ECONOMICS BULLETIN

Stochastic convergence among European economies

Mauro Costantini
University of Rome and ISAE

Claudio Lupi
University of Molise

Abstract

The aim of this paper is to test the stochastic convergence in real per capita GDP for 15 European countries using non−stationary panel data approaches over the period 1950−2003. Cross−sectional dependence is assumed due to the existence of strong linkages among European economies. However, tests derived under the assumption of cross−sectional independence are also carried out for completeness and comparison. We also split the whole sample into two sub−periods (1950−1976, 1977−2003) in order to take into account the effects of the first oil crisis (1973−1974) and to evaluate the robustness of the statistical analysis. Our results offer little support to the stochastic convergence hypothesis for the whole period, while suggest the presence of convergence in the first sub−period.

Citation: Costantini, Mauro and Claudio Lupi, (2005) "Stochastic convergence among European economies." *Economics Bulletin*, Vol. 3, No. 38 pp. 1−17
Submitted: April 22, 2005. **Accepted:** August 3, 2005.
URL: http://www.economicsbulletin.com/2005/volume3/EB−05C30008A.pdf

1. Introduction

Economic convergence/divergence issues have been much discussed in the literature in the last few decades. Several concepts of convergence have been proposed and different econometric methodologies have been employed. Cross-sections, time series, and panel data have been exploited in the empirical analyses. Most of the papers in this field tend to interpret the presence (absence) of convergence as a confirmation (falsification) of the validity of the neoclassical economic growth theory (Solow, 1956; Swan, 1956). Our viewpoint is that this interpretation is not granted and that this is possibly not even the crucial policy issue. We do not deny the value of theoretical models, we merely consider that econometrics and applied economics should have other scopes than merely confirming/refuting economic theories. Theoretical models are often too vaguely defined to be tested, so that theory falsification is often impossible in practice (Granger *et al.*, 1995; Granger, 1999). Even when theoretical models are precisely defined, their distance from reality may be so large that there might be a number of reasons why the theory can be refused on statistical grounds, but this fact alone does not necessarily entail that the theory has to be dismissed (Caldwell, 1982). Statistical tests can only answer very specific questions and only reject "sufficiently false" hypotheses: falsifying theories is a rather different matter (Keuzenkamp and Barten, 1995). Econometrics should probably have precise practical goals (Franses, 2002). It is not even obvious that the tested hypotheses are always the economically interesting ones (Summers, 1991). Theory testing might even have the perverse effect of becoming an obstacle to research progress. This may happen when theory testing hides potentially interesting observations to the researcher (Greenwald *et al.*, 1986). With reference to the neoclassical economic growth/convergence hypothesis, these problems are exacerbated by the fact that a number of alternative definitions have been offered for "convergence". In most cases, econometric tests cannot really discriminate among different hypotheses (see, *e.g.*, Durlauf, 2003). This is particularly true for definitions of convergence based on theoretical steady-state notions.

In the light of these considerations, we do not interpret our empirical results in direct connection with economic theory. Rather, we want to assess the presence of statistically verifiable stylized facts, without being too concerned with "deep parameters questions" (Summers, 1991). For us is more interesting to answer practical questions such as: "Are the European economies converging during the observed sample?". "Which kind of convergence, if any, do we observe?". To this end, a non-stationary panel data analysis is applied to real per capita GDP of the 15 European Union (EU) countries over the period 1950-2003.[1]

The literature on non-stationary panels include two distinct generations of tests (see, *e.g.*, Hurlin and Mignon, 2004). The first generation's papers assume that the cross-section units are independent each other (Levin and Lin, 1992 and 1993; Levin *et al.*, 2002 [LLC]; Im *et al.*,1997 and 2003 [IPS]; Maddala and Wu, 1999 [MW]; Choi, 2001). Due to the increasing empirical relevance of macro-panels and because of the evidence of co-movements in national business cycles (Backus and Kehoe, 1992), the second generation's tests relax the independence hypothesis and assume instead cross-sectional dependence (Choi, 2004 [CH]; Bai and Ng, 2004 [BNG]; Moon and Perron, 2004 [MP]; Pesaran, 2005 [PS]).

[1] Given the sample period considered in the empirical analysis and the lack of reliable time series for many countries, in this paper we do not consider the countries that joined the EU on May 1st, 2004.

In this paper, we apply both "independent" and "dependent" panel unit root tests, in order to compare the different results. Our empirical results convey little evidence of stochastic convergence among EU countries for the whole period 1950–2003, while suggest the presence of stochastic convergence only in the sub-period 1950–1976.

Throughout the paper i is the cross-section index ($i = 1, 2, \ldots, N$) and t is the time index ($t = 1, 2, \ldots, T$). "$\xrightarrow[\mathrm{H}_0]{d}$" denotes convergence in distribution under the null hypothesis.

The paper is organized as follows: Section 2 briefly describes the statistical framework and the econometric methods used in the empirical investigation. Section 3 reports the empirical results. Section 4 concludes.

2. Convergence and panel unit root tests

Traditional empirical tests of convergence broadly fall into two categories (Bernard and Durlauf, 1996). The first class of tests studies the cross-section correlation between initial per capita output and subsequent growth rates for a group of countries. A negative correlation is taken as evidence of convergence among countries, that is countries with low initial per capita output tend to grow faster than those with high initial per capita output (the so-called "β-convergence").[2] Since these tests assume homogeneous technological progress and speed convergence over time across individual units, the null hypothesis that is tested do not necessarily imply growth convergence or divergence in the neoclassical sense (Durlauf, 2003; Phillips and Sul, 2003). This class of tests has been also extended to panel data (Islam, 1995). However, the same sort of criticisms that apply to purely cross-section tests also apply to their panel versions.

The second class of tests studies the long-run behavior of differences in per capita output across countries in a time series framework (Bernard and Durlauf, 1995). In this approach, economic convergence implies that per capita GDP differences between two countries cannot contain stochastic trends (the so-called "stochastic convergence"). Because of the well known lack of power of univariate time series unit root tests, this approach has been lately extended using panel unit root tests (Evans and Karras, 1996; Bernard and Jones, 1996; Fleissig and Strauss, 2001).

However, both the two classes of tests fail to test the null hypothesis of economic convergence in the sense implied by the neoclassical economic growth theory. What they do test is the average diminution of cross-country disparities over a fixed time period (β-convergence) or the existence of an "equilibrium" relation among real per capita GDPs over an observed sample (stochastic convergence). Indeed, they have little to say about steady state convergence, any conclusion about this aspect being conditional on strong identifying assumptions.

In our analysis we adhere to Bernard and Durlauf's (1995) definition of stochastic convergence:

$$\lim_{t\to\infty} \mathrm{E}\left(\mathcal{Y}_{b,0+t} - \mathcal{Y}_{i,0+t} | \mathcal{I}_0\right) = 0 \tag{1}$$

where $\mathcal{Y}_{j,\tau}$ is real per capita output of country j at time τ, and $\mathcal{I}_0$ denotes the information set as of time 0. In other words, countries b and i converge if the long-term forecast of output

[2] For empirical applications see, among others, Barro, 1991; Barro and Sala-i-Martin, 1992; Mankiw *et al.* 1992.

for both countries are equal at fixed time. In a panel of countries stochastic convergence occurs if the difference, y_{it}, between the real per capita GDP of the "benchmark" country and that of each other country in the panel follows a zero-mean stationary process. Indeed, this difference being $I(1)$ in the observed sample is compatible with a number of alternative growth theories. The reasons why the presence of unit root components cannot be rejected may be completely different each other, and might even be related to the fact that we are observing the sample over a transition period *towards* equilibrium. What we want to test is simply if European national per capita GDPs tend to evolve roughly along similar "equilibrium" paths.

In this paper, the benchmark country is Germany. To test for stochastic convergence, we apply different panel unit root tests. In particular we consider three tests based on the cross-sectional independence hypothesis (MW, LLC, and IPS) and four cross-sectional dependent tests (CH, BNG, MP, and PS).

MW propose a new simple test based on Fisher's (1932) suggestion of combining the p-values p_i from individual Augmented Dickey-Fuller (ADF) tests. Under the unit root null hypothesis and the ancillary hypothesis of cross-sectional independence of the errors terms from the individual ADF equations, the panel unit root test statistic

$$\lambda = -2 \sum_{i=1}^{N} \log(p_i) \tag{2}$$

asymptotically has a chi-square distribution with $2N$ degrees of freedom, when $T \to \infty$ and N is fixed.

LLC formulate a panel unit root test procedure which consists of three steps. In the first step, the ADF regressions for each individual in the panel are carried out:

$$\Delta y_{it} = \delta_i y_{i,t-1} + \sum_{j=1}^{k_i} \theta_{ij} \Delta y_{i,t-j} + \alpha_{mi} d_{mt} + \varepsilon_{it} \tag{3}$$

where d_{mt} denotes the vector of deterministic variables and α_{mi} indicate the corresponding vector of coefficients for the specific model m ($m \in \{1, 2, 3\}$).[3] After having determined the order of the ADF regression, LLC run two auxiliary regressions of Δy_{it} and $y_{i,t-1}$ against $\Delta y_{i,t-j}$ (with $j = 1, \ldots, k_i$), and generate two orthogonolized residuals, $\hat{e}_{it}$ and $\hat{\nu}_{it}$. To control for heterogeneity across individuals, LLC derive the normalized residuals $\tilde{e}_{it}$ and $\tilde{\nu}_{it}$ by dividing by the standard error from equation (3): $\tilde{e}_{it} = \hat{e}_{it}/\hat{\sigma}_{\varepsilon i}$ and $\tilde{\nu}_{i,t-1} = \hat{\nu}_{i,t-1}/\hat{\sigma}_{\varepsilon i}$. The second step requires estimating the ratio of the long run to short run innovation standard deviation, $s_i = \sigma_{yi}/\sigma_{\varepsilon i}$, for each individual by using $\hat{s}_i = \hat{\sigma}_{yi}/\hat{\sigma}_{\varepsilon i}$. Finally, the pooled t-statistic is computed:

$$t^*_\rho = \frac{t_\rho - N\tilde{T}\hat{S}_N \hat{\sigma}^{-2}_{\tilde{\varepsilon}} STD(\hat{\rho}) \mu^*_{m\tilde{T}}}{\sigma^*_{m\tilde{T}}} \xrightarrow[\mathrm{H_0}]{d} \mathrm{N}(0,1) \quad \text{as } \tilde{T}, N \to \infty \tag{4}$$

where t_ρ is the t-statistic in the regression

$$\tilde{e}_{it} = \rho \tilde{\nu}_{it-1} + \tilde{\varepsilon}_{it}, \tag{5}$$

[3] The models are identified as follows: $m = 1$ denotes an ADF with no constant and trend; $m = 2$ indicates an ADF with the constant term; $m = 3$ denotes an ADF with constant and trend.

$\hat{S}_N$ is the estimated average standard deviation ratio, $\hat{S}_N = N^{-1}\sum_{i=1}^{N}\hat{s}_i$, $\mu^*_{m\tilde{T}}$ and $\sigma^*_{m\tilde{T}}$ are the mean and the standard deviation adjustments, and $STD(\hat{\rho}) = \hat{\sigma}_{\tilde{\varepsilon}}\left[\sum_{i=1}^{N}\sum_{t=2+p_i}^{T}\tilde{\nu}^2_{i,t-1}\right]^{-1/2}$. $\tilde{T} = T - \bar{k} - 1$ is the average number of observations per individual in the panel, with $\bar{k} = N^{-1}\sum_{i=1}^{N}k_i$ being the average lag order for the individual ADF regressions.

IPS propose a test based on the average of the ADF statistics computed for each individual in the panel. The IPS test is based on

$$\Delta y_{it} = \alpha_i + \beta_i y_{i,t-1} + \sum_{j=1}^{k}\delta_{ij}\Delta y_{i,t-1} + \xi_{it}\,. \tag{6}$$

The null hypothesis of a unit root can be now defined as $\mathrm{H}_0 : \beta_i = 0$ for all i against the alternatives $\mathrm{H}_1 : (\beta_i < 0,\ i = 1, 2, \ldots, N_1 < N$ and $\beta_j = 0,\ j = N_1 + 1, N_1 + 2, \ldots, N)$. Under this formulation of the alternative hypothesis β_i may differ across cross-sectional units. Therefore, the IPS test evaluates the null hypothesis that all the series contain a unit root against the alternative that some of the series are stationary. The IPS test simply uses the average of the N ADF individual t-statistics, $\breve{t}_{iT}$

$$\bar{t}_{NT} = N^{-1}\sum_{i=1}^{N}\breve{t}_{iT} \tag{7}$$

from which it is possible to write the standardized statistic

$$Z_{\bar{t}} = \frac{\sqrt{N}[\bar{t}_{NT} - \mathrm{E}(\bar{t}_T)]}{[\mathrm{var}(\bar{t}_T)]^{\frac{1}{2}}} \xrightarrow[\mathrm{H}_0]{d} \mathrm{N}(0,1) \quad \text{as } N \to \infty \tag{8}$$

where $E(\bar{t}_T)$ and $\mathrm{var}(\bar{t}_T)$ are respectively the theoretical mean and variance of $\bar{t}_{NT}$.[4] The finite sample critical values and p-values can be computed by bootstrap.

CH propose new panel unit root tests for cross-sectionally correlated panels. The cross-sectional correlation is modelled by error-component models. The test statistics are derived from combining p-values from the ADF test applied to each time series whose non-stochastic trend components and cross-correlations are eliminated by Elliott *et al.*'s (1996) GLS-based de-trending and the conventional cross-sectional demeaning panel data procedure. The proposed panel unit root tests are:

$$P_m = -\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\log(p_i + 1) \xrightarrow[\mathrm{H}_0]{d} \mathrm{N}(0,1) \ \text{ as } T, N \to \infty \tag{9}$$

$$Z = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Phi^{-1}(p_i) \xrightarrow[\mathrm{H}_0]{d} \mathrm{N}(0,1) \ \text{ as } T, N \to \infty \tag{10}$$

$$L^* = \frac{1}{\sqrt{\frac{\pi^2 N}{3}}}\sum_{i=1}^{N}\log\left(\frac{p_i}{1 - p_i}\right) \xrightarrow[\mathrm{H}_0]{d} \mathrm{N}(0,1) \ \text{ as } T, N \to \infty \tag{11}$$

[4] If the individual estimates of $\breve{t}_{iT}$ were unbiased, $\mathrm{E}(\bar{t}_{iT})$ would be zero. However, $\mathrm{E}(\bar{t}_{iT})$ and $\mathrm{var}(\bar{t}_T)$ can be computed by simulation.

where P_m test is a modification of Fisher (1932) inverse chi-square, $\Phi(\cdot)$ is the standard normal cumulative distribution function and p_i indicates the asymptotic p-value of the ADF-GLS test for the i-th unit.

BNG consider instead the factor model

$$y_{it} = D_{it} + \boldsymbol{\lambda}_i^{'} \boldsymbol{F}_t + e_{it} \tag{12}$$

where D_{it} is a polynomial trend function, $\boldsymbol{F}_t$ is an $r \times 1$ vector of common factors, and $\boldsymbol{\lambda}_i$ is a vector of factor loadings. Therefore, y_{it} is decomposed into three components: a deterministic one, a common component with factor structure and an idiosyncratic error component. The process for y_{it} may be non-stationary if one or more of the common factors are non-stationary, or the idiosyncratic error is non-stationary, or both. To test the stationarity of the idiosyncratic component, BNG propose pooling the individual ADF t-statistics computed over the de-factored estimated components $\hat{e}_{it}$ using the model with no deterministic trend, namely

$$\Delta \hat{e}_{it} = \delta_{i0} \hat{e}_{i,t-1} + \sum_{j=1}^{k_i} \delta_{i,j} \Delta \hat{e}_{i,t-j} + u_{it}. \tag{13}$$

Let $ADF_{\hat{e}}^c(i)$ be the ADF t-statistic from (10) for the i-th cross-section unit. The asymptotic distribution of the $ADF_{\hat{e}}^c(i)$ coincides with the Dickey-Fuller distribution for the case of no-constant. However, these individual time series tests have the same low power as those based on the initial series. BNG propose using pooled tests based on Fisher's type statistics defined as in Choi (2001) and Maddala and Wu (1999). Let $P_{\hat{e}}^c(i)$ be the p-value of the $ADF_{\hat{e}}^c(i)$. The test statistics are

$$Z_{\hat{e}}^c = \frac{-\sum_{i=1}^N \log P_{\hat{e}}^c(i) - N}{\sqrt{N}} \xrightarrow[\mathrm{H}_0]{d} \mathrm{N}(0,1) \quad \text{as } T, N \to \infty \tag{14}$$

and

$$P_{\hat{e}}^c = -2 \sum_{i=1}^N \log P_{\hat{e}}^c(i) \xrightarrow[\mathrm{H}_0]{d} \chi^2(2N) \quad \text{as } T \to \infty\,. \tag{15}$$

MP also develop several unit root tests in which the cross-sectional units are correlated. To model cross-sectional dependence, MP provide an approximate linear dynamic factor model in which the panel data are generated by both idiosyncratic shocks and unobservable dynamic factors that are common to all individual units but to which each individual reacts heterogeneously. In particular, in our analysis, we apply the following tests:

$$t_a^* = \frac{\sqrt{N} T (\hat{\rho}_{pool}^+ - 1)}{\sqrt{\frac{2\hat{\phi}_e^4}{\hat{\omega}_e^4}}} \xrightarrow[\mathrm{H}_0]{d} \mathrm{N}(0,1) \quad \text{as } T, N \to \infty \tag{16}$$

$$t_b^* = \sqrt{N} T (\hat{\rho}_{pool}^+ - 1) \sqrt{\frac{1}{NT^2} \operatorname{tr}(\boldsymbol{Y}_{-1} \boldsymbol{Q}_B \boldsymbol{Y}_{-1}') \frac{\hat{\omega}_e^2}{\hat{\phi}_e^4}} \xrightarrow[\mathrm{H}_0]{d} \mathrm{N}(0,1) \quad \text{as } T, N \to \infty \tag{17}$$

where $\hat{\rho}_{pool}^+$ is the bias-corrected version of the OLS pooled autoregressive parameter $\hat{\rho}_{pool} = \operatorname{tr}(\boldsymbol{Y}_{-1}' \boldsymbol{Y}) / \operatorname{tr}(\boldsymbol{Y}_{-1}' \boldsymbol{Y}_{-1})$. $\boldsymbol{Y}$ is the matrix of observations on y_{it}: $\boldsymbol{Y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N)$, with

$\boldsymbol{y}_i = (y_{i1}, \ldots, y_{iT})'$. $\boldsymbol{Y}_{-1}$ is the matrix of the corresponding lagged values. $\hat{\omega}^4_{\hat{e}}$ and $\hat{\phi}^4_{\hat{e}}$ are the estimates of the cross sectional average of the long run variance of residual $\hat{e}_{it}$ and the cross sectional average of $\omega^4_{e,i}$, respectively. $\boldsymbol{Q}_B$ is the projection matrix used to eliminate the common factor in the panel.[5]

To deal with the problem of cross-sectional dependence PS does not consider the deviations from the estimated common factors. Instead, he suggests augmenting the standard DF (or ADF) regression with the cross section averages of lagged levels and first-differences of the individual series. If the residual are not serially correlated, the regression used for the i-th cross-section unit is defined as:

$$\Delta y_{it} = \alpha_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_i \Delta \bar{y}_{t-1} + e_{it}, \tag{18}$$

where

$$\bar{y}_{t-1} = N^{-1} \sum_{i=1}^{N} y_{i,t-1},$$

$$\Delta \bar{y}_t = N^{-1} \sum_{i=1}^{N} \Delta y_{i,t}.$$

Let $t_i(N,T)$ be the t-statistic of the OLS estimate of b_i. The panel unit root tests are then based on the average of individual cross-sectionally augmented ADF statistics (CADF). PS presents also a truncated version of the test, denoted CADF*, in order to avoid undue influences of extreme outcomes that could emerge in the case of small T. In both cases, the individual CADF or CADF* statistics may be used to build modified versions of the IPS $\bar{t}_{NT}$ test (CIPS and CIPS* for cross-sectionally augmented IPS):

$$CIPS(N,T) = N^{-1} \sum_{i=1}^{N} t_i(N,T) \xrightarrow[\mathrm{H_0}]{d} \mathrm{N}(0,1) \text{ as } N,T \to \infty \tag{19}$$

$$CIPS^*(N,T) = N^{-1} \sum_{i=1}^{N} t_i^*(N,T) \xrightarrow[\mathrm{H_0}]{d} \mathrm{N}(0,1) \text{ as } N,T \to \infty \tag{20}$$

where the truncated CADF statistics are defined as

$$t_i^*(N,T) = \begin{cases} -K_1, & \text{if } t_i(N,T) \leq -K_1, \\ K_2, & \text{if } t_i(N,T) \geq K_1, \\ t_i(N,T), & \text{if } -K_1 < t_i(N,T) < K_2 \end{cases}$$

with K_1 and K_2 positive constants such that $\Pr\left(t_i(N,T) \in [K_1, K_2]\right)$ is close to 1.

3. Empirical results

The data used in this paper are annual log real per capita GDP in 1999 EKS dollars over the period 1950–2003. The data are taken from the Gröningen Growth and Development Center data base. The EU countries considered here are Austria, Belgium, Denmark,

[5]For further details see Moon and Perron, 2004.

Table 1: **Independent panel unit root tests: 1950-2003**

LLC panel test (a)			
		asymptotic	bootstrap
k	t_{ρ}^{*}	p-values	p-values
1	-0.956	0.169	0.274
2	-1.067	0.143	0.250
IPS panel test (b)			
		asymptotic	bootstrap
k	$Z_{\bar{t}}$	p-values	p-values
1	-1.508	0.066	0.174
2	-1.201	0.115	0.956
MW panel test (c)			
		asymptotic	bootstrap
k	λ	p-values	p-values
1	50.848	0.005	0.333
2	28.187	0.403	0.997

Notes:

k denotes lag lengths.The bootstrap procedure is described in the main text.

(a) Derived on the basis of equation (4) in the main text.

(b) Derived on the basis of equation (8) in the main text.

(c) Derived on the basis of equation (2) in the main text.

Table 2: **Independent panel unit root tests: 1950-1976**

LLC panel test (a)			
		asymptotic	bootstrap
k	t^*_ρ	p-values	p-values
1	-1.975	0.024	0.156
2	-1.796	0.036	0.272
IPS panel test (b)			
		asymptotic	bootstrap
k	$Z_{\bar{t}}$	p-values	p-values
1	-2.456	0.005	0.012
2	-2.031	0.021	0.014
MW panel test (c)			
		asymptotic	bootstrap
k	λ	p-values	p-values
1	50.482	0.000	0.000
2	29.168	0.431	0.345

Notes:

k denotes lag lengths. The bootstrap procedure is described in the main text.

(a) Derived on the basis of equation (4) in the main text.

(b) Derived on the basis of equation (8) in the main text.

(c) Derived on the basis of equation (2) in the main text.

Table 3: **Independent panel unit root tests: 1977-2003**

LLC panel test (a)			
k	t_ρ^*	asymptotic p-values	bootstrap p-values
1	-1.139	0.783	0.372
2	-1.140	0.929	0.250
IPS panel test (b)			
k	$Z_{\bar{t}}$	asymptotic p-values	bootstrap p-values
1	-1.508	0.066	0.174
2	-1.201	0.115	0.256
MW panel test (c)			
k	λ	asymptotic p-values	bootstrap p-values
1	3.988	0.987	0.921
2	3.465	0.998	0.891

Notes:

k denotes lag lengths. The bootstrap procedure is described in the main text.

(a) Derived on the basis of equation (4) in the main text.

(b) Derived on the basis of equation (8) in the main text.

(c) Derived on the basis of equation (2) in the main text.

Table 4: **Dependent panel unit root tests: 1950-2003**

CH panel tests (a)				
	P_m	Z	L^*	
	6.394 (0.000)	-3.131 (0.000)	-4.035 (0.000)	
MP panel tests (b)				
$\hat{r}$	t_a^*	t_b^*	t_a^{*B}	t_b^{*B}
1	-6.655 (0.000)	-5.151 (0.000)	-6.712 (0.000)	-5.202 (0.000)
BNG panel tests (c)				
$\hat{r}$	$Z_{\hat{e}}^c$	$P_{\hat{e}}^c$	$ADF_{\hat{F}}^c$	
1	-2.786 (0.300)	17.149 (0.934)	-0.295 (0.920)	
PS panel tests (d)				
$p*$	$CIPS$	$CIPS^*$		
1	-1.707 (0.430)	-1.707 (0.430)		
2	-1.532 (0.795)			
3	-1.553 (0.775)			

Notes:

The numbers in parenthesis are p-values.

(a) The tests are based on equations (9)-(11) in the main text. Lag selection in the ADF-GLS tests is determined using the BIC. (b) The tests are based on equations (16) and (17) in the main text. $\hat{r}$ is the number of common factors that in the MP test is assumed *a priori*. The t_a^* and t_b^* tests are based on de-factored data and computed using a quadratic spectral (QS) kernel. The t_a^{*B} and t_b^{*B} tests are computed using a Barlett (B) kernel. The optimal truncation parameter q_i for the QS kernel is $q_i = 1.3221[4\hat{\rho}_{i,1}^2 T_i/(1-\hat{\rho_{i,1}^2})^4]^{1/5}$, where $\hat{\rho}_{i,1}^2$ is the first-order autocorrelation estimate of the individual component $\hat{e}_{i,t}$ of unit i. The optimal truncation parameter for the B kernel is $q_i = 1.447[4\hat{\rho}_{i,1}^2 T_i/(1-\hat{\rho_{i,1}^2})^2]^{1/3}$. In both cases, the bandwidth parameters are chosen according to the Newey and West (1994) procedure and are set equal to $\lfloor 4(T/100)^{2/9} \rfloor$. ($c$) The tests are based on equations (14) and (15). For each variable, the number of estimated common factors ($\hat{r}$) is determined by the the modified BIC (BIC3) proposed by Bai and NG (2002), assuming a maximum number of factors equal to 5. For idiosyncratic components $\hat{e}_{it}$, the pooled unit root statistic test are reported. $P_{\hat{e}}^c$ is a Fisher's type statistic based on a p-valued of the individual ADF tests. Under the null hypothesis, $P_{\hat{e}}^c$ has a $\chi^2(2N)$ distribution when T tends to infinity and N is fixed. $Z_{\hat{e}}^c$ is the standardized Choi's type test statistic. Under the null hypothesis, $Z_{\hat{e}}^c$ has a $\mathrm{N}(0,1)$ distribution. For the $\hat{F}_t$ component, two different cases must be distinguished. If $\hat{r}=1$, only the standard ADF t-statistic, $ADF_{\hat{F}}^c$, is reported; if $\hat{r}>1$ the estimated number of independent stochastic trends in the common factors are reported ($\hat{r}_1(MQ_c)$ and $\hat{r}_1(MQ_f)$). $\hat{r}_1(MQ_c)$ is derived by correcting for serial correlation of arbitrary form non-parametrically; $\hat{r}_1(MQ_f)$ filters the factors under the assumption that they have a finite order VAR representation. These two tests are modified versions of the Q_c and Q_f tests developed by Stock and Watson (1998). (d) The CIPS test is the mean of individual cross-sectionally augmented ADF statistics (ADF) and is based on equation (19). $CIPS^*$ indicates the mean of truncated individual CADF statistics and it is based on equation (20). The truncated statistics are reported only for one lag since they are always equal to the non-truncated ones for higher lag lengths. p^* denotes the nearest integer of the mean of the individual lag lengths in ADF tests.

Table 5: **Dependent panel unit root tests: 1950-1976**

CH panel tests (a)				
	P_m	Z	L^*	
	3.596 (0.000)	-2.156 (0.000)	-2.743 (0.000)	
MP panel tests (b)				
$\hat{r}$	t_a^*	t_b^*	t_a^{*B}	t_b^{*B}
1	-11.775 (0.000)	-6.271 (0.000)	-12.290 (0.000)	-6.488 (0.000)
BNG panel tests (c)				
$\hat{r}$	$Z_{\hat{e}}^c$	$P_{\hat{e}}^c$	$\hat{r}_1(MQ_c)$	$\hat{r}_1(MQ_f)$
2	1.667 (0.041)	40.472 (0.047)	2	2
PS panel tests (d)				
$p*$	$CIPS$	$CIPS^*$		
1	-1.795 (0.430)	-1.795 (0.430)		
2	-1.485 (0.790)			
3	-1.296 (0.920)			

Notes:

The numbers in parenthesis are p-values.

(a) The tests are based on equations (9)-(11) in the main text. Lag selection in the ADF-GLS tests is determined using the BIC. (b) The tests are based on equations (16) and (17) in the main text. $\hat{r}$ is the number of common factors that in the MP test is assumed *a priori*. The t_a^* and t_b^* tests are based on de-factored data and computed using a quadratic spectral (QS) kernel. The t_a^{*B} and t_b^{*B} tests are computed using a Barlett (B) kernel. The optimal truncation parameter q_i for the QS kernel is $q_i = 1.3221[4\hat{\rho}_{i,1}^2 T_i/(1-\hat{\rho_{i,1}^2})^4]^{1/5}$, where $\hat{\rho}_{i,1}^2$ is the first-order autocorrelation estimate of the individual component $\hat{e}_{i,t}$ of unit i. The optimal truncation parameter for the B kernel is $q_i = 1.447[4\hat{\rho}_{i,1}^2 T_i/(1-\hat{\rho_{i,1}^2})^2]^{1/3}$. In both cases, the bandwidth parameters are chosen according to the Newey and West (1994) procedure and are set equal to $\lfloor 4(T/100)^{2/9} \rfloor$. (c) The tests are based on equations (14) and (15). For each variable, the number of estimated common factors $(\hat{r})$ is determined by the the modified BIC (BIC3) proposed by Bai and NG (2002), assuming a maximum number of factors equal to 5. For idiosyncratic components $\hat{e}_{it}$, the pooled unit root statistic test are reported. $P_{\hat{e}}^c$ is a Fisher's type statistic based on a p-valued of the individual ADF tests. Under the null hypothesis, $P_{\hat{e}}^c$ has a $\chi^2(2N)$ distribution when T tends to infinity and N is fixed. $Z_{\hat{e}}^c$ is the standardized Choi's type test statistic. Under the null hypothesis, $Z_{\hat{e}}^c$ has a $\mathrm{N}(0,1)$ distribution. For the $\hat{F}_t$ component, two different cases must be distinguished. If $\hat{r}=1$, only the standard ADF t-statistic, $ADF_{\hat{F}}^c$, is reported; if $\hat{r}>1$ the estimated number of independent stochastic trends in the common factors are reported ($\hat{r}_1(MQ_c)$ and $\hat{r}_1(MQ_f)$). $\hat{r}_1(MQ_c)$ is derived by correcting for serial correlation of arbitrary form non-parametrically; $\hat{r}_1(MQ_f)$ filters the factors under the assumption that they have a finite order VAR representation. These two tests are modified versions of the Q_c and Q_f tests developed by Stock and Watson (1998). (d) The CIPS test is the mean of individual cross-sectionally augmented ADF statistics (ADF) and is based on equation (19). $CIPS^*$ indicates the mean of truncated individual CADF statistics and it is based on equation (20). The truncated statistics are reported only for one lag since they are always equal to the non-truncated ones for higher lag lengths. p^* denotes the nearest integer of the mean of the individual lag lengths in ADF tests.

Table 6: **Dependent panel unit root tests: 1977-2003**

CH panel tests (a)				
	P_m	Z	L^*	
	-2.792	3.657	3.908	
	(0.993)	(0.997)	(0.932)	
MP panel tests (b)				
$\hat{r}$	t_a^*	t_b^*	t_a^{*B}	t_b^{*B}
1	-11.982	-5.106	-11.541	-5.282
	(0.000)	(0.000)	(0.000)	(0.000)
BNG panel tests (c)				
$\hat{r}$	$Z_{\hat{e}}^c$	$P_{\hat{e}}^c$	$\hat{r}_1(MQ_c)$	$\hat{r}_1(MQ_f)$
2	0.524	17.149	2	2
	(0.300)	(0.934)		
PS panel tests (d)				
$p*$	$CIPS$	$CIPS^*$		
1	-1.717	-1.717		
	(0.520)	(0.520)		
2	-1.153			
	(0.970)			
3	-1.325			
	(0.905)			

Notes:

The numbers in parenthesis are p-values.

(a) The tests are based on equations (9)-(11) in the main text. Lag selection in the ADF-GLS tests is determined using the BIC. (b) The tests are based on equations (16) and (17) in the main text. $\hat{r}$ is the number of common factors that in the MP test is assumed *a priori*. The t_a^* and t_b^* tests are based on de-factored data and computed using a quadratic spectral (QS) kernel. The t_a^{*B} and t_b^{*B} tests are computed using a Barlett (B) kernel. The optimal truncation parameter q_i for the QS kernel is $q_i = 1.3221[4\hat{\rho}_{i,1}^2 T_i/(1-\hat{\rho_{i,1}^2})^4]^{1/5}$, where $\hat{\rho}_{i,1}^2$ is the first-order autocorrelation estimate of the individual component $\hat{e}_{i,t}$ of unit i. The optimal truncation parameter for the B kernel is $q_i = 1.447[4\hat{\rho}_{i,1}^2 T_i/(1-\hat{\rho_{i,1}^2})^2]^{1/3}$. In both cases, the bandwidth parameters are chosen according to the Newey and West (1994) procedure and are set equal to $\lfloor 4(T/100)^{2/9} \rfloor$. ($c$) The tests are based on equations (14) and (15). For each variable, the number of estimated common factors ($\hat{r}$) is determined by the the modified BIC (BIC3) proposed by Bai and NG (2002), assuming a maximum number of factors equal to 5. For idiosyncratic components $\hat{e}_{it}$, the pooled unit root statistic test are reported. $P_{\hat{e}}^c$ is a Fisher's type statistic based on a p-valued of the individual ADF tests. Under the null hypothesis, $P_{\hat{e}}^c$ has a $\chi^2(2N)$ distribution when T tends to infinity and N is fixed. $Z_{\hat{e}}^c$ is the standardized Choi's type test statistic. Under the null hypothesis, $Z_{\hat{e}}^c$ has a $\mathrm{N}(0,1)$ distribution. For the $\hat{F}_t$ component, two different cases must be distinguished. If $\hat{r}=1$, only the standard ADF t-statistic, $ADF_{\hat{F}}^c$, is reported; if $\hat{r}>1$ the estimated number of independent stochastic trends in the common factors are reported ($\hat{r}_1(MQ_c)$ and $\hat{r}_1(MQ_f)$). $\hat{r}_1(MQ_c)$ is derived by correcting for serial correlation of arbitrary form non-parametrically; $\hat{r}_1(MQ_f)$ filters the factors under the assumption that they have a finite order VAR representation. These two tests are modified versions of the Q_c and Q_f tests developed by Stock and Watson (1998). (d) The CIPS test is the mean of individual cross-sectionally augmented ADF statistics (ADF) and is based on equation (19). $CIPS^*$ indicates the mean of truncated individual CADF statistics and it is based on equation (20). The truncated statistics are reported only for one lag since they are always equal to the non-truncated ones for higher lag lengths. p^* denotes the nearest integer of the mean of the individual lag lengths in ADF tests.

France, Finland, Germany, Greece, Ireland, Italy, Luxembourg, Netherland, Portugal, Spain, Sweden, and the United Kingdom.

The null hypothesis is that European economies do not converge stochastically or, in other words, that the pairwise differences of national GDPs with the benchmark GDP (Germany) contain a unit root. However, under the alternative we do not require the differences to be zero-mean stationary processes. In economic terms this implies that national GDPs may show “similar” growth paths without necessarily being on average at the same level.

Whenever possible, in this paper we use both asymptotic and bootstrap-based critical values. In order to estimate the bootstrap distribution under the null, for the tests derived under the hypothesis of cross-sectional independence we consider the data generating process (DGP)

$$\Delta y_{it} = \mu_i + \sum_{j=1}^{k_i} \delta_{ij} \Delta y_{i,t-j} + \epsilon_{it}. \tag{21}$$

The parametric bootstrap distributions for the test statistics are derived as follows: *a*) a sequence of length $T + R$ for the innovation vector $\tilde{\boldsymbol{\varepsilon}} \sim \mathrm{N}(0, \hat{\boldsymbol{\Sigma}})$ is drawn, where $\hat{\boldsymbol{\Sigma}} = T^{-1} \sum_{t=1}^{T} \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t'$ and $\hat{\boldsymbol{\varepsilon}}_t = (\hat{\epsilon}_{1t}, \ldots, \hat{\epsilon}_{Nt})'$ is the vector of OLS residuals from (21); *b*) pseudo-observations $\tilde{y}_{it}$ ($i = 1, \ldots, N,\ t = 1, \ldots, T + R$) are computed recursively, according to equation (21) using $\hat{\boldsymbol{\varepsilon}}_t$ and the estimated values of the coefficients $\hat{\mu}_i$ and $\hat{\delta}_{ij}$; *c*) the first R pseudo-observations are dropped and the test statistics are computed on the remaining ones; *d*) the procedure is repeated over 2000 replications, obtaining the desired distribution under the null.

Estimation details (lags, asymptotic p-values, etc.) are reported in the tables of results for each single test. However, as far as the independent panel tests are concerned, following standard practice, results for different numbers of lags are presented. Given that we are dealing with annual data, the maximum lag length has been fixed to 2. This choice is justified by the need to cope with the trade-off existing between size distortion and power (see *e.g.* MW). Most of the results of the independent panel tests presented in the tables are invariant with respect to lag selection, especially so for the bootstrap-based tests. As far as bandwidth parameters selection in dependent panel unit root tests is concerned, the Newey and West (1994) procedure is used and bandwidth is set equal to the largest integer less than $4(T/100)^{2/9}$.

Table 1 reports the independent panel unit roots for the entire sample 1950-2003. The null hypothesis of no stochastic convergence cannot be rejected in all cases, except for the MW test with $k = 1$. However, the simulated p-values confirm the evidence of unit roots in the differences between the benchmark GDP and that of the other countries also for this case.

Table 2 reports the results for the first sub-period (1950-1976) only. Fairly strong evidence of stochastic convergence is found. Only the bootstrap-based LLC test is not significant at conventional confidence levels.

Table 3 summarizes the results for the second sub-period (1977-2003). The null of no stochastic convergence cannot be rejected for all the panel unit root tests. None of the asymptotic and simulated p-values are significant at the 5% level.

Of course, the tests reported in tables 1-3 are subject to the criticism that they do not consider cross-sectional dependence. Therefore we turn now to panel unit root tests that are valid also when the panel units are not independent each other.

Table 4 provides the results for cross-sectional dependent panel unit root computed considering the entire temporal sample. The statistics for the CH and MP tests support evidence of stochastic convergence, while the opposite conclusion should be drawn from the BNG and PS tests.[6]

In Table 5 cross-sectional dependent panel unit-root results for the first sub-period are reported. In this case, there is a strong evidence of stochastic convergence among European Countries. These results are statistically significant, except for CIPS and CIPS* tests.

Table 6 describes panel unit-root results for the second sub-period. Here no stochastic convergence is found. Only the MP tests are statistically significant.

4. Conclusions

The study of stochastic convergence has received much attention in the last decade, after the publication of Bernard and Durlauf's (1996) seminal paper. Most of the empirical analyses have tested stochastic convergence in international economic contexts. This paper apply independent and cross-sectional dependent panel unit-root tests to evaluate stochastic convergence among the EU countries over the period 1950-2003. We also split the whole sample period into two sub-periods (1950-1976, 1977-2003) in order to test the robustness of our results to changes of the sample. The choice of the break is somewhat arbitrary, but we think that the proposed date is reasonable for different reasons. First, the break is just a few years after the first oil shock. Second, in 1973 Denmark, Ireland and the UK joined the European community, so that the number of countries linked by close partnership became nine after that date. Last, but not least, we needed two sub-samples of approximately the same dimension.

We do not intend to test if any particular growth model is consistent with available data, nor we want to derive asymptotic properties of European economic growth in the far future. We believe that these goals cannot be pursued with the available econometric techniques. We confine ourselves in studying the observable features of the growth process in European countries. In particular, we test for the presence of a unit root in the pairwise differences between German per capita real GDP (taken as a benchmark) and the per capita real GDP of the other EU countries. We interpret the absence of a unit root ("stochastic convergence") simply as the result of "balanced" growth paths in the EU. That is, that EU economies tend to grow together. However, this does not necessarily imply that regional disparities are going to disappear, given that we allow for the presence of a non-zero mean.

Our findings show some evidence in favor of stochastic convergence for the entire period only as far as the unit root tests with cross-sectional dependence are considered. However, the evidence is mixed being supported only by the CH and MP tests. The results seem more conclusive for the two sub-periods 1950-1976 and 1977-2003. As far as the first one is

[6] Gutierrez (2005) shows that MP tests have good size and power properties in finite samples for different specifications and different values of T and N, and that the BNG tests of the null hypothesis that idiosyncratic components are non-stationary have also good size and power, especially when the Dickey-Fuller-GLS version of the test is used, while the ADF test used to analyze the nonstationary properties of the common component has low power. Instead, the CH tests are largely oversized.

concerned, the majority of the tests suggest rejecting the null of a unit root; the opposite happens for the second sub-sample. This outcome seems at first sight at odd with intuition. Taken at face value, it means that economic convergence is rejected when the economic linkages are stronger. However, read in this direction, the results may be misleading. It may well be that after stronger economic links have been established, the EU economies started to move along a transitional path towards a new equilibrium. The tests used in the present paper are not designed to tackle this issue. This is why we are currently working to an extension of the research that considers also Phillips and Sul's (2003) as well as other approaches designed to take into account heterogeneity over time and cross-sectional units and the possibility that countries are experiencing transitional dynamics during the observed sample.

References

[1] Backus, D.K. and Kehoe, P.J., 1992. "International Evidence on the Historical Properties of Busyness Cycles", *American Economic Review*, 82, 864-888.

[2] Bai, J. and Ng, S., 2002. "Determining the Number of factors in Approximate Factor Models", *Econometrica*, 70, 191-221.

[3] Bai, J. and Ng S., 2004. "A PANIC Attack on Unit Root and Cointegration", *Econometrica*, 72, 1127-1177.

[4] Barro, R., 1991. "Economic Growth in a Cross-Section of Countries", *Quarterly Journal of Economics*, 106, 407-443.

[5] Barro, R. and Sala-i-Martin, X., 1992. "Convergence", *Journal of Political Economy*, 100,223-251.

[6] Bernard, A.B. and Durlauf, S. N., 1995. "Convergence in International Output", *Journal of Applied Econometrics*, 10, 97-108.

[7] Bernard, A.B. and Durlauf, S.N., 1996. "Interpreting tests of the convergence hypothesis", *Journal of Econometrics*, 71, 161-173.

[8] Bernard, A. B. and Jones, C.I., 1996. "Productivity Across Industries and Countries: Time Series Theory and Evidence", *Review of Economics and Statistics*, 78, 135-146.

[9] Caldwell, B., 1982. *Beyond Positivism: Economic Methodology in the Twentieth Century.* London: Unwin Hyman.

[10] Choi, I., 2001. "Unit Root Tests for Panel Data", *Journal of International Money and Banking*, 20, 249-272.

[11] Choi, I., 2004. "Unit Root Tests for Cross-Sectionally Correlated Panels", Forthcoming in Saikkonen, P. and I. Choi (Eds.) *Econometric Theory and Practice: Frontiers of Analysis and Applied Research. Essays in Honor of Peter C.B. Phillips.* Cambridge: Cambridge University Press.

[12] Durlauf, S.N., 2003. "The Convergence Hypothesis after 10 Years", University of Wisconsin at Madison, mimeo.

[13] Elliott, G., Rothemberg, T.J., and Stock, J.H., 1996. "Efficient Tests for an Autoregressive Unit Root", *Econometrica*, 64, 813-836.

[14] Evans, P. and Karras, G., 1996. "Convergence Revised". *Journal of Monetary Economics*, 37, 249-65.

[15] Fleissig, A. and Strauss, J., 2001. "Panel Unit-Root Tests of OECD Stochastic Convergence", *Review of International Economics*, 9, 153-162.

[16] Fisher, R.A., 1932. *Statistical Methods for Research Worker*, 4th edition. Edinburgh: Oliver & Boyd.

[17] Franses, P.H., 2002. *A Concise Introduction to Econometrics*. Cambridge: Cambridge University Press.

[18] Granger, C.W.J., 1999. *Empirical Modeling in Economics: Specification and Evaluation.* Cambridge: Cambridge University Press.

[19] Granger, C.W.J., King, M.L., and White, H., 1995. "Comments on Testing Economic Theories and the Use of Model Selection Criteria", *Journal of Econometrics*, 67, 173-187.

[20] Greenwald, A.G., Leippe, M.R., Pratkanis, A.R., and Braumgardner, M.H., 1986. "Under What Conditions Does Theory Obstruct Research Progress?", *Psychological Review*, 93, 216-229.

[21] Gutierrez, L., 2005. "Panel Unit Root Tests for Cross-Sectionally Correlated Panels: A Monte Carlo Comparison", *Oxford Bulletin of Economics and Statistics*, forthcoming.

[22] Hurlin, C. and Mignon, 2004. "Guide Pratique des Séries non Stationnaries en Panel. Partie I: Tests de Racine Unitarie", Document de Recherche LEO, 10.

[23] Im, K.S., Pesaran, M.H., and Shin, S., 1997. "Testing for unit roots in heterogeneous panels", Working Paper No. 9526, Department of Applied Economics, University of Cambridge.

[24] Im, K.S., Pesaran, M.H., and Shin, S., 2003. "Testing for unit roots in heterogeneous panels", *Journal of Econometrics*, 115, 53-74.

[25] Islam, N., 1995. "Growth empirics: a panel data approach", *Quarterly Journal of Economics*, 110, 1127-1170.

[26] Keuzenkamp, H.A., Barten, A.P., 1995. "Rejection without Falsification. On the History of Testing the Homogeneity Condition in the Theory of Consumer Demand", *Journal of Econometrics*, 67, 103-127.

[27] Levin, A., and Lin, C.F., 1992. "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties", Working Paper n.23, University of California, San Diego.

[28] Levin, A., and Lin, C.F., 1993. "Unit Root Tests in Panel Data: New Results", Working Paper n.56, University of California, San Diego.

[29] Levin, A., Lin, C.F., and Chu, C.S.J., 2002. "Unit root tests in panel data: Asymptotic and Finite-Sample Properties", *Journal of Econometrics*, 108, 1-25.

[30] Maddala, G.S., and Wu, S., 1999. "A comparative study of unit root tests with panel data and a new simple test", *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

[31] Mankiw, N. G., Romer, D., and Weil, D., 1992. "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics*, 107, 407-437.

[32] Moon, H.R., and Perron, B., 2004. "Testing for a Unit Root in Panels with Dynamic Factors", *Journal of econometrics*, 122, 81-126.

[33] Newey, W. and West, K., 1994. "Autocovariance Lag Selection in Covariance Matrix Estimation", *Review of Economic Studies*, Vol. 61, 613-653.

[34] Pesaran, H., 2005. "A simple panel unit root test in the presence of Cross Section Dependence", Cambridge University and USC, mimeo.

[35] Phillips, P.C.B., and Sul, D., 2003. "The Elusive Empirical Shadow of Growth Convergence", Yale University, mimeo.

[36] Solow, R., 1956. "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, 70, 65-94.

[37] Stock, J. H. and Watson, M.W., 1988. "Testing for Common Trends", *Journal of the American Statistical Association*, 83, 1097-1107.

[38] Summers, L.H., 1991. "The Scientific Illusion in Empirical Macroeconomics", in S.Hylleberg and M.Paldam (Eds.) *New Approaches to Empirical Macroeconomics.* Oxford: Basil Blackwell.

[39] Swan, T.W., 1956. "Economic Growth and Capital Accumulation", *Economic Record*, 32, 334-361.