Convergence test in the presence of structural changes: an empirical procedure based on panel data with cross-sectional dependence *

Abdou-Aziz NIANG[♭] Marie-Claude PICHERY[♮] Marcellin EDJO[†]

April 2010

This work is dedicated to the memory of Marcellin EDJO, an economist at the BCEAO, who, before disappearing prematurely December 14, 2008 in Dakar, had contributed to a first version of this paper. He also conducted other works addressing economic convergence in the framework of properties of non-stationary series.

abstract

This paper presents an essay on empirical testing procedure for economic convergence. Referring to the unit root test proposed by Moon and Perron (2004), we proposed a modified Evans (1996) testing procedure of the convergence hypothesis. The advantage of this modified procedure is that it makes possible to take into account cross-sectional dependences that affect GDP per capita. It also allows to take into account structural instabilities in these aggregates. The application of the procedure on OECD member countries and CFA zone member countries leads to accept the hypothesis of economic convergence for these two groups of countries, and it shows that the convergence rate is significantly lower in the OECD sample. However, the results of the tests applied to the Global sample composed by all countries in these two samples conclude a rejection of the convergence hypothesis.

Keywords : *$\beta-$convergence; Unit root; Panel data; Factor model; Cross-sectional dependence; Structural change*

JEL Classification: *C22; C23; O40; R11*

[*♭]PhD candidate in economics, University of Bourgogne (France).Email:Abdou-Aziz.Niang@u-bourgogne.fr. [♮]Professor of economics, University of Bourgogne (France).Email:marie-claude.pichery@u-bourgogne.fr. [†]Economist at BCEAO

1 Introduction

Since the works of Baumol (1986), Barro and Sala-i-Martin (1991, 1995), a large number of papers have focused on the analysis of convergence using generally two conventional approaches: $\beta-$convergence and σ-convergence. These two forms of convergence have many applications in time series properties. Indeed, the development of econometric analysis techniques and the availability of databases (Summers and Heston, 1991) covering large periods provide the opportunity to go beyond the cross-sectional analysis and exploit properties of nonstationary time series (Bernard and Durlauf 1995; Edjo 2003) to better inform the debate on economic convergence.

Convergence tests are also expanded in the framework of panel data analysis. The first tests in panel based primarily on the methodology used in cross-sectional analysis. One can cite the works of Islam (1995) and Berthelemy et al (1997). Then, like the procedure used in the individual time series, panel unit root tests are used to study economic convergence. This procedure based on panel unit root test is implemented by Quah (1992), Evans (1996), Evans and Karras (1996), Bernard and Jones (1996), Gaulier et al. (1999) among others. Indeed, the combination of the cross-section and time dimensions allows for more powerful tests. Now, there are essentially two generations of unit root tests. And most of the methodologies of the analysis of economic convergence using the properties of non-stationary series refer to the first generation that puts forward the hypothesis of independence between individuals (Levin and Lin 1993; Im, Pesaran and Shin 1997; Harris and Tzavalis 1999; Maddala and Wu 1999; Hadri 2000; Choi 2001). However, as pointed Hurlin and Mignon (2005), in applications of macro-economic convergence tests, this assumption of cross-section independence is particularly troublesome. The second generation of unit root tests (Choi 2002; Phillips and Sul 2003; Pesaran 2003; Bai and Ng 2004; Moon and Perron 2004) generally based on common factors models allows taking into account more general forms of cross-sectional dependences.

In this paper, the empirical procedure we propose is based precisely on unit root tests of the second generation and allows to take into account explicitly the dependences in the cross-sectional dimension. We focus on the fact that the cross-country correlation that may exist in the convergence equation is not only due to simple correlation of residuals, but also to the presence of one or more common factors that jointly affect the real GDP per capita of the countries. Therefore, the study of the convergence in panel based on the standard ADF model as advocated by Evans and Karras (1996) is no more suitable because it leads to tests with very low power (Strauss and Yigit, 2003).

Another issue addressed in this procedure is the existence of structural changes in per capita GDP. Works addressing structural changes in panel data with cross-sectional dependence are generally very rare. Examples include Bai and Carrion-i-Silvestre (2009) and Carrion-i-Silvestre and German-Soto (2009). As pointed out by Carrion-i-Silvestre et al. (2005), ignoring these shocks in the econometrics of panel data can lead to biases that lead to wrong conclusions. The financial and economic crises, economic reforms ..., are factors that may cause such shocks.

In the next section, we present the approach generally used to test convergence in nonstationary panel data. Then, we apply the procedure we propose which is inspired by this traditional approach. In section 4, we conduct monte carlo simulation to explore the impact of the proposed procedure on performance test. Section 5 presents the application conducted using a sample of OECD member countries and a sample of CFA zone member countries.

2 Convergence tests in panel data econometrics

The convergence hypothesis tests in panel data are generally based on the standard approach in cross-section that is to test whether economies with low initial income relative to their long-term position or steady state will grow faster than economies with high initial income. This involves applying ordinary least squares (OLS) to the equation

$$\frac{1}{T}\ln(y_{i,T}/y_{i,0}) = \alpha + \beta\ln(y_{i,0}) + \varphi\Xi_i + \xi_i \qquad \xi_i \backsim i.i.d(0,\sigma_\xi^2) \tag{1}$$

where y_i is real GDP per capita of country i, Ξ_i is a vector of controlled variables so as to maintain constant steady state of each economy i and ξ_i is the error term. The index T refers to the length of the time interval. α, β, φ are unknown parameters which have to be estimated. The convergence speed $\theta = -\ln(1+\beta T)/T$ is the speed required for each economy to reach its steady state. The null hypothesis tested is the lack of convergence against the alternative that some countries converge to a certain level of production initially different. If the estimated coefficient β is negative and significant, one can accept the hypothesis of convergence, which means that once the variables that influence growth are controlled, low-income economies tend to grow faster towards their own steady state. It is possible to deduce the time necessary for countries to fill half the gap separating them from their steady state, from the coefficient β. This half-life is given by the expression $\tau = -\ln(2)/\ln(1+\beta)$

However, OLS estimation of (1) is useful for inference only under certain conditions. Indeed, Evans and Karras (1996) explain that the estimators $\hat{\beta}$ and $\hat{\varphi}$ obtained by applying ordinary least squares to (1) are valid only if ξ_i and $y_{i,0}$ are uncorrelated and the constant term is generated as

$$\delta_i = \psi^{'} X_i \tag{2}$$

with $\psi \equiv (\lambda-1)\varphi/\beta$. In panel data, Evans and Karras's (1996) procedure based on unit root tests is a basic procedure for many studies on economic convergence tests. Considering a group of N countries, these authors show that the countries converge if deviations of the log GDP per capita from the international average are stationary for each country. Let y_{it} be the log GDP per capita of country i at the period t with $i = 1, ...N$; $t = 1, ...T$ and $\bar{y}_t$ the international average[1] of y_{it}. This is to test whether the data generating process $(y_{it} - \bar{y}_t)$ is stationary for all i

$$\lim_{h\to\infty}(y_{i,t+h} - \bar{y}_{t+h}) = \mu_i \tag{3}$$

Convergence occurs if for each i deviations of per capita GDP from the international average tend to a constant when $t \to \infty$. Specifically, the convergence hypothesis is accepted only if $y_{it} - \bar{y}_t$ are

[1] $\bar{y}_t = \sum_{i=1}^{N} y_{it}/N$

not stationary while the y_{it} are integrated of order 1. In such case, we have stochastic convergence. However, as stressed by Carrion-i-Silvestre and German-Soto (2009), stochastic convergence is a necessary but not sufficient condition to satisfy the definition of β-convergence. For this, we consider the data generating process proposed by Evans (1996)

$$y_{it} - \bar{y}_t = \delta_i + \lambda(y_{i,t-1} - \bar{y}_{t-1}) + u_{it} \tag{4}$$

where $\lambda \equiv (1 + \beta T)^{(1/T)}$ is inferior to 1 if the N economies converge and in this case $\beta \prec 0$. However, there is divergence if $\lambda = 1$ this also implies that $\beta = 0$. δ_i is a constant specific to each economy and the error term is serially uncorrelated. Moreover, Evans and Karras (1996) show that in the case where the error terms are correlated in the cross-sectional dimension, this specification implies serious problems of statistical inference. Or, international trade in goods and assets makes innovations probably correlated. In addition, given the specificity of countries in terms of technology, the parameter λ should be specific to each economy. Therefore, the ADF specification in panel with a heterogeneous autoregressive root is generally used as alternative

$$\Delta(y_{it} - \bar{y}_t) = \delta_i + \rho_i(y_{i,t-1} - \bar{y}_{t-1}) + \sum_{s=1}^{p} \gamma_{i,s}\Delta(y_{i,t-p} - \bar{y}_{t-p}) + u_{it} \tag{5}$$

The parameter ρ_i is negative if the economies converge and is equal to zero if they diverge. The roots of $\sum_s \gamma_{i,s} L^s$ are outside the unit circle. In the application we propose below, we use a general specification of equation (4) which allows better control of cross-sectional dependences of the term u_{it} and the specificity of the coefficient λ. It also takes into account possible structural changes affecting the parameter δ_i.

3 An alternative procedure

This section exposes the proposed procedure for testing β-convergence hypothesis which is equivalent to verify if $0 \preceq \lambda \prec 1$. To do so, the procedure is decomposed into two steps. In the first step we use the Moon and Perron (2004) procedure to produce a consistent modified pooled estimator of λ. Then we test stochastic convergence, a primary condition of β-convergence. This is to test nonstationarity of per capita GDP cross economies differences ($H_0 : \lambda = 1$). If stochastic convergence is verified, the second step consists of testing whether $\lambda = 0$ and determining the implied value of β.

3.1 The econometric specification

As mentioned previously, specification (4) is useful only under certain conditions and if they are not verified estimating consistently parameters of the model will be very challenging. These conditions are relative to the error term u_{it} and can be summarized in two general points related by Evans (1996). (*i*) u_{it} is a serially uncorrelated error term with a zero mean and finite and constant variance. (*ii*) Also, u_{it} is contemporaneously uncorrelated across countries. To deal with cross-sectional correlation of u_{it} we use the data generating process in Moon and Perron (2004) to define a general form of equation (4)

$$(y_{it} - \bar{y}_t) = \delta_i + \lambda_i(y_{i,t-1} - \bar{y}_{t-1}) + u_{it} \quad (6)$$

In this model, the correlation among the cross-sectional units of u_{it} are captured by using a factor model

$$u_{it} = \pi_i^{'} F_t + e_{it} \quad (7)$$

where F_t is a $(T \times r)$ matrix representing the common factors, π_i is a $(r \times 1)$ vector of factors loadings and the $(T \times 1)$ vector represents the idiosyncratic term.

The procedure is to first deal with these cross-sectional dependences by removing common factors. Then the null hypothesis of divergence is tested on the variable $y_{it} - \bar{y}_t$ previously de-factored. This is equivalent to test the null hypothesis of unit root $H_0 : \lambda_i = 1 \ \forall i$ against the alternative hypothesis of stationarity $H_1 : \lambda_i \prec 1$ for some individuals of the panel.

To take into account the structural changes that may affect the series, we propose a general form of equation (4) which admits the presence of one break in the constant

$$y_{it} - \bar{y}_t = \delta_i + \theta_i DU_{i,t} + \lambda_i(y_{i,t-1} - \bar{y}_{t-1}) + u_{it} \quad (8)$$

where $DU_{i,t} = 1$ for $t \succ T_i$ and 0 elsewhere. T_i denotes the break in the intercept for the $i - th$ individual. Let $y_{0,it} - \bar{y}_{0,t} = \lambda_i(y_{i,t-1} - \bar{y}_{t-1}) + u_{it}$ with $y_{0,i,0} - \bar{y}_{0,0} = 0$, the first-differenced form of equation (8) is

$$\Delta(y_{it} - \bar{y}_t) = \theta_i I(T_i)_t + \Delta(y_{0,it} - \bar{y}_{0,t}) \quad (9)$$

where $I(T_i)$ are impulses such that $I(T_i)_t = 1$ for $t = T_i + 1$ and 0 elsewhere. Following[2] Bai and Carrion-i-Silvestre (2009), we ignore these impulses since they take into account a few unusual events and their effect is asymptotically negligible.

Thus, replacing $\Delta(y_{it} - \bar{y}_t)$ by Δy_{it}^c and $\Delta(y_{0,it} - \bar{y}_{0,t})$ by $\Delta y_{0,it}^c$ in equation (9), we can define $\hat{y}_{0,it}^c = \sum_{s=2}^{t} \Delta\hat{y}_{is}^c = (y_{it} - \bar{y}_t) - (y_{i1} - \bar{y}_1)$ the cumulative sum of $y_{0,it}^c$. That is

$$\hat{y}_{0,it}^c = \lambda_i \hat{y}_{i,t-1}^c + (u_{it} - u_{i1}) \quad (10)$$

Finally for $t = 2, ..., T$, we define

$$\hat{y}_{0,it}^c = \lambda_i \hat{y}_{i,t-1}^c + \hat{u}_{it} \quad (11)$$

where $\hat{u}_{it} = \pi_i^{'} f_t + \varepsilon_{it}$ with $f_t = (F_t - F_1)$ and $\varepsilon_{it} = (e_{it} - e_{i1})$. Indeed, the $\hat{y}_{0,it}^c$ series preserve the same nonstationarity property as the original series $y_{0,it}^c$ (Bai and Ng, 2004) and has the advantage to not suffering from the presence of structural change. Thus, we face the simple case of a test without break. In section 4 we proceed to monte carlo simulations to verify whether this procedure can affect the performance of the test.

Now, considering all individuals of the panel, the matrix form of equation (11) under the null hypothesis is

$$\hat{Z}_0 = \hat{Z}_{0,-1} + f\pi^{'} + \varepsilon \quad (12)$$

[2]These authors adopt this procedure in their modified Sargan-Bhargava (MSB) tests which takes into account structural changes and common factors.

where $\hat{Z}_0$, $\hat{Z}_{0,-1}$ and ε are $(T-1)\times N$ matrices. $\hat{Z}_0$ is the matrix of individual observations of $\hat{y}^c_{0,it}$ and $\hat{Z}_{0,-1}$ the lagged observations matrix. π is the $N\times r$ matrix of factor loadings.

To remove common factors, Moon and Perron (2004) propose an orthogonalization procedure that is similar to Phillips and Sul's (2003) method. Considering the equation (12), they use the projection matrix that allows de-factoring the data by right-multiplying by the matrix projection. With $\tilde{Z}_0 = \hat{Z}_0\hat{Q}_\pi$ and $\tilde{Z}_{0,-1} = \hat{Z}_{0,-1}\hat{Q}_\pi$, equation (12) becomes

$$\tilde{Z}_0 = \tilde{Z}_{0,-1} + \tilde{\varepsilon} \tag{13}$$

Then, for each individual i with $i = 1, ..., N$, the de-factored form of model (11) is

$$\tilde{y}^c_{0,it} = \lambda_i \tilde{y}^c_{i,t-1} + \tilde{\varepsilon}_{it} \tag{14}$$

where $\tilde{\varepsilon}_{it}$ is uncorrelated across country accordance to condition (ii). Note that the projection matrix is obtained by principal component analysis developed by Bai and Ng (2002). This allows to estimate the number r of common factors and the factor loadings matrix π.

3.2 Testing stochastic convergence

The implementation of the testing procedure of Moon and Perron (2004) requires to estimate the projection matrix used to eliminate dependences and to define a consistent estimator of λ. In the next sub-sections we present the method for estimating the projection matrix, the estimation of λ and the tests statistics of the null hypothesis $\lambda = 1$.

3.2.1 Estimation of the projection matrix

The matrix of estimated factors $\tilde{f}$ is equal to $\sqrt{T-1}$ times the eigenvectors corresponding to the r largest eigenvalues of the $(T-1)\times(T-1)$ matrix $\hat{u}\hat{u}^{'}$. Considering the normalization $\pi^{'}\pi/N = I_r$ and $f^{'}f/(T-1) = I_r$, the matrix of factor loadings can be obtained by ordinary least squares $\tilde{\pi}^{'} = (\tilde{f}^{'}\tilde{f})^{-1}\tilde{f}^{'}\hat{u} = \tilde{f}^{'}\hat{u}/(T-1)$. Then, we can use a re-scaled estimator defined as

$$\hat{\pi} = \tilde{\pi}\left(\frac{1}{N}\tilde{\pi}^{'}\tilde{\pi}\right)^{1/2} \tag{15}$$

Furthermore, to estimate r we use the IC_1 and BIC_3 information criteria of Bai and Ng (2002). The BIC_3 criterion is a modification of the usual BIC which perform better in small samples ($N \leq 20$). Let [3] $V(r,f)$ the sum of squared residuals (divided by $N(T-1)$) of the regression of $\hat{u}_{it}$ on the r factors for each i. If $N \leq 20$, we can use[4] $V(r,f) + r_{max}g_{BIC}(N,T)$, where $g_{BIC}(N,T)$ is the penalty function. Bai and Ng (2002) show that in this case, r can be estimated consistently with $g_{BIC}(N,T) = \frac{(N+T-1-r)ln(N(T-1))}{N(T-1)}$ by minimizing[5]

$$BIC_3(r) = V(r,\tilde{f}) + r\hat{\sigma}^2_\epsilon(r_{max})\left(\frac{(N+T-1-r)ln(N(T-1))}{N(T-1)}\right) \tag{16}$$

[3] $V(r,f)$ is the variance of the idiosyncratic component estimated with the maximum number of factors
[4] r_{max} is the maximum number of factors
[5] $\hat{\sigma}^2_\epsilon$ is the variance of the estimated idiosyncratic components

For IC_1 the penalty function is $g_{IC}(N,T) = \frac{N+T-1}{N(T-1)} ln\left(\frac{N(T-1)}{N+T-1}\right)$ and the problem consists of minimizing

$$IC_1(r) = ln\left(V(r,\tilde{f})\right) + r\hat{\sigma}_\epsilon^2(r_{max})\frac{N+T-1}{N(T-1)} ln\left(\frac{N(T-1)}{N+T-1}\right) \quad (17)$$

An estimate of the projection matrix Q_π which allows to get de-factored data is given by

$$\hat{Q}_\pi = I_N - \hat{\pi}\left(\hat{\pi}'\hat{\pi}\right)^{-1}\hat{\pi}' \quad (18)$$

In the next subsection, this matrix is used to define a pooled estimator of λ that we denote $\hat{\lambda}^*$ and then we construct the Moon and Perron (2004) test statistics.

3.3 Estimation of λ and construction of test statistics

The test statistics can be constructed using the modified pooled OLS estimator of the autoregressive root. Note that this estimator is adjusted to take account condition (i). Thus the possible serial correlation of the idiosyncratic residual $\tilde{\varepsilon}_{it}$ is controlled. Let ϕ_e be the sum of positive autocovariances of the idiosyncratic component and $\hat{\lambda}^*$ the modified pooled OLS estimator of λ which is defined as

$$\hat{\lambda}^* = \frac{trace\left(\hat{Z}_{0,-1}\hat{Q}_\pi\hat{Z}_0'\right) - N(T-1)\hat{\phi}_e}{trace\left(\hat{Z}_{0,-1}\hat{Q}_\pi\hat{Z}_{0,-1}'\right)} \quad (19)$$

Two test statistics of the null hypothesis $\lambda = 1$ are constructed by Moon and Perron (2004) from the pooled estimator. They are noted t_a and t_b and both follow a standard normal law

$$t_a^* = \frac{(T-1)\sqrt{N}(\hat{\lambda}^* - 1)}{\sqrt{2\hat{\nu}_e^4/\hat{\omega}_e^4}} \longrightarrow N(0,1) \quad (20)$$

$$t_b^* = (T-1)\sqrt{N}(\hat{\lambda}^* - 1)\sqrt{\frac{1}{N(T-1)^2} trace\left(\hat{Z}_{0,-1}\hat{Q}_\pi\hat{Z}_{0,-1}'\right)\frac{\hat{\omega}_e^2}{\hat{\nu}_e^4}} \longrightarrow N(0,1) \quad (21)$$

ω_e^2 and ν_e^4 respectively correspond to the means on N of the individual long-term variances $\nu_{e,i}^4$ and of squared individual long-term variances $\phi_{e,i}^4$ of the idiosyncratic component e_{it}. Let $\hat{\Gamma}_i(j)$ be the residual empirical autocovariance

$$\hat{\Gamma}_i(j) = \frac{1}{T}\sum_{t=1}^{T-j}\hat{e}_{it}\hat{e}_{i,t+j}$$

From $\hat{\Gamma}_i(j)$, it is possible to construct an estimator of the individual long-term variances[6]

$$\hat{\omega}_{e,i}^2 = \frac{1}{N}\sum_{j=-T+1}^{T-1}\omega(q_i,j)\hat{\Gamma}_i(j)$$

$$\hat{\phi}_{e,i} = \sum_{j=1}^{T-1}\omega(q_i,j)\hat{\Gamma}_i(j)$$

[6] $q_i = 1.3221\left[\frac{4\hat{\psi}_{i,1}^2 T_i}{(1-\hat{\psi}_{i,1})^4}\right]^{1/5}$ with $\hat{\psi}_{i,1}$ the first-order autocorrelation estimate of $\hat{e}_{it}$; $\omega(q_i,j) = \frac{25}{12\pi^2\kappa^2}\left[\frac{sin(6\pi\kappa/5)}{6\pi\kappa/5} - cos\left(\frac{6\pi\kappa}{5}\right)\right]$ with $\kappa = \frac{j}{q_i}$

This allows to define the estimates of the means of the individual long-term variances as follows

$$\hat{\omega}_e^2 = \frac{1}{N}\sum_{i=1}^{N}\hat{\omega}_{e,i}^2 \qquad \hat{\phi}_e = \frac{1}{N}\sum_{i=1}^{N}\hat{\phi}_{e,i} \qquad \hat{\nu}_e^4 = \frac{1}{N}\sum_{i=1}^{N}(\hat{\omega}_{e,i}^2)^2$$

The test statistics are obtained by substituting the estimated values of these variances in the expressions of t_a^* and t_b^*. If the realization of the statistic t^* is lower than the normal critical level, we accept the hypothesis of stochastic convergence for all N countries.

3.3.1 Analyzing β-convergence

In this subsection, the aim is to estimate the implied value of β given by $\hat{\beta} = \left((\hat{\lambda}^*)^T - 1\right)/T$ in order to analyze β-convergence. For this purpose we use $\hat{\lambda}^*$, the consistent estimator of λ. However, we previously need to test the nullity of λ, and for this we proceed to three steps.

Step 1: We use the recumulated first differenced form of equation (4) and obtain

$$\hat{y}_{0,it}^c = \lambda\hat{y}_{i,t-1}^c + \hat{u}_{it} \tag{22}$$

where the variables are defined as in equation (11). Then, for each i, we normalize the $\hat{y}_{0,it}^c$ series by the OLS regression standard error $\hat{\sigma}_{\hat{u}i}$ to control for heterogeneity across countries. The normalized series is $\hat{S}_{0,it} = \hat{y}_{0,it}^c/\hat{\sigma}_{\hat{u}i}$

Step 2: Considering the normalized model, we have

$$\hat{S}_{0,it}^c = \lambda_s\hat{S}_{i,t-1}^c + \hat{\nu}_{it} \tag{23}$$

where $\hat{\nu}_{it} = \hat{u}_{it}/\hat{\sigma}_{\hat{u}i}$. Applying the procedure presented in subsection 3.1 we obtain the de-factored form of the normalized model

$$\tilde{S}_{0,it} = \lambda_s\tilde{S}_{i,t-1}^c + \tilde{\nu}_{it} \tag{24}$$

Step 3: Let $\hat{S}_0$ the matrix of observations $\hat{S}_{0,it}^c$ and $\hat{S}_{0,-1}$ the matrix of lagged observations. Using λ_s^* the modified pooled estimator of the normalized equation obtained by replacing $\hat{Z}$ by $\hat{S}$ in (19), we calculate the t-statistic

$$t_{\lambda_s}^* = \frac{\lambda_s^*}{\sigma_\lambda^*} \tag{25}$$

where

$$\sigma_\lambda^* = \hat{\sigma}_{\tilde{\varepsilon}^*}\left(\sum_{i=1}^{N}\sum_{t=2}^{T}(\tilde{S}_{i,t-1}^c)^2\right)^{-1/2}$$

and

$$\hat{\sigma}_{\tilde{\varepsilon}^*} = \sqrt{trace\left((\tilde{S}_0 - \lambda_s^*\tilde{S}_{0,-1})(\tilde{S}_0 - \lambda_s^*\tilde{S}_{0,-1})'\right)/N(T-1)}$$

$\tilde{S}_0$ and $\tilde{S}_{0,-1}$ are respectively the matrices of $\tilde{S}_{0,i,t}^c$ and $\tilde{S}_{0,i,t-1}^c$. We compare this statistic with the appropriate critical value. However, since we don't now the limiting distribution of $t_{\lambda_s}^*$, it is approximated by simulations. We use ordinary least squares to estimate the parameters of the null model $y_{it} - \bar{y}_t = \delta_i + \theta_i DU_{i,t} + u_{it}$ where $u_{it} = \pi_i' F_t + e_{it}$. For this, we fit $y_{it} - \bar{y}_t$ to country fixed effect and break (if any). Then the residual from this OLS regression are used to effectuate the principal component analysis described in sub-section 3.2.1 in order to get parameters of common and idiosyncratic components of u_{it}. We generate 10,000 data sets for the fitted null model. For each of the generated data sets, we use the alternative model $y_{it} - \bar{y}_t = \delta_i + \theta_i DU_{i,t} + \lambda(y_{i,t-1} - \bar{y}_{t-1}) + u_{it}$ and proceed to the three steps above to estimate the modified OLS pooled estimator for the normalized model and to determine the test statistic $t^*_{\lambda_s}$. With a sample of 10,000 values of $t^*_{\lambda_s}$ we obtain critical values which correspond to quantiles 5% and 10%. Then, $t^*_{\lambda_s}$ is compared to these critical values.

3.4 Monte Carlo Simulations : Exploring the test performance of the difference-recumulation approach

This section presents the results of Monte Carlo simulations which investigate whether the difference-recumulation procedure used in this paper affects the Moon and Perron (2004) test performance of the null hypothesis $\lambda = 1$ in the case of the presence of single break. That is precisely to show that the transformations effectuated for testing unit root will not impact negatively on the size and power of the test. We will conduct two experiments using MATLAB 6.5 and for simplicity of notation we replace $y_{it} - \bar{y}_t$ by x_{it}.

The experiment 1 reproduces the same conditions in Moon and Perron (2004), case "single factor, fixed effects not estimated".

Experiment 1:

$$
\begin{aligned}
x_{it} &= \delta_i + x_{0,i,t} \\
x_{0,i,t} &= \lambda_i x_{0,i,t-1} + u_{it} \\
x_{0,i,0} &= 0
\end{aligned}
$$

Let $\mu_i = T_i/T$ be the break fraction for every i. The break points are randomly positioned with break fractions following $\mu_i \sim [0.2, 0.8]$ and we have

Experiment 2:

$$
\begin{aligned}
x_{it} &= \delta_i + \theta_i DU_{i,t} + x_{0,i,t} \\
x_{0,i,t} &= \lambda_i x_{0,i,t-1} + u_{it} \\
x_{0,i,0} &= 0
\end{aligned}
$$

For this second experiment, using the same transformations presented in sub-section 3.1, we define $\hat{x}_{0,it} = \sum_{s=2}^{t} \triangle x_{0,it} = x_{it} - x_{i1}$ the estimate of $x_{0,it}$. In both experiments, the error term has a factor structure and we adopt the data generating process in Bai and Ng (2002)

$$u_{it} = \sum_{j=1}^{r} \pi_{ij} F_{tj} + \sqrt{r} e_{it}$$

However, we only consider the case of a single common factor in which common shocks are *iid* standard normal

$$(F_{ij}, \pi_{ij}, e_{it}) \sim iidN(0, I_3)$$

Note also that in both experiments $\delta_i \sim N(0,1)$. To study the size, we set $\lambda_i = 1$ for all i. For power, we have considered values of λ_i that are not far from the null hypothesis of unit root. Thus, under the alternative, the parameter λ is specific to each individual and has an average value equal to 0.99. The number of common factors is estimated with the procedure of Bai and Ng (2002) and using the BIC_3 and IC_1 criteria. The maximum number of factors is equal to 8. Simulations are conducted using 1000 replications with $N = \{10, 20\}$ and $T = \{100, 300\}$ and we consider the 5% significance level.

[TABLE 1 HERE]

Table 1 presents the results for power and size in each empirical experiment described above. For these two data generating processes, the properties of size and power of t_a^* and t_b^* tests are studied by considering the percentage of replications in which the unit root hypothesis is rejected. This table gives also the average number of factors estimated using the selection criteria and the average true number of factors which is equal to 1. As expected in our analysis, the results show that this procedure does not affect the finite-sample properties of Moon and Perron's (2004) test. The properties of size and power of the two experiments are very similar especially for the t_b^* test which provides the best performance statistics. In addition, these results show that the proposed transformations in the procedure have no impact in the results of estimating the number of common factors. The average number of common factors remains the same for both experiments and irrespective of the couple (N, T) considered. Finally, note that the BIC_3 criterion gives a precise estimate of the number of common factors when $N = 20$.

4 Application

4.1 Data

The data are from the World Development Indicators (WDI) of World Bank Group. These are annual real per capita GDP covering the period 1975-2008. To compare results for developed and poor countries we consider two samples. The first sample *OECD* include 20 OECD members countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, United Kingdom, Canada, United States, Japan and New Zealand. The second sample called *CFA* is composed by 8 CFA Zone members countries identified as co-moving countries in Diagne and Niang (2008). These countries are given by the following list: Benin, Burkina Faso, Cameroon, Congo Rep., Cote d'Ivoire, Niger, Senegal and Togo. A global sample called *GLOBAL* and composed by these two groups of countries is also considered. Thus, this last sample consists of 28 countries including poor and rich countries.

4.2 Results

Table 1 presents the results of convergence tests for these three samples. Note that we have previously tested the existence of breaks in our data using the test procedure of Bai and Perron (1998). We also tested the null hypothesis of cross-sectional independence using the Pesaran (2004) test in panel data which is robust to breaks. The results of these tests are given in Appendix and show that we face problems of structural changes and cross-section dependences in our samples. By implementing the proposed empirical procedure to take into account these problems, the results show that the *OECD* sample have converged over the period 1975-2008 and admit a number of common factors equal to 6. The p-values associated to tests statistics t_a^* and t_b^* are respectively lower than the 10% level, indicating rejection of the null hypothesis of divergence for these countries. Thus, the parameter $\hat{\lambda}^*$ is lower than unit with a value $\hat{\lambda}^* = 0.9815$. Results of the tests based on $t_{\lambda_s}^*$ show that $\lambda \succ 0$.[7] Consequently, the implied value $\hat{\beta} = -0.0140$. These results allow defining the speed of convergence and the half-life for the countries of this sample. The convergence rate is 1.88% and the corresponding half-life is 49 years.

[TABLE 2 HERE]

The results for the countries of CFA Zone show that both t_a^* and t_b^* tests accept the hypothesis of stochastic convergence for these countries at the 1% level. Also, CFA Zone member countries have converged during the 1975-2008 period with a speed of convergence higher than that of countries in OECD. Indeed, with a parameter $\hat{\beta} = -0.0295$ the speed of convergence is 11%, implying a half-life equal to 23 years. For this sample, the number of common factors selected by BIC_3 criterion is equal to 6. Given the small size of this sample, we set $r_{max} = 6$ unlike the other samples where the maximum number of common factors allowed is 8. However, results for this sample (CFA) must be analyzed with caution. Indeed, as shown by simulation results, the reduced size of the cross-sectional dimension tends to make it difficult the estimation of $\hat{r}$.

Regarding the full sample (*GLOBAL*), the null hypothesis of divergence was accepted. The probabilities associated to t_a^* and t_b^* are higher than the standard levels of 5% and 10%. Moreover, for this case the number of factors is estimated using the IC_1 criterion which suggests the presence of 7 factors.

5 Conclusion

This study presented a testing procedure of economic convergence in panel. Based on the approach proposed by Evans (1996), we implemented an application of this convergence test procedure drawing on recent work by Moon and Perron (2004). This procedure allows to focus on cross-sectional dependences and structural changes which, if ignored, can lead to bias. This approach goes beyond the standard approach of considering these phenomena as nuisance parameters.

The application of the procedure on *OECD* and *CFA* samples lead to accept the hypothesis of economic convergence for each of these two groups of countries, with convergence rates respectively

[7] Following Evans and Karras (1996) we take for granted that $\lambda \geq 1$

equal to 1.88% and 11%. However, the results of the test applied to the full sample (*GLOBAL*) led to the rejection of the convergence hypothesis

References

[1] Bai J, Carrion-i-Silvestre J L (2009) Structural Changes, Common Stochastic Trends, and Unit Roots in Panel Data. *Review of Economic Studies, Blackwell Publishing* 76:471-501.

[2] Bai J, Ng S (2004) A PANIC Attack on Unit Roots and Cointegration, *Econometrica* 72:1127-1178.

[3] Bai J, Ng S (2002) Determining the Number of Factors in Approximate Factor Models, *Econometrica* 70:191-221.

[4] Barro R J, Sala-i-Martin X (1991) Convergence across States and Regions, *Brookings Papers on Economic Activity* 1:107-182.

[5] Baumol W J (1986) Productivity Growth, Convergence and Welfare : What the Long Run Data Show, *American Economic Review* 76:1072-1085.

[6] Bernard A, Jones C (1996) Productivity Across Industries and Countries: Times Series Theory and Evidence, *The Review of Economics and Statistics* 78:135-146.

[7] Bernard A B, Durlauf S N (1995). Convergence of International Output, *Journal of Applied Econometrics* 10:97-108.

[8] Berthélemy J C, Dessus S, Varoudakis A (1997) Capital Humain et Croissance : le Rôle du Régime Commercial, *Revue Economique* 48:419-427.

[9] Carrion-i-Silvestre J L, del Barrio-Castro T, Lopez-Bazo E (2005) Breaking the panels: An application to the GDP per capita, *Econometrics Journal* 8:159-175.

[10] Diagne A, Niang A A (2008) Co-mouvements économiques dans les pays de la Zone CFA : Une analyse par le modèle factoriel dynamique généralisé, Laboratoire d'Economie et de Gestion, *Université de Bourgogne.*

[11] Edjo, E. M. (2003), Analyse économétrique de la croissance, de la convergence et des changements structurels dans les pays de la zone CFA : une approche par les séries temporelles, thèse Université de Bourgogne

[12] Evans P (1996) Using cross-country variances to evaluate growth theories, *Journal of Economic Dynamics and Control* 20:1027-1049.

[13] Evans P, Karras G (1996). Convergence Revisited, *Journal of Monetary Economics* 37:249-265.

[14] Gaulier G, Hurlin C, Jean-Pierre P (1999). Testing Convergence: A Panel Data Approach, *Annales d'Economie et de statistique* 17:55-56.

[15] Hadri K (2000) Testing for Unit Roots in Heterogeneous Panel Data, *Econometrics Journal* 3:148-161.

[16] Harris R D F, Tzavalis E (1999) Inference for Unit Roots in Dynamic Panels where the Time Dimension is Fixed, *Journal of Econometrics* 91:201-226.

[17] Hurlin C, Mignon V (2006) Une Synthèse des Tests de Racine Unitaire sur Données de Panel, *Economie et Prévision* 169:253-294.

[18] Im K S, Pesaran M H, Shin Y (2003) Testing for Unit Roots in Heterogenous Panels, *Journal of Econometrics* 115:53-74.

[19] Islam N (1995) Growth Empirics: A Panel Data Approach, *Quarterly Journal of Economics* 110:1127-1170.

[20] Lluis Carrion-i-Silvestre J, German-Soto V (2009). Panel data stochastic convergence analysis of the Mexican regions, *Empirical Economics* 37:303-327.

[21] Maddala G S, 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.

[22] Moon H R, Perron B (2004) Testing for a Unit Root in Panels with Dynamic Factors, *Journal of Econometrics* 122:81-126.

[23] Pesaran H M (2007) A Simple Panel Unit Root Test in the Presence of Cross Section Dependence, *Journal of Applied Econometrics* 22:265-312

[24] Phillips P C B, Sul D (2003) Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence, *Econometrics Journal* 6:217-259.

[25] Solow R M (1956) A Contribution to the Theory of Economic Growth, *Quarterly Journal of Economics* 70:65-94.

[26] Strauss J, Yigit T (2003) Shortfalls of Panel Unit Root Testing, *Economics Letters* 81:309-313.

[27] Summers R, Heston A (1991) The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988, *The Quarterly Journal of Economics*, 106:327-368.

[28] Swan T W (1956) Economic Growth and Capital Accumulation, *Economic Record* 32:334-361.

Table 1: Monte Carlo simulations results

(N,T)			Experiment 1			Experiment 2	
		t_a^*	t_b^*	mean r	t_a^*	t_b^*	mean r
Size							
$(10,100)$	True r	12.1	7.7	1.00	13.1	7.7	1.00
	BIC_3	29.2	23.1	5.86	30.7	23.2	5.85
	IC_1	40.1	32.8	8.00	41.5	33.3	8.00
$(20,100)$	True r	12.1	7.3	1.00	11.5	6.7	1.00
	BIC_3	12.1	7.3	1.00	11.5	6.7	1.00
	IC_1	12.1	7.3	1.00	11.5	6.7	1.00
$(10,300)$	True r	12.8	8.0	1.00	12.6	7.7	1.00
	BIC_3	15.9	10.8	1.81	16.4	10.0	1.81
	IC_1	39.5	33.3	8.00	40.6	33.4	8.00
$(20,300)$	True r	12.2	7.7	1.00	11.8	7.8	1.00
	BIC_3	12.2	7.7	1.00	11.8	7.8	1.00
	IC_1	12.2	7.7	1.00	11.8	7.8	1.00
Power							
$(10,100)$	True r	59.2	45.7		60.6	45.4	
	BIC_3	62.5	54.0		63.5	53.2	
	IC_1	64.1	56.0		64.6	53.7	
$(20,100)$	True r	76.3	68.3		75.2	67.0	
	BIC_3	76.3	68.3		75.2	67.0	
	IC_1	76.3	68.3		75.2	67.0	
$(10,300)$	True r	91.2	85.3		90.5	83.8	
	BIC_3	90.5	84.9		89.9	82.6	
	IC_1	84.4	76.2		81.8	73.9	
$(20,300)$	True r	97.7	96.7		96.4	95.1	
	BIC_3	97.7	96.7		96.4	95.1	
	IC_1	97.7	96.7		96.4	95.1	

Notes : For size, t_a^* and t_b^* columns give the percentage of replications in which the null hypothesis of a unit root is rejected for 5% level. The number of factors is either set to 1 (the true number)or estimated using the information criteria suggested by Bai and Ng (2002). The last two columns provide the mean number of estimated factors. For Power, entries represent the percentage of replications in which the null hypothesis of a unit root is rejected.

Table 2: Estimations results

Samples:	*CFA*	*OECD*	*GLOBAL*
Estimated $\hat{\lambda}^*$			
	0.8962	0.9815	1.0000
Stoch. conver. test ($H_0 : \lambda = 1$)			
t_a^*	-4.4704***	-1.4906*	2.0033
	[0.0000]	[0.0680]	[0.9978]
t_b^*	-2.3658***	-1.3779*	2.8484
	[0.0090]	[0.0841]	[0.9978]
β-convergence analyse			
$t_{\lambda_n}^*$	116.77	219.7	
Critical $t_{\lambda_n}^*$ (5%)	16.372	24.305	
implied $\hat{\beta}$	-0.0295	-0.0140	
convergence rate $\hat{\theta}$	11%	1.88%	
half-life $\hat{\tau}$	23	49	
Common factors			
$\hat{r}$	6	6	7
r_{max}	6	8	8

Notes: $\hat{\beta} = \left[\left(\hat{\lambda}^*\right)^T - 1\right]/T$. Following Bai and Ng (2002) the maximum number of factors is set $r_{max} = 8int\left[\left(min\left\{N, T\right\}/100\right)^{1/4}\right]$. The values in brackets correspond p-values. * (resp. **, ***)denote statistically significant at the 10% (resp. 5%, 1%) significance level.

Table 3: Breaks and dates, sample *CFA*

	y_{it}		$y_{it} - \bar{y}_t$	
countries	breaks	dates	breaks	dates
Benin				
Burkina Faso				
Cameroun			+	1988
Congo				
Ivory Coast				
Niger				
Senegal				
Togo				

Notes : numbers and dates of breaks are estimated

following the procedure of Bai and Perron (1998).

We consider the case of a single break.

The sign (+) indicates presence of beak.

Table 4: Breaks and dates, sample *OECD*

	y_{it}		$y_{it} - \bar{y}_t$	
countries	breaks	dates	breaks	dates
Austria				
Belgium				
Canada				
Denmark				
Finland				
France				
Germany				
Greece			+	2002
Hungary				
Ireland				
Italy	+	2003	+	1994
Japan	+	1993	+	1995
Netherlands				
New Zealand				
Norway				
Portugal				
Spain			+	1987
Sweden				
United Kingdom				
United States				

Notes: see notes of table 2

Table 5: Cross-section Dependence (CD)

	CD Statistics					
	y_{it}			$y_{it} - \bar{y}_t$		
ADF(p) regressions	$p = 0$	$p = 1$	$p = 2$	$p = 0$	$p = 1$	$p = 2$
CFA	5.78	4.63	4.65	-3.32	-2.95	-2.63
OECD	25.10	23.89	23.45	-2.88	-2.87	-2.63
GLOBAL	19.72	17.43	17.48	0.53	1.92	2.24

Notes: CD corresponds to the Pesaran's (2004) statistic. The test statistic is based on the average of pair-wise Pearson's correlation coefficients of the estimated residuals from the ADF-type regression equations. We consider different orders p of the ADF regression. The statistic is compared to the standard normal distribution. The null hypothesis of cross-section independence is rejected if $|CD| \succeq 1.96$.