1.11	1.26
Norway	0.81	0.82
Sweden	0.58	1.30
Switzerland	1.03	0.84
U.K.	0.58	0.58
Australia	0.69	0.75
New Zealand	0.85	0.85
Canada	0.97	0.96
U.S.A.	0.57	0.46
Japan	0.61	0.92
Asymptotic S.E.	0.177	0.177
Wald test statistic	1.24e+16 (0.0)	1.62e+16 (0.0)

Table 2: Log-Periodogram Estimates of d, (OECD, 1885-1994). The estimation procedure is described in the Appendix. The Wald test statistic is distributed as a χ^2 with 15 degrees of freedom under the hypothesis $H_0 : d_1 = d_2 = ... = d_{16}$. $P-$values are reported in parenthesis.

tries as a group, we use an induced test based on the sequential Bonferroni approach[11]. We want to test for the existence of a number d_0 strictly less than one, such that all the parameter of fractional integration of the OECD countries in the sample are less or equal than d_0. For an overall level of signi cance of 10 percent, we examine the country with the highest ex post probability of rejecting the null hypothesis and set the signi cance level using the total number of countries examined.

The results of this procedure are reported in Figure 7. The horizontal line represents the 10 percent critical value of the t-statistics such that the null hypothesis is rejected. The x-axis represents the coefficient d_0 considered under the null. The negatively sloped line shows the actual t-statistics calculated for different null hypotheses. Figure 7 shows that it exists a non empty set of values of d_0, strictly less than one, such that the null hypothesis that the parameter of fractional integration of all the OECD countries are less than d_0 can not be rejected at the 10 percent signi cant level. We also note how this set always lays above the value 1/2.

The empirical results can be summarized as follows:

- GDP per capita of all the countries in the sample exhibit long memory ($d > 0$). In our framework this suggests that the economy behaves as an aggregation of Solow models rather than as a Solow model itself.
- The hypothesis that all the OECD countries are non stationary still mean reverting ($0.5 < d < 1$) can not be rejected using the induced test based on the Bonferroni procedure (Figure 7).
- We found the 2% rate of convergence in the form of a parameter of fractional integration strictly between 0.5 and 1.
- The rates of convergence are very low and similar across countries even if the rate of convergence is not *uniform* as the null hypothesis that

[11]This procedure yields a conservative yet consistent test (Gourieroux and Monfort, 1989, Property 19.7). The exact test for one-side multivariate hypothesis (Gourieroux and Monfort, 1989, Chapter 21) is not implementable when the number of constraints is greater than two.

Figure 7: The test statistic for the null $H_0 : d_i < d_0$, $i = 1, ..., 16$ is plotted for different values of d_0. The horizontal line represents the critical value for a 10% signi cance level.

the coefficients of fractional integration are constant across countries is strongly rejected (see Table 2) .

- As the order of integration of different OECD countries are different, time series tests of convergence based on cointegration are mispeci ed.

- We conclude that there is unconditional convergence across OECD countries and the rates of convergence are pretty similar even if the test reject the null of exact equality of the coefficients.

7 Conclusions

In this paper we embed standard approaches to test for β-convergence in a more general framework. In order to do so we join different strands of literature, the aggregation theory of dynamic economic models, the theory of long memory processes and the literature on the empirics of growth.

We give striking evidence that the (de-trended) per capita GDP is well approximated at the low frequencies by a long memory process displaying nonstationarity together with mean reversion, stressing the importance of capturing in the very long the true rate of convergence. We then nd primitive conditions under which long memory arises naturally as the result of aggregating heterogeneous units in the same economy and we then apply it to an extension of the Solow-Swan model augmented by cross-sectional heterogeneity.

Finally we draw robust inference on the possibility of conditional and unconditional β-convergence among the OECD countries and as a result we support the conclusion of the well established Barro type of regression and we reconcile both time series and cross sectional evidence.

Some questions still remain open. In particular, we stress how the degree of persistence differs among OECD countries. This drives the question of whether the underlying economic structure of OECD countries are different and asks for a further investigation of what country speci c economic mechanism make long memory to arise in real world.

8 Appendix 1. The logperiodogram estimator

Following Robinson (1995), let us suppose that the time series under study is given by the G dimensional real valued vector $Z_t = (Z_{1,t}, \ldots, Z_{G,t})'$. The (g,h)th element of the spectral density matrix $f(\lambda)$ is denoted by $f_{gh}(\lambda)$. For (C_g, d_g) , $g = 1, ..., G$ satisfying $0 < C_g < \infty$ and $\mid d_g \mid < 1/2$ it is assumed that[12]

$$f_{gg}(\lambda) \sim C_g \lambda^{-2d_g} \quad \text{as } \lambda \to 0^+ .$$

This represent the only assumption on the shape of the spectrum which motivates the semiparametric nature of the estimator of the $G{+}G$ parameters (C_g, d_g) $g = 1, \ldots, G$ beside integrability to ensure stationarity.

The periodogram[13] for the g-th component $Z_{gt}, t = 1, \ldots, N$, N being the sample size, is denoted by

$$I_g(\lambda) = \frac{1}{2\pi N} \mid \sum_{t=1}^{N} Z_{gt} e^{it\lambda} \mid^2 , \; g = 1, \ldots, G . \tag{18}$$

De ning the fourier frequency $\lambda_j = \frac{2\pi j}{N}$ one has to de ne the log-periodogram

$$Y_{gk} = ln(I_g(\lambda_k)) , \; g = 1, \ldots, G , k = l + 1, \ldots, m . \tag{19}$$

The positive integer m is the user-chosen bandwidth number and the positive integer $l$is the user-chosen trimming number[14]. In this context there is just the need to say that the asymptotic results require that m and l both tend to in nity with N but more slowly together with $l/m \to 0$. Then de ning the unobservable random variables U_{gk} by the following set of regressions

$$Y_{gk} = c_g - d_g(2ln\lambda_k) + U_{gk} \, g = 1, \ldots, G \;\; , k = l + 1, \ldots, m. \tag{20}$$

where $c_g = ln\, C_g + \psi(1)$ which involves the digamma function $\psi(z) = (d/dz)ln\Gamma(z)$, with $\Gamma(.)$ being the gamma function.

[12] Basically as in *De nition 1* for each component Z_{gt} .

[13] Practically one will consider the periodogram at the fourier frequencies only thus making irrelevant to demean the series by the sample mean.

[14] We refer to Robinson (1995) for a thorough discussion on the concepts and the roles played in the asymptotic theory by these two user-chosen numbers.

Then the OLS estimates of $c = (c_1, \dots, c_G)'$ and $d = (d_1, \dots, d_G)'$ are given by $\tilde{c}, \tilde{d}$

$$
\begin{aligned}
&\begin{bmatrix} \tilde{c} \\ \tilde{d} \end{bmatrix} = vec(Y'X(X'X)^{-1}) \,, \\
&X \stackrel{def}{=} (X_{l+1}, \dots, X_m)' \,, \; Y \stackrel{def}{=} (Y_1, \dots, Y_G)' \,, \\
&X_k \stackrel{def}{=} (1, -2ln\lambda_k)' \,, \; Y_k = (Y_{g,l+1}, \dots, Y_{g,m})' \,.
\end{aligned}
$$

Denoting as usual the OLS residuals as

$$
\tilde{U}_k = Y_k - \tilde{c} + \tilde{d}(2ln\lambda_k) \,, \; k = l+1, \dots, m \,, \tag{21}
$$

and the matrix of sample variances and covariances

$$
\tilde{\ } = \frac{1}{m-l} \sum_{i=l}^{m} \tilde{U}_k \tilde{U}_k', \tag{22}
$$

one gets that the OLS standard errors for $\tilde{d}_g$, $g = 1, \dots, G$ are given by the square root of the $(G+g)$th diagonal element of the matrix $(Z'Z)^{-1} \quad \tilde{\ }$.

This estimating procedure allows for cross equations restrictions such as that all (or some of) the G series are characterized by a common parameter of long memory that is

$$
d_g = \delta, \qquad g = 1, \dots, G \,,
$$

or in matrix formulation

$$
d = Q\delta,
$$

where $Q = (1, 1, \dots, 1)'$ is a $G \times 1$ vector and δ is a scalar representing the unknown common long memory parameter. Thus the GLS estimator $\hat{c}$ and $\hat{d}$ is given by

$$
\begin{bmatrix} \hat{c} \\ \hat{\delta} \end{bmatrix} = \left\{ \begin{bmatrix} I_G & 0 \\ 0 & Q' \end{bmatrix} (X'X \quad \tilde{\ }^{-1}) \begin{bmatrix} I_G & 0 \\ 0 & Q \end{bmatrix} \right\}^{-1} \begin{bmatrix} I_G & 0 \\ 0 & Q' \end{bmatrix} vec(\tilde{\ }^{-1} Y'X) \,.
$$

When there are no restriction we set $Q = I_G$ and we obtain again the OLS estimator[15].

[15] Also to obtain a consistent estimate of $C_g \; g = 1, \dots, G$ one has to consider the relation $C_g = exp(c_g - \psi(1))$.

Under certain regularity conditions (Robinson 1995) among which Gaussianity of the process Z_t the following asymptotic results are obtained, which allows to perform standard inference on the OLS and GLS estimators. For the OLS Robinson established

$$\begin{bmatrix} \frac{m^{1/2}}{ln\,n}(\tilde{c}-c) \\ 2m^{1/2}(\tilde{d}-d) \end{bmatrix} \rightarrow_d N\left(0, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \right), \tag{23}$$

and for the GLS

$$\begin{bmatrix} \frac{m^{1/2}}{ln\,n}(\hat{c}-c) \\ 2m^{1/2}(\hat{d}-d) \end{bmatrix} \rightarrow_d N\left(0, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad Q(Q' \quad ^{-1}Q)^{-1}Q'\right). \tag{24}$$

One obtains a consistent estimate of by using (22) . Considering each $\tilde{d}_g$ individually the general result in (23) becomes

$$2m^{1/2}(\tilde{d}_g - d_g) \rightarrow_d N(0, \frac{\pi^2}{6}).$$

The results allow us to make use of all the regression theory. In particular one can build a Wald test for linear restriction expressed by

$$H_0: \qquad Pd = 0,$$

where P is a $H \times G$ matrix of rank $H < G$. The test statistic is given by

$$\tilde{d}'P'[(0,P)\left\{(X'X)^{-1} \quad \tilde{} \right\}\begin{pmatrix} 0 \\ P' \end{pmatrix}]^{-1}P\tilde{d}, \tag{25}$$

that under H_0 is asymptotically distribuited like a central χ^2 with H degrees of freedom.

Estimating Procedure.

Firstly we detrend the data tting either a country speci c or a common trend. The former has been estimated with OLS, the latter with GLS. We then evaluate the order of integration of the residuals[16]. A preliminary analysis of the parameters d_g , $g = 1, \ldots, G$ gives estimated values greater than 1/2

[16]It is a reasonable question to ask if the properties of the theoretical disturbances carries over to the ones of the residuals after detrending the data with either the country speci c or the common trend (see i.e. Nelson and Kang 1981). There are good reasons to believe that it does once a semi-parametric frequency domain approach is undertaken. Nelson and Kang (1981) shows that the regression of a driftless random walk against a time trend delivers residuals exhibiting a periodogram with a single peak at a period equal to 0.83 of sample size thus asymptotically at frequency zero, as one would expect. In words, the memory of the process is entirely re ected in the residuals.

thus out of the admissible region for the asymptotic results to be valid. If we first differenced the data the estimates would be totally independent on the type of β-convergence we are considering (conditional and unconditional). In fact the periodogram evaluated at the Fourier frequencies is independent of any shift of location. For this reason we prefer to difference fractionally the data before estimating, by multipying them by $(1-L)^q$, $q = 0.5$. Obviously in doing so we have to approximate a series with finite sum. Our choice of $q = 0.5$ reflects the trade- off between differentiating "enough" (big q) to obtain estimates in the stationary region and minimizing the approximation from using a sum instead of a series (small q) [17]. To initialize the fractional filter $(1-L)^q$ we use the first 10 observations in the sample.

Choice of Trimming and Bandwidth.

In our application we choose a trimming coefficient equal to two, $l = 2$, so that we avoid the first periodogram ordinate (see also footnote 20). Unfortunately a complete theory for the optimal choice of the trimming and the bandwidth is still missing for this estimator but it seems that choosing a trimming bigger than one increases the performance of this estimator in finite samples (Hurvich and Ray 1994). Because of this reason we report estimates based on the the same criterium as the one used by Diebold and Rudebusch (1989) for their univariate analysis, that is $m = T^{0.525}$ after checking for robustness under alternative bandwidths [18]. It is nevertheless important to point out that the empirical results are very robust to changes in the the choice of the trimming and the bandwidth.

[17] Even if not formally proved, we follow the empirical literature of long memory processes conjecturing that asymptotically this approximation becomes negligible. Also the results are globally robust with respect to the choice of q.

[18] The results are available under requests. We defer to Beran (1994) for a review on parametric and semi non-parametric estimation in a long memory framework.

References

[1] Barro, J. (1991), Economic Growth in a Cross Section of Countries , *Quarterly Journal of Economics*, 2, pp. 407-443.

[2] Barro, J. and Sala-I-Martin, X. (1991), Convergence across States and Regions , *Brookings Paper on Economic Activity*, 1, pp. 107-182.

[3] Barro, J. and Sala-I-Martin, X. (1995), *Economic Growth*, McGraw-Hill, Advanced Series in Economics.

[4] Beran, J. (1994) *Statistics for Long Memory Processes*, Chapman and Hall.

[5] Bernard, A. and Durlauf, S. (1996), Interpreting Tests of the Convergence Hypothesis , *Journal of Econometrics*, 71(1-2), pp. 161-173.

[6] Brockwell P., and Davis, R. (1991), *Time Series: Theory and Methods*, Springer Verlag New York.

[7] Campell, J. and Mankiw, N. (1987), Are Output Fluctuations Transitory? , *Quarterly Journal of Economics*, 4, pp. 857-880.

[8] Chan, N. and Terrin, N. (1995), Inference for Unstable Long-Memory Processes with Applications to Fractional Unit Root Autoregressions , *Annals of Statistics,* Vol. 23 No. 5, pp. 1662-1683.

[9] Cochrane, J. (1988), How Big is the Random Walk in GNP? , *Journal of Political Economy*, 96, pp. 893-920.

[10] Coe, T. and Helpman, E. (1993), International *R&D* Spillovers , *NBER Working Paper n.4444*, August.

[11] Davidson, J. (1994) *Stochastic Limit Theory,* Oxford University Press 1994.

[12] Den Haan, W.J. (1995), Convergence in Stochastic Growth Models. The Importance of Understanding why Income Levels Differ , *Journal of Monetary Economics*, Vol.35 ,pp. 65 82.

[13] Dickey, D.A. and Fuller, W. (1979), Distribution of the Estimators for Autoregressive Time Series with a Unit Root , *Journal of American Statistical Association* , 74, pp. 427-431.

[14] Diebold, F. and Rudebusch, G. (1991), Long Memory and Persistence in Aggregate Output ,*Journal of Monetary Economics,* 24, pp. 189-209.

[15] Diebold, F. and Rudebusch, G. (1989), Long Memory and Persistence in Aggregate Output ,*Journal of Monetary Economics,* 24, pp. 189-209.

[16] Diebold, F. and Senhadji, A. (1996), Deterministic vs. Stochastic Trend in U.S. GNP, yet Again ,*American Economic Review, 86,*pp. 1291-1298.

[17] Durlauf, S.N. and Phillips, P.C.(1988) Trends Versus Random Walks in Time Series Analysis , *Econometrica*, Vol. 56-6, pp.1333-1354.

[18] Durlauf, S. (1991), Multiple Equilibria and Persistence in Aggregate Fluctuations , *American Economic Review Papers and Proceedings,* May, Vol. 81 no.2, pp. 70-74.

[19] Geweke, J. and Porter Hudak, S. (1983), The Estimation and Application of Long Memory Time Series Models , *Journal of Time Series Analisys*, 4, pp.221-238.

[20] Gil-Alana, L. and Robinson, P. M. (1997), Testing Unit Root and other Non-Stationary Hypotheses in Macroeconomic Time Series , *Journal of Econometrics,* Vol.80-2, October, pp.223-239.

[21] Gourieroux, C. and Monfort, A. (1989), *Statistics and Econometric Models*, Cambridge Press.

[22] Granger, C.W.J. (1966), The Typical Spectral Shape of an Economic Variable , *Econometrica* 34,no.1, pp. 150-161.

[23] Granger, C.W.J. (1980), Long Memory Relationships and the aggregation of Dynamic Models , *Journal of Econometrics* 14, pp. 227-238.

[24] Granger, C.W.J. and Joyeux, R. (1980), An Introduction to Long-Memory Time Series Models and Fractionally Differencing , *Journal of Time Series Analysis* Vol.1, no.1, pp. 15-29.

[25] Hurwich, C.M. and Ray, B.K. (1995), Estimation of the memory Parameter for Non-stationary or Non-invertible Fractionally Integrated Processes , *Journal of Time Series Analysis,* Vol.16, No.1, pp. 17-41.

[26] Jones,C. (1995a), Time Series Tests of Endogenous Growth Models , *Quarterly Journal of Economics,*110, pp. 495-525.

[27] Jones,C. (1995b), *R&D* Based Models of Economic Growth , *Journal of Political Economy*, 103-4, pp. 759-783.

[28] Levin, A. and Lin, C.F. (1992), Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties ,*Discussion Paper 92-92, UCSD.*September.

[29] Lippi, M. and Reichlin, L. (1992), Permanent and Transitory Components in Macroeconomics in Thygesen,N., Velupilai,K. and Zambelli, S. eds. *Business Cycles:Theories, Evidence and Analisys,* Macmillan Academic and Professional Ltd.

[30] Lippi, M. and Reichlin, L. (1994), Diffusion of Technical Change and the decomposition of output into Trend and Cycles , *Review of Economic Studies,* 61, pp.19-30.

[31] Maddison, A. (1995), Monitoring the World Economy 1820:1992 , *OECD, Paris.*

[32] Mankiw, N. Romer, R. and Weil, D. (1992), A Contribution to the Empirics of Economic Growth ,*Quarterly Journal of Economics,* 107-2, pp. 407-437.

[33] Nelson, C. and Kang, H. (1981), Spurious Periodicity in Inappropriately Detrended Time Series , *Econometrica*, Vol. 49-3, pp.741 751.

[34] Nelson, C. and Kang, H. (1984), Pitfalls in the Use of Time as an Explanatory Variable in Regression , *Journal of Business and Economic Statistics*, Vol. 2-1, pp. 73 82.

[35] Nelson, C. and Plosser, C. (1982), Trends and Randon Walks in Macro-economic Time Series: some Evidence and Implications , *Journal of Monetary Economics*,10, pp. 139-162.

[36] Priestley, M. (1965), Evolutionary Spectra and Non-Stationary Processes , *Journal of the Royal Statistical Association,* No. 27, pp. 204-237.

[37] Quah, D. (1992a), The relative Importance of Permanent and Transitory Components: Identi cation and some Theoretical Bounds , *Econometrica 60,* pp. 107-118.

[38] Quah, D. (1992b), International Patterns of Growth: Persistence in Cross-Country Disparities , *Working Paper, London School of Economics.*

[39] Quah, D. (1993a), Empirical Cross-Section Dynamics in Economic Growth , *European Economic Review,* 37-2/3, pp. 426-434.

[40] Quah, D. (1993b), Galton s fallacy and tests of the convergence hypothesis , *Scandinavian Journal of Economics,* 95(4), December, pp. 427-443.

[41] Quah, D. (1995), Empirics for Economic Growth and Convergence , *Discussion Paper 253, Centre for Economic Performance .*

[42] Robinson, P.M. (1978), Statistical Inference for a Random Coefficient Autoregressive Model , *Scandinavian Journal of Statistics* 5, pp. 163-168.

[43] Robinson, P.M. (1986), On a Model for a Time Series of Cross-Sections , *Journal of Applied Probability*,23A, pp. 113-126.

[44] Robinson, P. (1994), Time Series with Strong Dependence , in *Advances in Econometrics, Sixth World Congress*, ed. by C.A. Sims, Vol. 1, pp. 47- 95, Cambridge University Press.

[45] Robinson, P.M. (1995), Log-Periodogram Regression of Time Series with Long Range Dependence , *Annals of Statistics* , 23-3, pp. 1048-1072.

[46] Rudebusch, G. (1993), The Uncertain Unit Root in Real GNP , *American Economic Review,* 83, pp. 264-272.

[47] Rudin, W. (1973), *Functional Analysis*, McGraw-Hill New York.

[48] Seneta, E. (1976), *Regularly Varying Functions,* Springer, Berlin.

[49] Sowell, F. (1990), The Fractional Unit Root Distribution , *Econometrica,* 58-2, pp. 495-505.

[50] Yong, C. (1974), *Trigonometric Series*, Cambridge University Press.