A novel test of time series convergence*

Jose Olmo[§] Marcos Sanso-Navarro[‡] María Vera-Cabello[§]

February 1, 2023

Abstract

This paper extends the definition of time series convergence beyond the conventional parity condition by allowing for positive cointegration ($\beta > 0$, with β the cointegration coefficient) between pairs of unit roots. The paper also develops a novel one-sided test of time series convergence and presents its asymptotic properties under the null hypothesis of no convergence and the alternative hypothesis. The test is robust to general forms of weak dependence in the transitory components and does not require the estimation of the cointegration coefficient. These features are illustrated in a Monte-Carlo simulation exercise for a battery of ARMA(1,1) innovations of the unit root processes. As a byproduct, we propose a methodology to detect convergence clubs. This procedure is based on centrality measures of network dependence given by the degree and betweenness. The empirical application analyzes regional data on population and per-capita income at the NUTS-2 level from France, Italy and Spain. Our results uncover the presence of different convergence clubs for population dynamics and convergence to a single regime for per-capita income.

Keywords: asymptotic theory, hypothesis testing, networks, time series convergence, unit root tests, convergence clubs.

*This work has received financial support from Gobierno de Aragón (Grants LMP21-21 and S39-20R ADETRE Research Group), and Ministerio de Ciencia e Innovacion/Agencia Estatal de Investigación (Grants PID2019-104326GB-I00 and PID2020-112773GB-I00). Jose Olmo also acknowledges financial support from Fundación Agencia Aragonesa para la Investigación y el Desarrollo.

[§]Corresponding author. Address: Departamento de Análisis Económico, Universidad de Zaragoza. Gran Vía 2, 50005, Zaragoza (Spain). Department of Economics, University of Southampton. E-mail: joseolmo@unizar.es and j.b.olmo@soton.ac.uk

[‡]Departamento de Análisis Económico & IEDIS, Universidad de Zaragoza

[§]Centro Universitario de la Defensa de Zaragoza

1 Introduction

One salient topic in the literature on economic growth and development, see Barro (1991) and Mankiw, Romer, and Weil (1992) as seminal contributions, is the analysis of economic convergence (Abramovitz 1986; Baumol 1986). This hypothesis suggests that economies should converge to a single steady-state equilibrium. In practice, however, we observe the occurrence of different equilibria across countries with different macroeconomic conditions and historical backgrounds, leading to the presence of convergence clubs and poverty traps (Azariadis and Drazen 1990; Durlauf and Johnson 1995; Galor 1996; Quah 1996a, 1996b).

Testing for the presence of economic convergence empirically requires a formal definition that varies depending on whether the interest is on cross-sectional analysis or time series studies. In this paper, we revisit the concept of time series convergence (Bernard and Durlauf 1995, 1996; Durlauf, Romer, and Sims 1989; Pesaran 2007), which takes advantage of the concept of long-run dependence and cointegration, see Granger (1981), Engle and Granger (1987), Johansen (1988, 1991), and Johansen and Juselius (1990), as seminal examples. According to this notion, two economies converge if their per-capita income move together in the long run such that the difference is a stationary process. This definition of convergence is appealing both from an economic and a statistical point of view because it combines concepts of economic theory with well known statistical procedures. In fact, and as acknowledged by Kong, Phillips, and Sul (2019), the concept of cointegration is closely related to the concept of convergence.

A standard testing procedure consists of applying residual-based cointegration tests for unit root processes. Prominent examples are Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests, see Said and Dickey (1984) and Phillips and Perron (1988), respectively. As a result, time series convergence tests suffer from the same problems of unit root tests in finite samples, namely, the existence of size distortions under several forms of serial dependence. This has been documented in early studies by Schwert (1989), Ng and Perron (2001) and Perron and Ng (1996). The related literature has developed modifications of the original ADF and PP tests that, under different assumptions, solve the size problem. These methods involve, in general, complex corrections of the test statistic for removing the effect of serial dependence, the choice of suitable estimators of nuisance parameters affecting the test distribution, or the application of modified information criteria to select the truncation lag in residual-based tests.

The first contribution of our study is to develop a test of time series convergence between pairs of unit roots that is based on the idea of cointegration. This definition extends the standard notion of time series convergence by allowing for a cointegration relationship between the variables $(1, -\beta)$ that may be different from the conventional $(1, -1)$ combination postulated by Bernard and Durlauf (1996) and Pesaran (2007), among others. We differentiate between positive $(\beta > 0)$ and negative $(\beta < 0)$ cointegration. Convergence is characterized by the presence of positive cointegration such that both unit roots are generated by a common unit root factor.

To test this form of time series convergence, we propose a hypothesis test similar to the residual-based tests of cointegration extant in the literature. The null hypothesis is composite and given by one of the following two conditions: (i) the processes are two independent unit roots or (ii) there is negative cointegration between the time series $(\beta < 0)$. The alternative hypothesis is characterized by the presence of positive cointegration $(\beta > 0)$. These hypotheses are tested using a novel test statistic constructed as the Euclidean distance between the standardized versions of the unit roots. Under the presence of convergence, the Euclidean distance between both time series converges to zero in probability. The absence of convergence is characterized by two possible scenarios. If the unit roots are stochastically independent, the Euclidean distance between the standardized versions of both processes converges to a limiting distribution that is free of nuisance parameters, and critical values can be universally tabulated. The second scenario is characterized by two unit root processes that are negatively cointegrated. In this case the Euclidean distance converges to four in probability. In contrast to standard residual-based tests of cointegration, our test statistic does not require suitable modifications to account for the presence of serial and mutual correlation between the innovations of the unit root processes.

The finite-sample performance of the convergence test is analyzed in a Monte-Carlo simulation exercise for unit root processes with different degrees of persistence in the transitory components. The empirical size of the test is close to the nominal size at the 5% significance level for all data generating processes. The power of the test to reject the null hypothesis of no convergence is also very high across different sample sizes, achieving values close to one for $T = 100$. Our test is compared with standard ADF and PP tests for different specifications of ARMA(1,1) models. These simulation results illustrate the typical size distortions characterizing the ADF and PP methods and offer strong support to our testing approach for the hypothesis of convergence as an alternative to standard residual-based cointegration tests. It is also important to remark that testing for cointegration is a necessary condition for the presence of convergence between time series but not sufficient.

The second contribution of the study is to extend the analysis of time series convergence to a system of n unit roots. The main purpose of the multivariate analysis is to detect the existence of convergence clubs (see Quah (1997) and Quah (1996a, 1996b)). We define a convergence club as a group of time series with long-run dynamics driven by the same non-stationary common factors and such that the common factors of different clubs are mutually independent. The novelty of our approach is to apply network measures of centrality such as the degree of a node and its betweenness. A node is interpreted as a unit of the set of n unit roots. Two nodes are related by an edge if the corresponding pair of non-stationary time series converges. The betweenness statistic is particularly important in this setting to assess the sensitivity of the formation of clubs to specific units. Nodes with large betweenness belong to several clusters indicating spurious convergence results between different pairs of time series. We remove the spurious edges from the network to reduce the betweenness. This strategy allows us to obtain clusters of time series that are self contained with no spillovers to other clusters satisfying, in turn, our definition of convergence club.

These methodologies are illustrated in an empirical application investigating the existence of economic and demographic regional convergence in France, Italy and Spain at the NUTS-2 level. Our flexible characterization based on the concept of positive cointegration uncovers more convergence relationships than conventional formulations of convergence based on the parity between time series. More specifically, we find convergence clubs for regional population that are not observed when testing for a one-to-one relationship.

The rest of the paper is organized as follows. Section 2 reviews the literature on economic convergence and discusses different measures of convergence and related hypothesis tests.

Section 3 introduces our definition of pairwise convergence and convergence clubs. The section also proposes a statistical test of convergence, and derives its asymptotic properties under the null and alternative hypotheses. In Section 4, we carry out an exhaustive Monte-Carlo simulation exercise. The critical values of the convergence test are tabulated and its finite-sample performance is studied both under the null and alternative hypotheses. The performance of our test is also compared to that of standard residual-based tests. Section 5 contains an empirical application to the study of economic and demographic regional convergence in selected European countries. Section 6 concludes. The mathematical proofs can be found in the appendix.

2 A review of convergence tests of economic growth

The rise of economic growth theory in the 1980s and, especially, the 1990s led to the formulation of two empirical questions: (i) what factors explain observed growth rate differences across countries or regions?, and (ii) do differences between economies decrease over time? (Durlauf, Johnson, and Temple 2005). The latter issue is labelled convergence, and has several statistical notions, see Durlauf, Johnson, and Temple (2009). The first of them, mainly related to neoclassical growth models, is known as the $\beta-$convergence hypothesis and refers to the concept of catching-up. This notion is tested using regression analysis and tries to disentangle whether initial levels of income per capita are inversely related to subsequent growth.

A second notion is $\sigma-$convergence and refers to the decrease in the dispersion of income per capita across economies over time[1]. Although initial studies focused on this evolution, regression-based tests have also been proposed (Cannon and Duck 2000; Friedman 1992). Nonetheless, they have been shown to be difficult to interpret if the data generating process is not invariant, and in the presence of unit roots (Bliss 1999). This connects with the time series approach to convergence[2], which is of a statistical nature and not directly related to any particular growth theory. Therefore, this approach can be applied to other variables

[1] In fact, and as shown by Young, Higgins, and Levy (2008), $\beta-$convergence is a necessary but not a sufficient condition for $\sigma-$convergence.

[2] Another alternative is the approach that consists of analyzing the income distribution and its dynamics; see Quah (1993a, 1993b) and Bianchi (1997), among many others. The main techniques that are applied with this aim are nonparametric and stochastic dominance methods, transition matrices, and mixture models (Durlauf, Johnson, and Temple 2009).

such as prices, wages or unemployment rates. Bernard and Durlauf (1996) introduces a formal definition of time series convergence that is based on the limit of the expected output gap. According to these authors, convergence takes place if the difference in per-capita income between a pair of economies is a stationary stochastic process, property that is tested using cointegration techniques. Bernard and Durlauf (1995, 1996) also show that time series-based tests are associated with a weaker notion of convergence than cross-sectional ones, in terms of the permanent or transitory character of contemporary output differences. Cross-sectional tests are more appropriate for economies that are far from their steady state. Silva Lopes (2016) shows that the power of time series tests depends, to a great extent, on the specification of the deterministic component, see also Carvalho and Harvey (2005) and Harvey and Carvalho (2005).

Hobijn and Franses (2000) establish three alternative definitions of convergence in a time series context and develop an algorithm to endogenously select convergence clubs. Similarly, and in a panel data framework, Phillips and Sul (2007, 2009) propose an algorithm based on a log-t regression for the study of relative $\sigma-$convergence. Under the assumption of a common factor, this convergence notion requires the ratio of two time series to tend to unity in the long run. In contrast to traditional time series tests, Phillips and Sul (2007, 2009)'s method does not suffer from the small sample problems of unit root and cointegration tests, and is appropriate under temporal transitional heterogeneity. Kong, Phillips, and Sul (2019) introduce the notion of weak $\sigma-$convergence in order to capture convergent behavior in panel data that does not involve stochastic or divergent deterministic trends. Beylunioğlu, Yazgan, and Stengos (2020) develop a set of statistical criteria for club formation combining unit root tests and graph theory concepts (cliques).

Pesaran (2007) extends the definition of convergence in Bernard and Durlauf (1996) by introducing the concept of probabilistic convergence. This probabilistic version of output convergence does not assume that the economies are identical, but the time series should be cotrended and cointegrated with vector $(1 \;\; -1)$. To study this notion, Pesaran (2007) proposes a convergence test that considers all possible pairs of output gaps across economies, not requiring to establish a reference unit. These results are exploited to establish convergence clubs.

A recent framework to characterize time series convergence is proposed by Garcia-Hiernaux and Guerrero (2021). These authors consider convergence as steady-state or catching-up, as well as both strong and weak versions, and model a wide range of transition paths. In a price context, these authors claim that two time series converge if they are cointegrated with vector $(1 \ -\beta)$. Nevertheless, they assume that $\beta = 1$ – i.e., goods are homogeneous with a extremely large elasticity of substitution – in order to base their methodology on the implementation of univariate unit root tests, rather than on cointegration analysis. The following section extends this approach by proposing a test of cointegration that requires the cointegration coefficient to be positive.

3 A novel test of time series convergence

This section presents a novel test of convergence between pairs of unit roots that is robust to the presence of weak dependence in the innovations of the unit root processes. Another novel feature of the test is that it does not require estimation of the cointegration coefficients. Before introducing the test we motivate it and provide some background and definitions.

3.1 Background

We consider n time series $x_t = (x_{1t}, \ldots, x_{nt})'$ and study pairwise convergence. A general specification of any pair of unit roots in the system is

$$\begin{cases} x_{it} & = \alpha_{x_i} + \theta_i f_{it} + u_{it} \\ f_{it} & = \pi_{if} + f_{i,t-1} + v_{it}, \end{cases} \tag{1}$$

where f_{it} are the unit root processes generating the stochastic trend in the processes x_{it}; u_{it} and v_{it} are the corresponding innovations to each process that are assumed to be mutually independent for all leads and lags. Under this assumption the innovations u_{it} characterize the transitory component of x_{it}. The vector of innovations $u_t = (u_{1t}, \ldots, u_{nt})$ exhibits weak dependence and is defined as $u_t = C(L)\varepsilon_t$, where $C(z) = \sum_{j=0}^{\infty} z^j$ with $C(0) = I_n$ (the $n \times n$ identity matrix); $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})'$ is an *iid* vector normally distributed with mean zero and covariance matrix Ω_ε. Furthermore, we impose $\sum_{j=1}^{\infty} j|C_j| < \infty$ to limit the amount of serial dependence in u_t. The above process generates a deterministic trend such that each component of the vector of time series x_t can be expressed as

$$x_{it} = \alpha_{x_i} + \pi_i t + \theta_i f_{it}^o + u_{it}, \tag{2}$$

where $\pi_i = \pi_{if}\theta_i$ and $f_{it}^o = \sum_{s=1}^{t} v_{is}$ is a random walk process without drift. In this setting, convergence is defined as follows.

Definition 1: *The unit root processes x_{it} and x_{jt} converge if there exists a positive coefficient β_{ij} such that $x_{it} - \beta_{ij}x_{jt}$ is stationary.*

Using this definition, we characterize convergence between a pair of unit roots (x_{it}, x_{jt}) by the existence of a common factor $f_{ij,t} = \pi_{ij,f} + f_{ij,t-1} + v_{ij,t}$ driving the deterministic and stochastic trends of both processes. Then,

$$\begin{cases} x_{it} = \alpha_{x_i} + \pi_i t + \theta_i f_{ij,t}^o + u_{it}, \\ x_{jt} = \alpha_{x_j} + \pi_j t + \theta_j f_{ij,t}^o + u_{jt}, \end{cases} \tag{3}$$

where $\pi_i = \pi_{ij,f}\theta_i$, $\pi_j = \pi_{ij,f}\theta_j$ and $f_{ij,t}^o = \sum_{s=1}^{t} v_{ij,s}$ is a random walk without drift. The properties of the innovation sequences are defined in model (1). The convergence hypothesis entails the following relationship between the parameters: $\theta_i = \beta_{ij}\theta_j$ and $\pi_i = \beta_{ij}\pi_j$, with $\beta_{ij} > 0$.

3.2 Testing pairwise convergence

Let $z_{ij,t}(\beta) = x_{it} - \beta x_{jt}$. The hypothesis of pairwise convergence is represented as

$$\begin{cases} H_0 : z_{ij,t}(\beta) \sim I(1), \text{ for all } \beta \in \mathbb{R} \quad OR \quad z_{ij,t}(\beta) \sim I(0), \text{ for some } \beta \in \mathbb{R}^-, \\ H_A : z_{ij,t}(\beta) \sim I(0), \text{ for some } \beta \in \mathbb{R}^+. \end{cases} \tag{4}$$

In the same spirit of residual-based tests such as Dickey-Fuller and Engle-Granger procedures, see also Pesaran (2007) in the context of time series convergence, the null hypothesis corresponds to absence of convergence and the alternative hypothesis to the existence of time series convergence.

This section proposes a statistical test of time series convergence based on the Euclidean distance between the standardized versions of the unit root processes x_{it} and x_{jt}. The test is based on the recent contribution of Olmo (2022) that develops a simple test of cointegration that accommodates weak dependence in the innovation sequences. In this paper, we specialize this test and adapt it to the hypothesis of pairwise convergence between two unit root processes. A major advantage of this test is that knowledge or estimation of the cointegration coefficient is not required. This is possible by the standardization of the unit root processes x_{it} and x_{jt} prior to testing for cointegration.

Let x_{rt}^d be the detrended process of x_{rt} defined as $x_{rt}^d = x_{rt} - \alpha_{x_r} - \pi_r t$, for $r = i, j$, and let $\widetilde{x}_{rt}$ be its sample counterpart defined as $\widetilde{x}_{rt} = x_{rt} - \overline{x}_r - \widehat{\pi}_r \left(t - \frac{T+1}{2}\right)$, for $r = i, j$, with $\overline{x}_r$ the sample mean and $\widehat{\pi}_r$ the OLS estimator of π_r obtained from the regression of x_{rt} on $(1, t)$. Similarly, let $\widehat{\sigma}^2_{\widetilde{x}_r} = \frac{1}{T}\sum_{t=1}^{T} \widetilde{x}^2_{rt}$ be the sample variance such that the standardized unit root process is $y_{rt} = \frac{\widetilde{x}_{rt}}{\widehat{\sigma}_{\widetilde{x}_r}}$. The proposed test statistic for the hypothesis of pairwise convergence is

$$\widehat{D}_{ij,T} = \frac{1}{T}\sum_{t=1}^{T} \left(y_{it} - y_{jt}\right)^2. \tag{5}$$

Proposition 1. If the processes x_{it} and x_{jt} are defined as in (3) with $\beta_{ij} = \theta_i/\theta_j > 0$, then $\widehat{D}_{ij,T} \xrightarrow{p} 0$ as $T \to \infty$.

The absence of convergence between the processes x_{it} and x_{jt} yields different asymptotic results depending on the relationship between the unit roots. If these processes are negatively cointegrated ($\beta_{ij} < 0$), we obtain the following result.

Corollary 1. If the processes x_{it} and x_{jt} are defined as in (3) with $\beta_{ij} = \theta_i/\theta_j < 0$, then $\widehat{D}_{ij,T} \xrightarrow{p} 4$ as $T \to \infty$.

In contrast, if the time series x_{it} and x_{jt} are characterized by two independent unit root processes f^o_{it} and f^o_{jt} as in (2), then the test statistic $\widehat{D}_{ij,T}$ has the following limiting distribution.

Proposition 2. Let $x_{rt} = \alpha_{x_r} + \pi_r t + \theta_r f^o_{rt} + u_{rt}$ for $r = i, j$, with f^o_{it} and f^o_{jt} two mutually independent unit root processes. Let u_{it} and u_{jt} be the corresponding innovation sequences exhibiting weak and mutual dependence as described in (2). Then,

$$\widehat{D}_{ij,T} \xrightarrow{d} 2\left(1 + B_\pi A_{\pi_i} A_{\pi_j} - B_\pi Z_{ij}\right), \tag{6}$$

with

$$Z_{ij} = \frac{\int_0^1 W_i(r)W_j(r)dr - \int_0^1 W_i(r)dr \int_0^1 W_j(\tau)d\tau}{\left[\int_0^1 W_i(r)^2 dr - \left(\int_0^1 W_i(r)dr\right)^2\right]^{1/2} \left[\int_0^1 W_j(r)^2 dr - \left(\int_0^1 W_j(r)dr\right)^2\right]^{1/2}}, \tag{7}$$

$A_{\pi_i} = \frac{\sqrt{12}\int_0^1 (r-\frac{1}{2})W_i(r)dr}{\left[\int_0^1 W_i(r)^2 dr - \left(\int_0^1 W_i(r)dr\right)^2\right]^{1/2}}$, $B_\pi = (1 - A^2_{\pi_i})^{-1/2}(1 - A^2_{\pi_j})^{-1/2}$, and $W_i(r)$ and $W_j(r)$ are two independent Brownian motions.

The asymptotic distribution in (6) does not depend on nuisance parameters such as the long-run variance of the transitory components u_{it} and u_{jt}. These asymptotic results enable a simple testing procedure for the hypothesis of pairwise convergence. Thus, there is evidence to reject the hypotheses of (i) independent unit roots and (ii) negative cointegration if $\widehat{D}_{ij,T} < c_\alpha$, with c_α the $\alpha-$quantile of the asymptotic distribution (6), and α the corresponding significance level.

Critical values of the one-sided test of time series convergence are tabulated by simulation. These values are obtained by computing the test statistic $\widehat{D}_{ij,T}$ for B draws of two independent unit root processes $x_{it} = x_{i,t-1} + v_{it}$, with $v_{it} \sim WN(0,1)$. Importantly, the test is robust to the presence of serial correlation in the innovations v_{it}, thus, it is not necessary to account for such dependence in the simulation of the critical values.

3.3 Network measures and convergence clubs

The seminal contribution of Pesaran (2007) to study time series convergence also discusses multi-country convergence. This definition requires pairwise convergence across the n time series in the system. The scope of this approach is limited in practice as it can efficiently handle only a small number of series simultaneously. To be able to analyze the convergence properties of a large number of units, whilst at the same time avoiding the pitfalls that surround the use of a given benchmark, Pesaran (2007) adopts a pairwise approach that considers the unit root and trending properties of all $n(n-1)/2$ possible combinations.

This section discusses an alternative definition of convergence clubs and proposes a novel approach based on network statistics to detect these formations. We also assess the sensitivity of the clusters to the influence of specific time series. More formally, a convergence club is defined as a group of units with long-run dynamics driven by the same vector of common factors. Let $x_t = (x_t^{(1)}, \dots, x_t^{(K)})'$ with $x_t^{(k)} = (x_{1t}^{(k)}, \dots, x_{n_k t}^{(k)})'$ a group of n_k series forming cluster k and such that $\sum_{k=1}^{K} n_k = n$. The time series in cluster k have the following specification:

$$x_t^{(k)} = \alpha_x^{(k)} + \pi^{(k)} t + \theta^{(k)} f_t^{o(k)} + u_t^{(k)}, \quad \text{for } k = 1, \dots, K, \tag{8}$$

with $\alpha_x^{(k)} = (\alpha_{x_1}^{(k)}, \dots, \alpha_{x_{n_k}}^{(k)})'$ and $\pi^{(k)} = (\pi_1^{(k)}, \dots, \pi_{n_k}^{(k)})'$ the vectors of deterministic components and $\theta^{(k)}$ a $1 \times n_k$ vector of factor loadings associated to the scalar non-stationary common factor $f_t^{o(k)}$. The common factors driving the long-run dynamics of different clusters are mutually independent such that any linear combination between them is a unit root. The error term for each cluster is defined as $u_t^{(k)} = (u_{1t}^{(k)}, \dots, u_{n_k t}^{(k)})'$ and may exhibit weak dependence as defined in (2).

To identify the clusters of time series, we follow a procedure similar in spirit to the method proposed in Pesaran (2007). We construct an $n \times n$ interaction matrix reflecting all the pairwise combinations between the time series. Let $1(\widehat{D}_{ij,T})$ be an indicator function that takes a value of one if the hypothesis of convergence is not rejected for (x_{it}, x_{jt}). This is given by the condition $\widehat{D}_{ij,T} < c_\alpha$. Otherwise, the indicator function takes a value of zero in the entry (i, j) of the interaction matrix. The entries of the interaction matrix can be interpreted as a network. Each time series in the vector x_t is a node. Two time series share an edge if their dynamics converge in the long run. The degree of a node is defined as the number of connections with the rest of the network. More formally, let $Z_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} 1(\widehat{D}_{ij,T})$. In self-contained clusters without spillovers to other clusters, this statistic is the same for all time series in the same cluster and such that $Z_i = n_k - 1$. In this setting, the interaction matrix accepts a representation into a $K \times K$ block-diagonal matrix, with each block given by an $n_k \times n_k$ submatrix of ones and comprised by all time series in the same cluster (convergence club).

In practice, some time series are grouped into two or more different clusters. This fact contradicts the definition of convergence club in (8) that establishes a unique one-to-one relationship between the K clusters of time series and the common factors $f_t^{o(k)}$ for $k = 1, \dots, K$. This apparent contradiction is empirically possible due to the occurrence of type I error of the test (5). This error is interpreted as not rejecting the hypothesis of convergence between two time series H_A when these series are in fact not positively cointegrated. This may be due to the presence of negative cointegration between x_{it} and x_{jt} or simply to the absence of any long-run relationship between the variables exhibiting unit root behavior. The presence of these influencing nodes linking different clusters can be detected in the interaction matrix as those entries with a value of one located outside the block-diagonal matrices. The following matrices illustrate this for 8 time series grouped into four clusters. The left matrix corresponds to a network with four self-contained groups. In contrast, in the right matrix, clusters 1 and 2 given by (x_1, x_2) and (x_3, x_4), respectively, are related through the edge between x_1 and x_4 in bold font. This entry suggests the presence of convergence between the time series $\{x_1, x_3, x_4\}$, which contradicts the fact that the pairs (x_1, x_3) and (x_2, x_3) do not converge.

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ \mathbf{1} & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

The betweenness of a network is a useful empirical measure to detect the presence of these influencing observations spuriously linking different clusters. This network statistic measures the proportion of shortest paths (fewest steps from i to j) containing a given node h. To formalize this concept, let $s_h(i,j)$ denote the proportion of shortest paths containing node h such that $B(h)$ denotes the corresponding betweenness of that node. More formally,

$$B(h) = \sum_{\substack{i=1 \\ i \neq h}}^{n} \sum_{\substack{j=1 \\ j \neq h,i}}^{n} s_h(i,j). \tag{9}$$

This statistic measures how much influence a node has over connections between others. In the above example the betweenness is close to zero for all the nodes in the left matrix. In contrast, the betweenness of nodes 1 and 4, corresponding to time series x_1 and x_4, is greater than zero. This value of the betwenness statistic is due to the edge relating x_1 and x_4. To correct for the spurious presence of time series in more than one convergence club a possibility is to remove from the interaction matrix those edges outside the block-diagonal matrices.

4 Monte-Carlo simulation

This section studies the finite-sample properties of the convergence test developed above. The test is characterized by the statistic $\widehat{D}_T$ and the critical values are obtained by simulation of two independent unit root processes as discussed below. The aim of the simulation exercise is to investigate the performance of the test under different degrees of time series persistence in the innovations of the unit root processes, and compare it against standard unit root tests for the hypothesis of cointegration such as the Augmented Dickey-Fuller (ADF) test, see Said and Dickey (1984), and the Phillips-Perron (PP) test in Phillips and Perron (1988).

To tabulate the critical values of the convergence test, we consider the following data generating process characterizing the null hypothesis. Let

$$x_{it} = x_{i,t-1} + v_{it}, \text{ with } v_{it} \sim N(0,1), \text{ for } i = 1, 2,$$

with v_{1t} and v_{2t} mutually independent. To avoid the effect of the initial values of the different processes, we discard the first 500 observations from each of them. The critical values are the empirical percentiles of the simulated distribution of the test statistic $\widehat{D}_T$ defined in (5), computed from a sample size $T = 5000$ and $B = 1000$ iterations. The critical values at 1%, 5% and 10% significance level are 0.3944, 0.6498 and 0.9013, respectively. We should note that these critical values are used under both *iid* and general forms of weak dependence in the innovation vector u_t in (1).

Table 1 reports the empirical size and power of the ADF and PP tests, both allowing for two lags, and of $\widehat{D}_T$. Under the null hypothesis ($\beta = 0$), there is no time series convergence and x_{1t} and x_{2t} are two independent unit root processes. The data generating process

(DGP) is

$$x_{1t} = f^o_{1t} + u_{1t},$$
$$x_{2t} = f^o_{2t} + u_{2t},$$

with $f^o_{it} = f^o_{i,t-1} + v_{it}$, for $i = 1, 2$, where $v_{it} \sim$ iid $N(0,1)$ are mutually independent, and u_{1t} and u_{2t} are the innovations with the following ARMA(1,1) structure

$$u_{it} = \rho u_{i,t-1} + \varepsilon_{it} + \theta \varepsilon_{i,t-1},$$

for both sequences; $\varepsilon_{1t}, \varepsilon_{2t}$ are two *iid* $N(0,1)$ random variables and $\{\rho, \theta\}$ are the coefficients characterizing both ARMA(1,1) processes. The alternative hypothesis of convergence is represented by the processes

$$x_{1t} = f^o_t + u_{1t},$$
$$x_{2t} = \beta f^o_t + u_{2t},$$

with $f^o_t = f^o_{t-1} + v_t$ a unit root process without drift characterizing the common stochastic trend driving the dynamics of both processes. Importantly, the cointegration coefficient β needs to be positive, otherwise, the two processes are cointegrated but do not converge. We show below the power of the test under $\beta < 0$, which is close to zero, confirming the validity of the test statistic under this characterization of the null hypothesis.

We closely follow the simulation exercises in Perron and Ng (1996) and Ng and Perron (2001), and consider a battery of ARMA(1,1) processes given by all possible combinations of $\rho = \{-0.8, -0.5, 0, 0.5, 0.8\}$ and $\theta = \{-0.8, -0.5, 0, 0.5, 0.8\}$. These combinations include the case of *iid* innovations and also strong persistence given by values of the parameters close to the unit circle. As discussed by the above authors, it is well documented that both ADF and PP tests suffer finite-sample distortions for the moving-average polynomial with a large negative root and also when the autoregressive root takes a large positive value. The simulation exercise in this section shows that this problem is also present when testing for time series convergence between two unit root processes using the ADF and PP tests, but not with the test $\widehat{D}_T$ introduced above.

Table 1: Empirical size of convergence tests for different values of β under the null hypothesis H_0 at 5% significance level.

ρ	T	ADF	PP	$\widehat{D}_T$	ADF	PP	$\widehat{D}_T$	ADF	PP	$\widehat{D}_T$	ADF	PP	$\widehat{D}_T$	ADF	PP	$\widehat{D}_T$
			$\theta = -0.8$			$\theta = -0.5$			$\theta = 0$			$\theta = 0.5$			$\theta = 0.8$	
								$\beta = 0$								
-0.8	50	0.02	0.90	0.01	0.04	0.88	0.01	0.04	0.74	0.03	0.07	0.45	0.05	0.05	0.28	0.06
-0.8	100	0.03	0.90	0.01	0.02	0.88	0.02	0.06	0.69	0.06	0.06	0.35	0.08	0.09	0.24	0.06
-0.8	500	0.05	0.87	0.04	0.02	0.83	0.05	0.05	0.59	0.06	0.07	0.28	0.08	0.07	0.20	0.09
-0.5	50	0.03	0.76	0.03	0.04	0.71	0.04	0.05	0.47	0.05	0.07	0.29	0.06	0.05	0.22	0.04
-0.5	100	0.04	0.75	0.04	0.05	0.63	0.06	0.06	0.44	0.07	0.07	0.23	0.07	0.08	0.24	0.08
-0.5	500	0.05	0.65	0.09	0.05	0.50	0.07	0.07	0.32	0.08	0.09	0.21	0.09	0.08	0.19	0.10
0.0	50	0.05	0.55	0.05	0.04	0.41	0.06	0.04	0.26	0.06	0.07	0.19	0.05	0.08	0.26	0.04
0.0	100	0.06	0.45	0.07	0.04	0.38	0.06	0.08	0.27	0.09	0.08	0.21	0.09	0.10	0.25	0.06
0.0	500	0.06	0.35	0.08	0.04	0.29	0.06	0.08	0.22	0.08	0.10	0.22	0.09	0.09	0.24	0.09
0.5	50	0.06	0.34	0.06	0.07	0.31	0.07	0.08	0.20	0.07	0.09	0.13	0.06	0.11	0.14	0.04
0.5	100	0.06	0.34	0.06	0.06	0.23	0.08	0.09	0.18	0.09	0.12	0.20	0.05	0.16	0.23	0.05
0.5	500	0.08	0.26	0.08	0.09	0.18	0.09	0.14	0.21	0.09	0.16	0.24	0.08	0.19	0.29	0.05
0.8	50	0.07	0.29	0.06	0.08	0.22	0.06	0.06	0.09	0.07	0.09	0.08	0.04	0.08	0.07	0.05
0.8	100	0.05	0.23	0.07	0.10	0.22	0.06	0.12	0.15	0.07	0.14	0.13	0.05	0.16	0.14	0.02
0.8	500	0.07	0.20	0.08	0.07	0.15	0.08	0.16	0.20	0.07	0.28	0.32	0.05	0.29	0.33	0.07
								$\beta = -1$								
-0.8	50	0.74	1.00	0.00	0.76	1.00	0.00	0.87	1.00	0.00	0.99	1.00	0.00	0.95	1.00	0.00
-0.8	100	0.88	1.00	0.00	0.91	1.00	0.00	0.99	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
-0.8	500	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
-0.5	50	0.92	1.00	0.00	0.96	1.00	0.00	0.97	1.00	0.00	0.96	1.00	0.00	0.86	1.00	0.00
-0.5	100	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
-0.5	500	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
0.0	50	1.00	1.00	0.00	1.00	1.00	0.00	0.97	1.00	0.00	0.90	1.00	0.00	0.73	1.00	0.00
0.0	100	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
0.0	500	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
0.5	50	1.00	1.00	0.00	0.97	1.00	0.00	0.69	0.99	0.00	0.49	0.77	0.00	0.41	0.71	0.00
0.5	100	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	0.99	1.00	0.00	0.94	1.00	0.00
0.5	500	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
0.8	50	0.96	1.00	0.00	0.46	0.99	0.00	0.23	0.40	0.00	0.20	0.14	0.00	0.10	0.12	0.00
0.8	100	1.00	1.00	0.00	0.97	1.00	0.00	0.71	0.90	0.00	0.56	0.64	0.00	0.41	0.60	0.00
0.8	500	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
								$\beta = -2$								
-0.8	50	0.96	1.00	0.00	0.95	1.00	0.00	0.99	1.00	0.00	0.98	1.00	0.00	0.97	1.00	0.00
-0.8	100	0.99	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
-0.8	500	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
-0.5	50	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	0.97	1.00	0.00	0.92	1.00	0.00
-0.5	100	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
-0.5	500	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
0.0	50	1.00	1.00	0.00	1.00	1.00	0.00	0.98	1.00	0.00	0.84	1.00	0.00	0.78	1.00	0.00
0.0	100	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
0.0	500	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
0.5	50	1.00	1.00	0.00	0.97	1.00	0.00	0.72	1.00	0.00	0.54	0.79	0.00	0.39	0.72	0.00
0.5	100	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	0.98	1.00	0.00	0.94	1.00	0.00
0.5	500	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
0.8	50	0.98	1.00	0.00	0.54	0.99	0.00	0.25	0.41	0.00	0.18	0.18	0.00	0.13	0.12	0.00
0.8	100	1.00	1.00	0.00	0.96	1.00	0.00	0.75	0.93	0.00	0.58	0.66	0.00	0.46	0.62	0.00
0.8	500	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00

Table 1 reports the size of the ADF and PP tests applied to the residuals of the cointegration equation between x_{1t} and x_{2t}. Both methodologies consider two lags and are computed using Matlab routines. The null hypothesis is represented by two different scenarios: *i*) $\beta = 0$, that corresponds to two independent unit root processes, and *ii*) $\beta < 0$, that corresponds to a cointegration relationship between x_{1t} and x_{2t} given by a negative coefficient. The top panel of Table 1 reports the empirical size of the tests for $\beta = 0$ and all possible combinations of ARMA(1,1) processes under consideration. The results confirm the strong size distortions of the PP test across values of the parameter space, and smaller but still significant distortions for the ADF test. This is particularly the case for those combinations given by the joint condition $\rho, \theta > 0$. In contrast, the columns for the one-sided test $\widehat{D}_T$ report estimates of the nominal size $\alpha = 0.05$ in the interval $(0.03, 0.09)$ across models and sample sizes. These results show an adequate performance of the test $\widehat{D}_T$ when the unit root processes are independent.

These empirical findings provide strong support to the approximation of the asymptotic critical value at the 5% significance level given by 0.6498, obtained by simulation of two independent unit root processes. Related to this is the study of the performance of the test when there is negative cointegration between x_{1t} and x_{2t}. Corollary 1 shows that, for $\beta < 0$, the test statistic $\widehat{D}_T$ converges to 4 in probability as the sample size increases. The large difference between the critical value 0.6498 and the limit of the test statistic for $\beta < 0$ that is equal to four implies a strong performance of the test in these cases. The middle and bottom panels in Table 1 report the rejection probability of the different tests for $\beta = -1$ and $\beta = -2$. The columns corresponding to $\widehat{D}_T$ report values of zero, implying that negative values of β are correctly classified under the null hypothesis of no convergence. In contrast, the ADF and PP tests obtain values close to one, implying the rejection of the null hypothesis of no cointegration. In order for these tests to be meaningful in this scenario, one needs to jointly assess the p-value of the ADF and PP tests along with the sign of the cointegration coefficient such that negative and statistically significant values of β can be identified with the absence of time series convergence. Importantly, in contrast to these methods, the test statistic $\widehat{D}_T$ does not require knowledge nor estimation of the cointegration coefficient β for testing the convergence hypothesis.

Table 2 reports the empirical power for two models that reflect convergence between the time series x_{1t} and x_{2t}. Both processes are positively cointegrated. The case $\beta = 1$ in the top panel corresponds to the conventional hypothesis of convergence between two unit root processes and is given by the condition $x_{1t} - x_{2t}$ being stationary. The three methods (ADF, PP and $\widehat{D}_T$) perform very satisfactorily under this scenario with the empirical power increasing with the sample size. It is worth noting, though, that the power of the

PP test is not size-adjusted so it may overestimate the true power of the test for rejecting the null hypothesis under the presence of convergence between the variables. The bottom panel reports the empirical power for $\beta = 2$. The results are similar to the empirical power reported in the upper panel but the magnitude of the power is slightly higher as the cointegration coefficient increases.

Table 2: Empirical power of convergence tests at 5% significance level.

ρ	T	ADF	PP	$\widehat{D}_T$	ADF	PP	$\widehat{D}_T$	ADF	PP	$\widehat{D}_T$	ADF	PP	$\widehat{D}_T$	ADF	PP	$\widehat{D}_T$
			$\theta = -0.8$			$\theta = -0.5$			$\theta = 0$			$\theta = 0.5$			$\theta = 0.8$	
								$\beta = 1$								
-0.8	50	0.73	1.00	0.21	0.75	1.00	0.32	0.89	1.00	0.59	0.96	1.00	0.86	0.95	1.00	0.91
-0.8	100	0.89	1.00	0.41	0.94	1.00	0.54	0.99	1.00	0.86	1.00	1.00	0.99	1.00	1.00	0.99
-0.8	500	1.00	1.00	0.95	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.5	50	0.95	1.00	0.53	0.95	1.00	0.64	0.98	1.00	0.87	0.96	1.00	0.93	0.90	1.00	0.89
-0.5	100	1.00	1.00	0.77	1.00	1.00	0.89	1.00	1.00	0.98	1.00	1.00	0.99	1.00	1.00	0.99
-0.5	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0	50	0.99	1.00	0.80	1.00	1.00	0.85	0.96	1.00	0.93	0.86	1.00	0.87	0.71	1.00	0.80
0.0	100	1.00	1.00	0.96	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.96
0.0	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	50	1.00	1.00	0.90	0.95	1.00	0.92	0.70	0.99	0.84	0.56	0.78	0.69	0.38	0.68	0.50
0.5	100	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	0.98	0.98	1.00	0.90	0.95	1.00	0.78
0.5	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	50	0.96	1.00	0.92	0.44	0.99	0.88	0.26	0.46	0.67	0.16	0.15	0.43	0.09	0.09	0.31
0.8	100	1.00	1.00	0.99	0.97	1.00	0.98	0.74	0.92	0.85	0.59	0.68	0.63	0.41	0.59	0.45
0.8	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.96
								$\beta = 2$								
-0.8	50	0.93	1.00	0.37	0.96	1.00	0.51	0.98	1.00	0.80	0.99	1.00	0.97	0.96	1.00	0.99
-0.8	100	1.00	1.00	0.61	1.00	1.00	0.76	1.00	1.00	0.97	1.00	1.00	1.00	1.00	1.00	1.00
-0.8	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.5	50	0.99	1.00	0.73	1.00	1.00	0.83	1.00	1.00	0.96	0.97	1.00	0.99	0.90	1.00	0.98
-0.5	100	1.00	1.00	0.91	1.00	1.00	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.5	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0	50	1.00	1.00	0.94	1.00	1.00	0.97	0.95	1.00	0.99	0.85	1.00	0.97	0.72	1.00	0.93
0.0	100	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99
0.0	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	50	1.00	1.00	0.97	0.98	1.00	0.99	0.73	0.99	0.96	0.55	0.77	0.88	0.40	0.68	0.76
0.5	100	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.98	0.95	1.00	0.93
0.5	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	50	0.97	1.00	0.98	0.48	1.00	0.96	0.26	0.44	0.82	0.18	0.17	0.61	0.12	0.10	0.47
0.8	100	1.00	1.00	1.00	0.98	1.00	1.00	0.76	0.92	0.96	0.61	0.65	0.81	0.43	0.64	0.70
0.8	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99

5 Empirical application

This section explores the application of our novel test to uncover the presence of convergence clubs in economic (per-capita income) and demographic (population) variables. With this aim, we have extracted data from Eurostat for the set of NUTS-2 regions comprising France, Italy and Spain over the period 1980 to 2021. These countries are divided into a large number of NUTS-2 regions: France contains 27 regions, Italy is divided into 21 regions, and Spain into 19 regions. The regions are reported in Table 3.

Table 3: NUTS-2 regions for France, Italy and Spain

	France	Italy	Spain
1	Ile de France	Piemonte	Galicia
2	Centre - Val de Loire	Valle d'Aosta/Vallée d'Aoste	Principado de Asturias
3	Bourgogne	Liguria	Cantabria
4	Franche-Comté	Lombardia	PaÃs Vasco
5	Basse-Normandie	Abruzzo	Comunidad Foral de Navarra
6	Haute-Normandie	Molise	La Rioja
7	Nord-Pas de Calais	Campania	Aragón
8	Picardie	Puglia	Comunidad de Madrid
9	Alsace	Basilicata	Castilla y León
10	Champagne-Ardenne	Calabria	Castilla-la Mancha
11	Lorraine	Sicilia	Extremadura
12	Pays de la Loire	Sardegna	Cataluña
13	Bretagne	Provincia Autonoma di Bolzano/Bozen	Comunidad Valenciana
14	Aquitaine	Provincia Autonoma di Trento	Illes Balears
15	Limousin	Veneto	Andalucía
16	Poitou-Charentes	Friuli-Venezia Giulia	Región de Murcia
17	Languedoc-Roussillon	Emilia-Romagna	Ciudad Autónoma de Ceuta
18	Midi-Pyrénées	Toscana	Ciudad Autónoma de Melilla
19	Auvergne	Umbria	Islas Canarias
20	Rhone-Alpes	Marche	
21	Provence-Alpes-Cote d'Azur	Lazio	
22	Corse		
23	Guadeloupe		
24	Martinique		
25	Guyane		
26	La Réunion		
27	Mayotte		

For illustrative purposes, we report first the results of the conventional tests of pairwise convergence proposed in the literature and based on the existence of parity between the time series. These tests rely on the unit root hypothesis for the difference between time series ($x_{it} - x_{jt} = c + \varepsilon_{ij,t}$, with c a constant and $\varepsilon_{ij,t}$ an error term) and are usually implemented through the ADF statistic. Figure 1 presents the case of Italy and Figure 2 describes the case of Spain. Top panels report the analysis of per-capita income and bottom panels report the analysis of population. The construction of convergence clubs is done as discussed above (see also Pesaran (2007)). For each variable and country, we construct an $n \times n$ interaction matrix reflecting all the pairwise combinations between per-capita income or population, with n the number of NUTS-2 regions. The element (i, j) of the interaction matrix takes a value of one if the stationarity condition for $\varepsilon_{ij,t}$ above is not rejected using the ADF test. The interaction matrix is used to construct a graph reflecting all the combinations as edges of a network. The maps in these figures represent with the same color those regions that share edges, forming clusters. The analysis in Figures 1 and 2 suggests that the presence of conventional convergence in per-capita income and population is very limited across regions for the two countries. Unreported results for the case of France show similar findings, however, for this country, per-capita income does not exhibit any cointegration relationship across NUTS-2 regions.

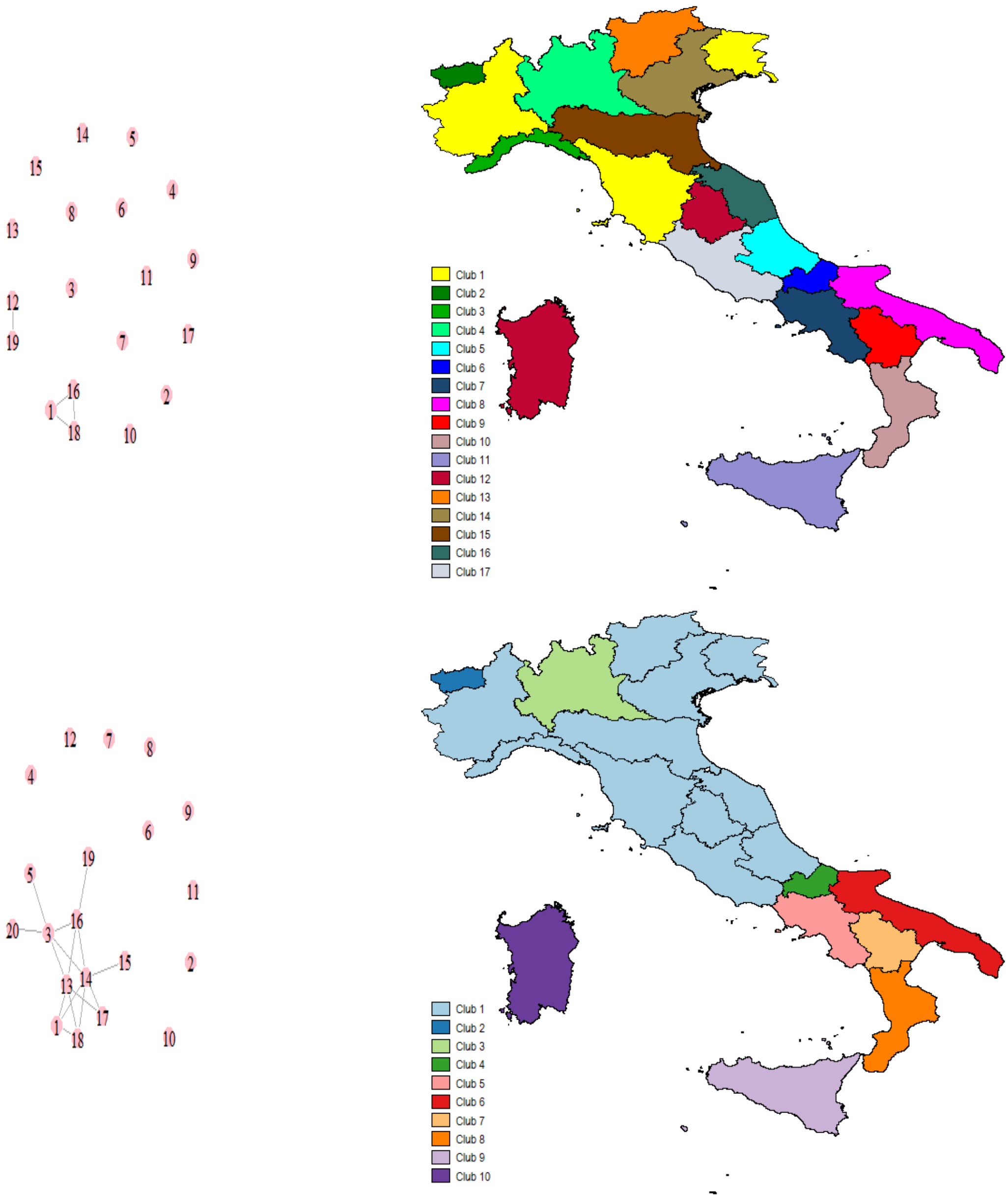

Figure 1: These panels present the NUTS-2 regions of Italy in several clusters characterized by pairwise convergence tests given by the hypothesis $x_{it} = x_{jt}$, with i, j denoting different NUTS-2 regions. Top panels for per-capita income and bottom panels for population. Left panel for the network representation of the interaction matrix reflecting the pairwise relationships between the regions uncovered by the ADF test for the residuals of the regression $x_{1t} - x_{2t} = c + \varepsilon_t$ at 1% significance level. Right panel for the representation of these relationships as a map.

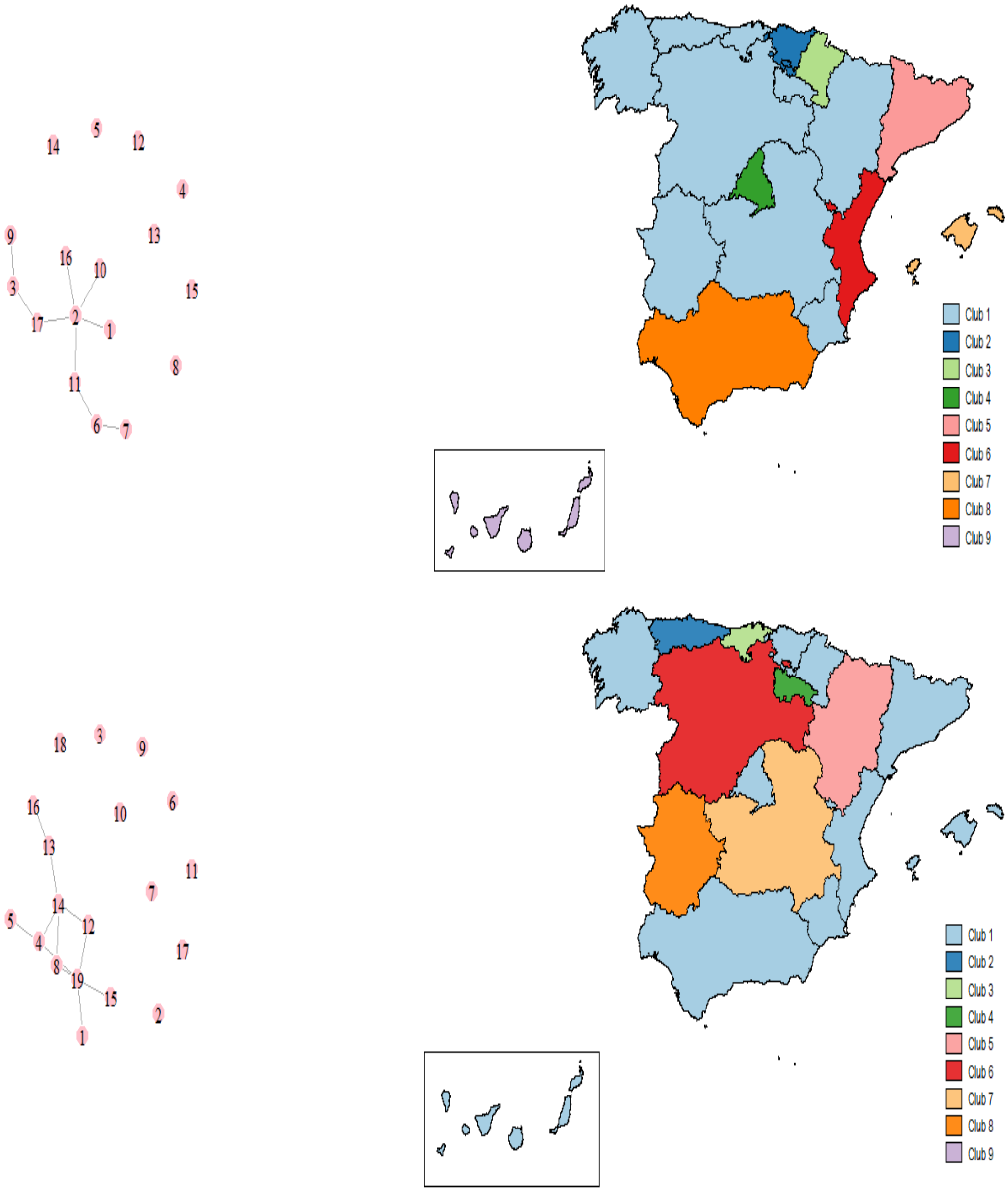

Figure 2: These panels present the NUTS-2 regions of Spain in several clusters characterized by pairwise convergence tests given by the hypothesis $x_{it} = x_{jt}$, with i, j denoting different NUTS-2 regions. Top panels for per-capita income and bottom panels for population. Left panel for the network representation of the interaction matrix reflecting the pairwise relationships between the regions uncovered by the ADF test for the residuals of the regression $x_{1t} - x_{2t} = c + \varepsilon_t$ at 1% significance level. Right panel for the representation of these relationships as a map.

In what follows, we repeat the exercise but allowing for a more flexible definition of pairwise convergence given by the combination $(1, -\beta)$, with β a positive coefficient. The testing procedure is as before: (i) compute the test statistic $\widehat{D}_{ij,T}$ for all pairwise combinations of regions in each country; (ii) construct the interaction matrix and network

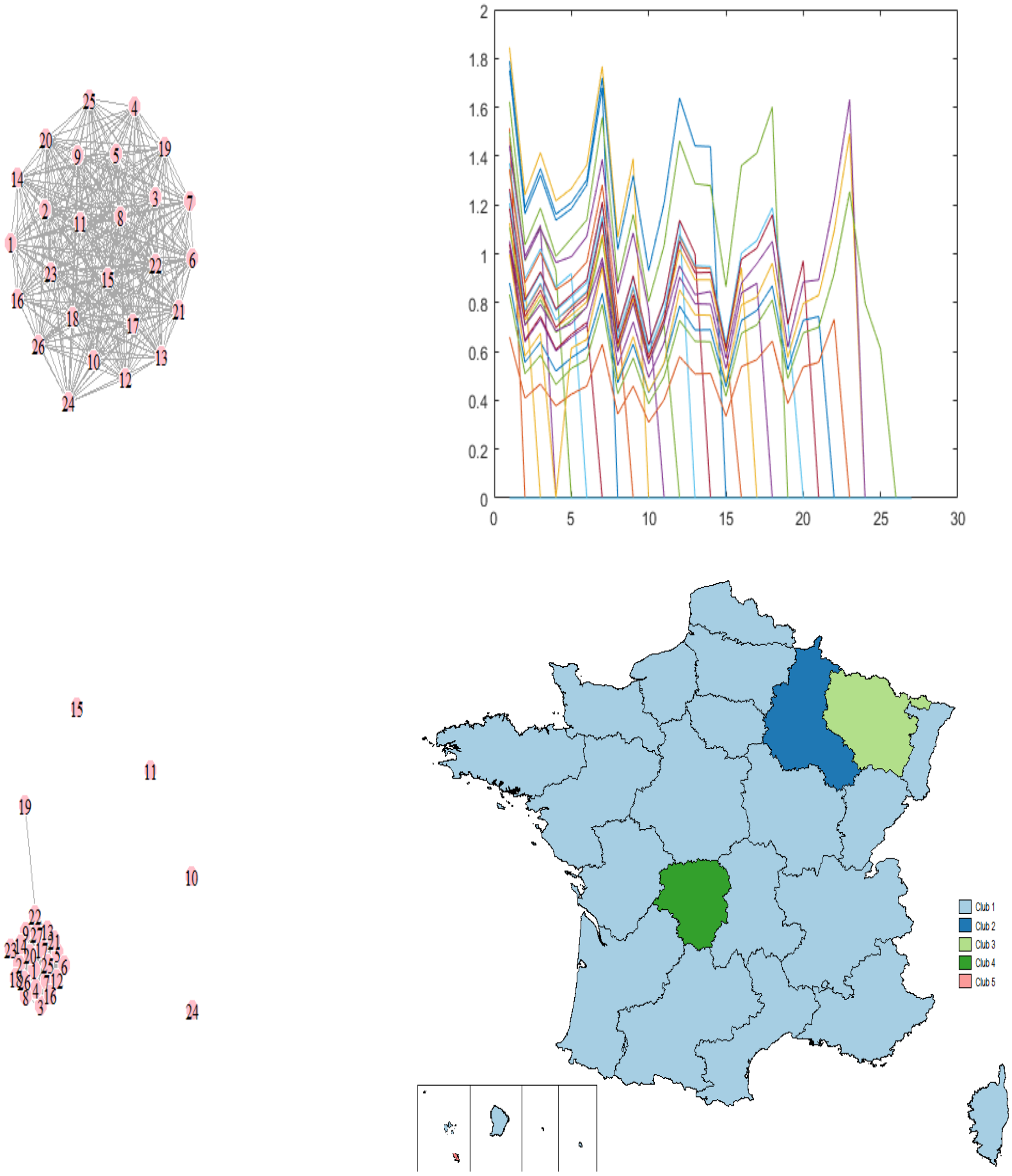

Figure 3: These panels present the NUTS-2 regions of France in several clusters characterized by pairwise convergence tests given by the hypothesis $x_{it} = \beta_0 + \beta x_{jt}$, with i, j denoting different NUTS-2 regions. Left panels present the graphs with the network analysis. Each number represents a NUTS2-region. Top left panel for per-capita income and bottom left panel for population. The top right panel represents the OLS estimates of the cointegration vector for per-capita income for each of the 27 regions. The bottom right panel reports a map of France divided into NUTS2-regions. Each color represents a different cluster as per the graph analysis in the bottom left panel. The pairwise relationships between regions are obtained from testing at 1% significance level the convergence hypothesis in (4) using the test $\widehat{D}_T$ defined in (5).

graph, and; (*iii*) plot the map for each country at NUTS-2 level. We focus on the test of convergence carried out at the 1% significance level. A suitable critical value is 0.3944, as calculated in the Monte-Carlo simulation section. Figure 3 presents the study of clusters for France. The graph on the top left panel contains the network analysis for per-capita income. All of the regions are interconnected with a positive cointegration coefficient. The top right panel reports the value of such parameter for each of the 27 NUTS-2 regions. Each line corresponds to a specific region and represents the OLS estimates of the cointegration coefficient with respect to the remaining regions. The estimates oscillate around a value of one implying that the standard one-to-one relationship may be a bit strong, but the values of per-capita income across regions are close to satisfying that condition. The bottom panels correspond to the study of population. The bottom left panel displays the network analysis explained above associated to the interaction matrix. The bottom right panel presents the corresponding representation as a map. The population of most regions converges as defined in this paper. However, in contrast to the study of per-capita income, the cointegration coefficient varies significantly across pairs of regions. The convergence test and the map imply that regional population grows at the same rate even if the levels of population are quite different. There are just a few regions with very different dynamics. These differences are studied in Figures 6 and 7 discussed below. Table 4 reports the degree and betwenness statistics for the analysis of regional population in France. The degree of each region is in most cases equal to 21 that corresponds to the number of regions that belong to the same cluster. The degree for those regions not belonging to the major convergence club is close to zero. In contrast, the betweenness is close to zero in all cases implying that there are no influencing observations biasing the classification of regions into clusters.

Table 4: Summary statistics of network analysis for population (France).

NUTS-2 region	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Degree	21	21	19	21	21	20	21	21	21	0	0	21	21	21
Betweenness	0.2	0.2	0	0.2	0.2	0.1	0.2	0.2	0.2	0	0	0.2	0.2	0.2

NUTS-2 region	15	16	17	18	19	20	21	22	23	24	25	26	27
Degree	0	21	21	21	1	21	21	21	19	0	21	21	21
Betweenness	0	0.2	0.2	0.2	0	0.2	0.2	0.2	0	0	0.2	0.2	0.2

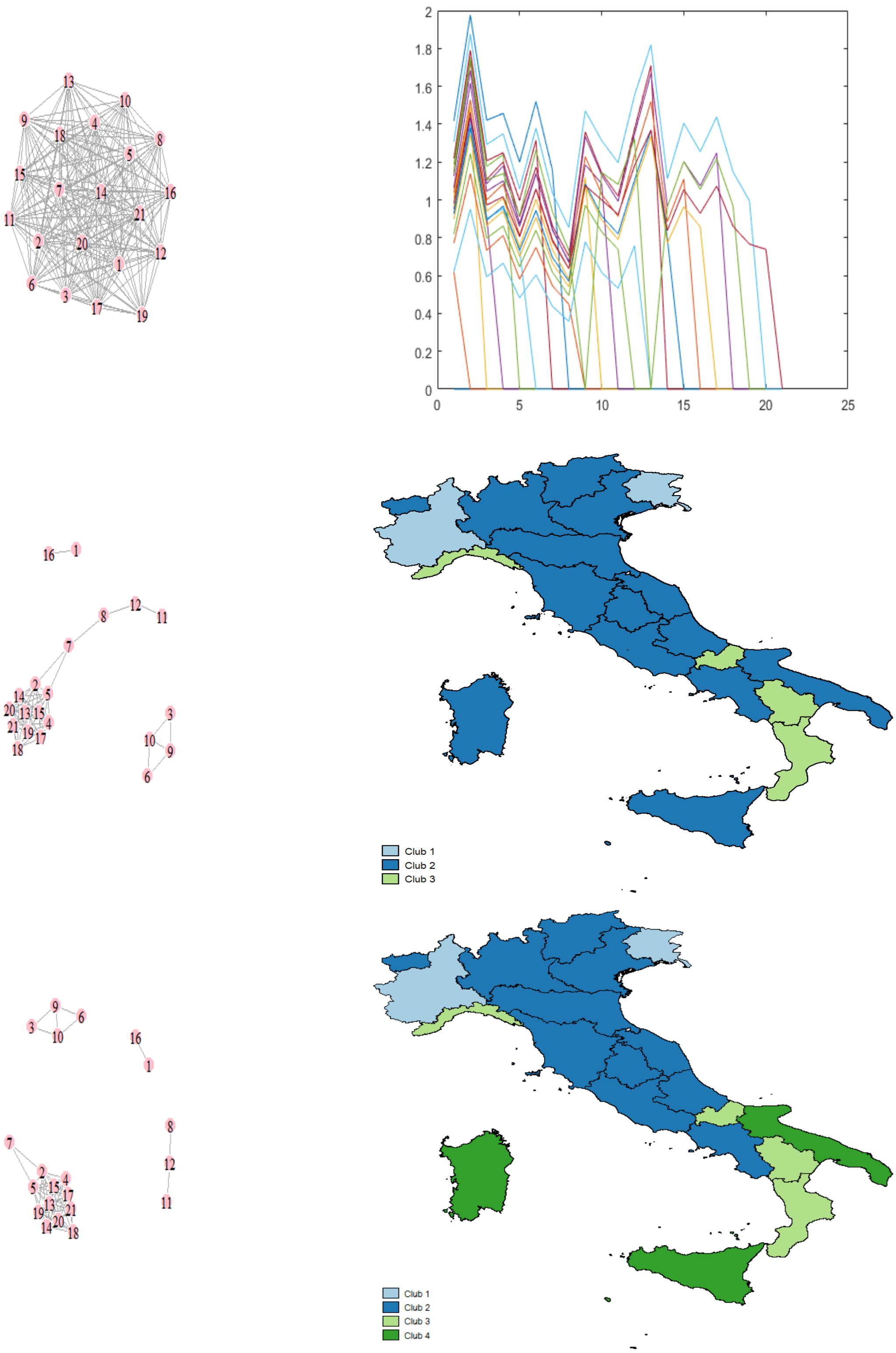

Figure 4: These panels present the NUTS-2 regions of Italy in several clusters characterized by pairwise convergence tests given by the hypothesis $x_{it} = \beta_0 + \beta x_{jt}$, with i, j denoting different NUTS-2 regions. Left panels present the graphs with the network analysis. Each number represents a NUTS2-region. Top left panel for per-capita income; middle and bottom left panels for population. The top right panel represents the OLS estimates of the cointegration vector for per-capita income for each of the 21 regions. The middle and bottom right panels present a map of Italy divided into NUTS2-regions. Each color represents a different cluster as per the graph analysis in the middle and bottom left panels, respectively. Middle panel correspond to the analysis with all connections and bottom panels correspond to the analysis removing the connections from Region 7 (Campania). The pairwise relationships between regions are obtained from testing at 1% significance level the convergence hypothesis in (4) using the test $\widehat{D}_T$ defined in (5).

Figure 4 presents the case of Italy. The top panel corresponds to the analysis of per-capita income. The findings are very similar to those obtained for France. All regions converge in per-capita income and the cointegration coefficients oscillate around a value of one, suggesting that convergence of per-capita income is close to the conventional one-to-one relationship. Population dynamics are studied in the middle and bottom panels. The middle left panel presents the network analysis of the interaction matrix and the middle right panel the corresponding map. Italian NUTS-2 regions are divided into three clusters. Most regions belong to the same cluster implying that the dynamics of population are positively cointegrated for most regions. There are, however, a few southern regions that reject the convergence hypothesis reflecting a different rate of population growth over the last forty years. This is illustrated in more detail in Figure 6 below.

Table 5: Summary statistics of network analysis for population (Italy).

NUTS-2 region	1	2	3	4	5	6	7	8	9	10	11
Degree	1	10	2	10	10	2	3	2	3	3	1
Betweenness	0	**18**	0	0.8	**18**	0	**33**	**24**	0.5	0.5	0
NUTS-2 region	12	13	14	15	16	17	18	19	20	21	
Degree	2	10	10	10	1	10	8	10	10	10	
Betweenness	**13**	0.8	0.8	0.8	0	0.8	0	0.8	0.8	0.8	

Table 6: Summary statistics of reduced network analysis for population (Italy).

NUTS-2region	1	2	3	4	5	6	7	8	9	10	11
Degree	1	10	2	10	10	2	2	1	3	3	1
Betweenness	0	4.5	0	0.4	4.5	0	0	0	0.5	0.5	0
NUTS-2region	12	13	14	15	16	17	18	19	20	21	
Degree	2	10	10	10	1	10	8	10	10	10	
Betweenness	1	0.4	0.4	0.4	0	0.4	0	0.4	0.4	0.4	

To study further the formation of these convergence clubs, we also present in Tables 5 and 6 the degree and betweenness for all the NUTS-2 regions in Italy. Table 5 considers the unrestricted analysis with all regions in the interaction matrix. There are a few regions reporting large values of the betweenness statistic. According to our definition of convergence club in (8), these units spuriously link several clusters and can entail the missclassification of units into multiple clusters. This is the case for region 7 (Campania) that reports a betweenness of 33 and region 8 (Puglia) that reports a betweenness of 24. Most of the remaining regions report values smaller than one. The bottom panels of Figure 4 report the convergence analysis for Italy's regional population but removing the connections from the interaction matrix departing from Campania. In this case, we find four groups of regions converging at different rates. This is also reflected in Table 6 that considers the reduced network. The degree is very similar to the analysis in Table 5, however, the betweenness has been drastically reduced for all regions.

Figure 5 presents the results for the analysis of regional convergence in Spain. The study of per-capita income is not discussed further, as it is analogous to the cases of France and Italy. The bottom panel presents the analysis of population. The Eastern regions exhibit similar population growth rates and are characterized by positive pairwise cointegration relationships. The Western and Northern regions show different patterns. This is further confirmed by the dynamics of population in the bottom panels of Figure 6. Table 7 presents the summary statistics for the degree and betwenness for the analysis of population. The results are consistent with the findings in Figure 5 and clearly reflect the presence of a few differentiated convergence clubs.

Figure 6 reports the dynamics of standardized log population for representative groups of regions. The top left panel analyzes the case of France. The top right panel presents the dynamics of region 19 (Auvergne). Although both panels show an increasing trend in log population over time, the dynamics of Auvergne present significant discrepancies at the beginning of the period with respect to the rest. The middle panels of Figure 6 consider the case of Italy. The left panel presents the dynamics of standardized log population for regions belonging to the largest cluster and the right panel presents the demographic evolution for Liguria, Basilicata and Calabria (in green in the map of Italy). In contrast to the dynamics of the remaining regions, the population of these three regions has steadily declined over the evaluation period. The bottom panels consider the case of Spain. The left panel reports the dynamics of standardized log population for the converging regions whereas the right panel reports the dynamics of region 2 (Principado de Asturias), which are similar to those observed for region 11 (Castilla y León). Interestingly, population in the East of Spain has been steadily growing over the last forty years whereas regions in

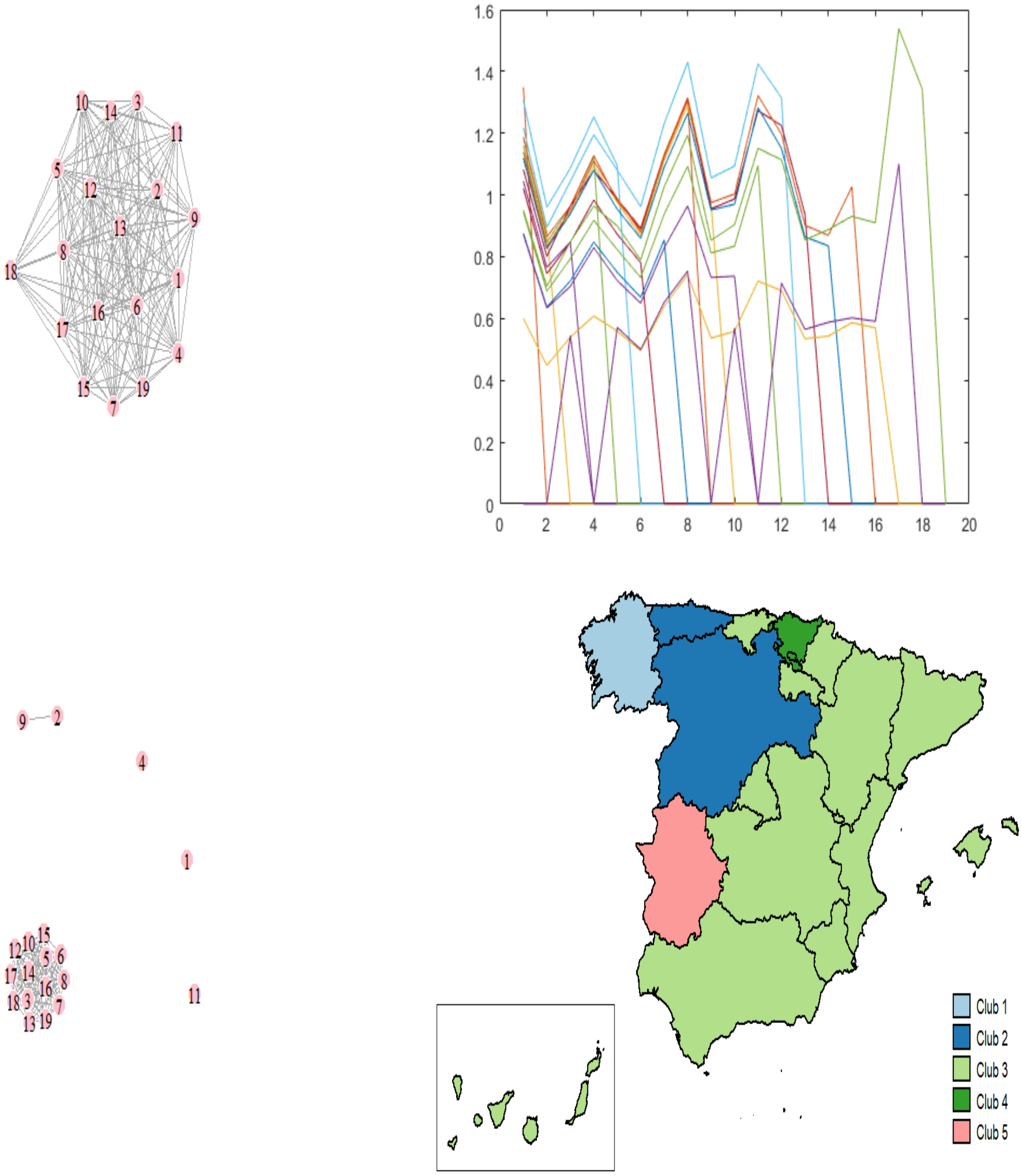

Figure 5: These panels present the NUTS-2 regions of Spain in several clusters characterized by pairwise convergence tests given by the hypothesis $x_{it} = \beta_0 + \beta x_{jt}$, with i, j denoting different NUTS-2 regions. Left panels present the graphs with the network analysis. Each number represents a NUTS2-region. Top left panel for per-capita income and bottom left panel for population. The top right panel represents the OLS estimates of the cointegration vector for per-capita income for each of the 19 regions. The bottom right panel reports a map of Spain divided into NUTS2-regions. Each color represents a different cluster as per the graph analysis in the bottom left panel. The pairwise relationships between regions are obtained from testing at 1% significance level the convergence hypothesis in (4) using the test $\widehat{D}_m$ defined in (5).

the North West have witnessed ups and downs, with a sharp decline in population after 2008-2009.

Table 7: Summary statistics of network analysis for population (Spain).

NUTS-2region	1	2	3	4	5	6	7	8	9	10
Degree	0	1	13	0	13	13	13	13	1	13
Betweenness	0	0	0	0	0	0	0	0	0	0
NUTS-2region	11	12	13	14	15	16	17	18	19	
Degree	0	13	13	13	13	13	13	13	13	
Betweenness	0	0	0	0	0	0	0	0	0	

Figure 7 presents a similar analysis but focusing on standardized log per-capita income. Top panel corresponds to France, middle panel is for Italy and bottom panel is for Spain. The dynamics are very homogeneous across NUTS-2 regions for the three countries confirming graphically the regional convergence in per-capita income once we allow for heterogeneity in the cointegration coefficients.

6 Conclusion

This paper extends the standard definition of time series convergence in per-capita income, given by the parity condition $(1, -1)$, between pairs of unit root variables, by allowing for a flexible cointegration relationship $(1, -\beta)$, with $\beta > 0$. The paper also introduces a novel test to statistically assess this hypothesis. The main novelty of our method is that it is a test of positive cointegration against the composite null hypothesis given by (i) two independent unit roots or (ii) negative cointegration between the variables. Another important feature of the proposed test of convergence is that it is robust to general forms of weak dependence in the transitory component of the unit root processes and produces more accurate empirical size, in finite samples, than standard residual-based tests of cointegration. Its implementation does not require knowledge or estimation of the cointegration coefficient.

As a byproduct of the above analysis, we propose a methodology to detect convergence clubs. This procedure is based on centrality measures of network dependence given by the degree and betweenness. The degree indicates the number of individuals in a cluster whereas the betweenness of a node indicates if such observation spuriously belongs to more than one cluster. We propose a simple procedure to remove those spurious relationships

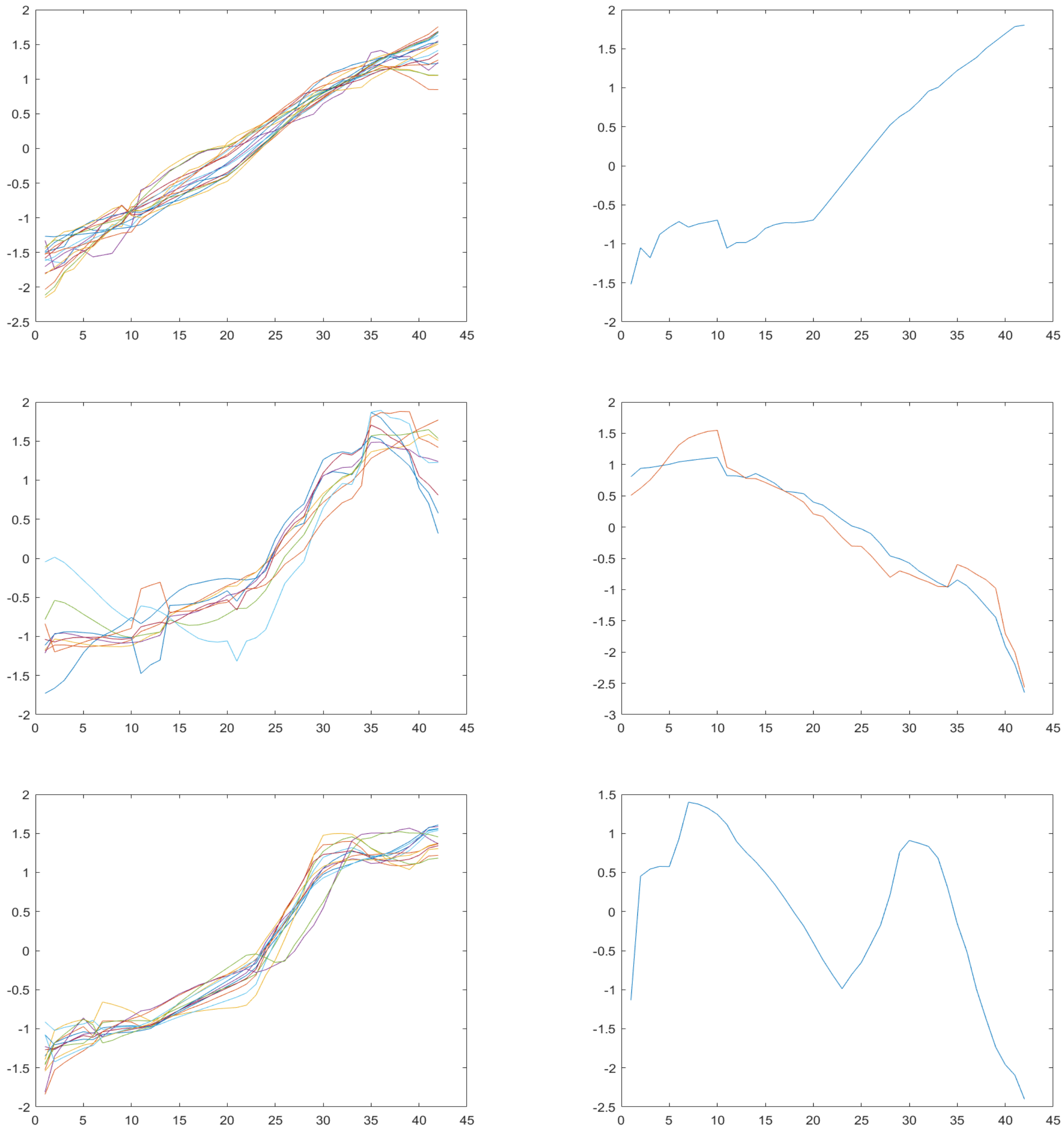

Figure 6: Dynamics of standardized log population over 2002-2020 for representative regions in each country. Left panels report the dynamics of regions with convergent patterns and right panels the dynamics of regions that do not converge. Top panels study the case of France, middle panels focus on Italy and bottom panels report the case of Spain. The top right panel presents the dynamics of Region 19 (Auvergne). Middle right panel presents the dynamics of population for Liguria, Basilicata and Calabria. Bottom right panel reports the dynamics of Region 2 (Principado de Asturias).

and obtain self-contained clusters of regions. These clusters are interpreted as convergence clubs.

The proposed methods are illustrated with European regional data on per-capita income and population from France, Italy and Spain at the NUTS-2 level. The outputs of the

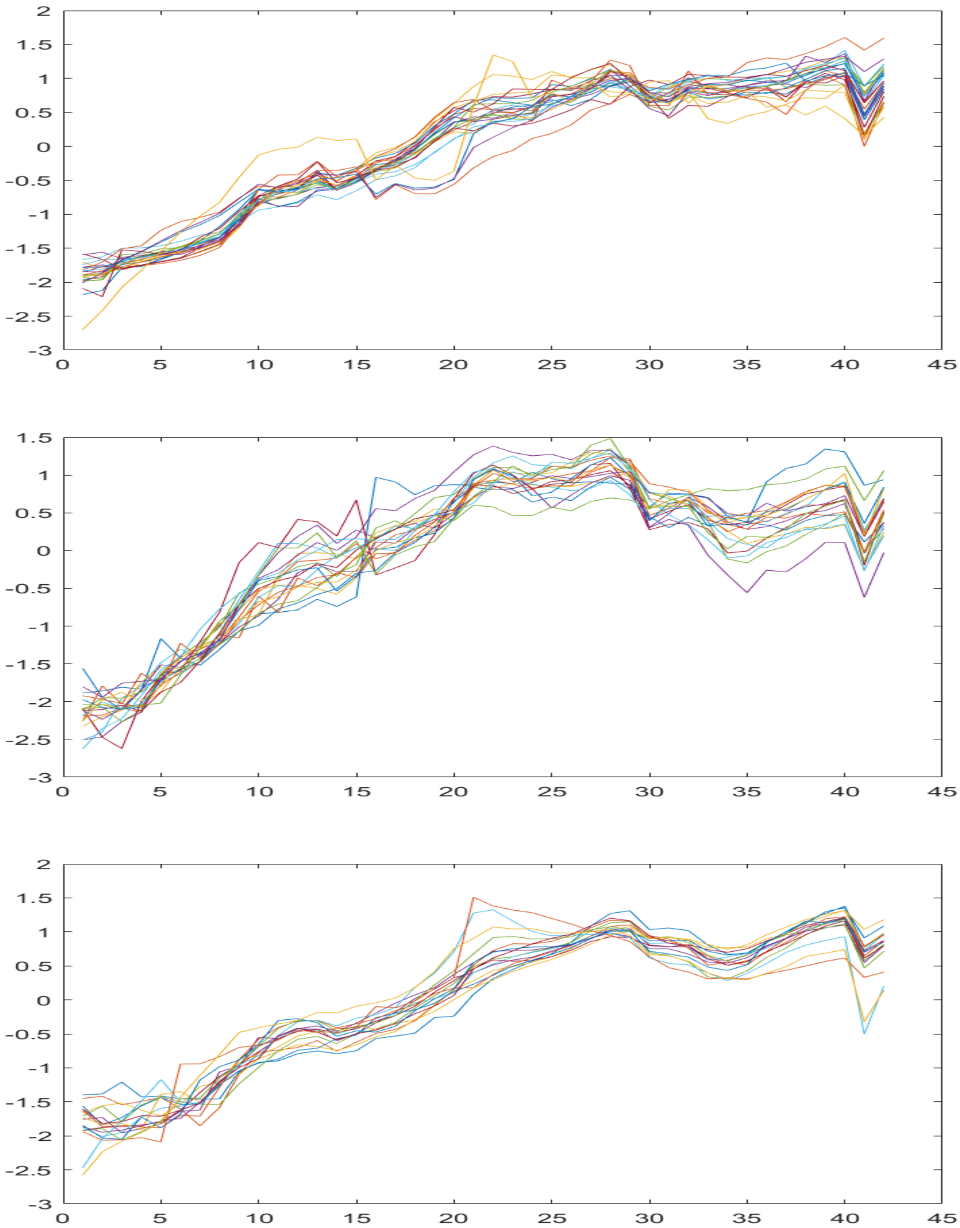

Figure 7: Dynamics of standardized log per-capita income over 2002-2020 for representative groups of regions in each country. Top panel reports the case of France; middle panel corresponds to Italy and bottom panel presents the case of Spain.

pairwise convergence tests are used as inputs of an interaction matrix that uncovers the presence of clusters of regions exhibiting convergence. The empirical results show similar findings across countries. Regional per-capita income converges for most regions. The analysis of long-run population dynamics is more complex. Our analysis shows two types of regional clusters. Population grows at similar rates in most regions for the three countries, however, we have identified regions with declining population patterns, in particular, for

Italy and Spain. Both types of regions are clustered geographically suggesting that the phenomenon of depopulation affects regions with specific geographical, environmental and climatological characteristics.

References

Abramovitz, Moses. 1986. "Catching up, forging ahead, and falling behind". *Journal of Economic History* 46 (2): 385–406. doi:10.2307/2122171.

Azariadis, Costas, and Allan Drazen. 1990. "Threshold externalities in economic development". *Quarterly Journal of Economics* 105 (2): 501–526. doi:10.2307/2937797.

Barro, Robert J. 1991. "Economic growth in a cross section of countries". *Quarterly Journal of Economics* 106 (2): 407–443. doi:10.2307/2937943.

Baumol, William J. 1986. "Productivity growth, convergence, and welfare: What the long-run data show". *American Economic Review* 76 (5): 1072–1085. doi:10.2307/1816469.

Bernard, Andrew B., and Steven N. Durlauf. 1995. "Convergence in international output". *Journal of Applied Econometrics* 10 (2): 97–108. doi:10.1002/jae.3950100202.

— . 1996. "Interpreting tests of the convergence hypothesis". *Journal of Econometrics* 71 (1): 161–173. doi:10.1016/0304-4076(94)01699-2.

Beylunioğlu, Fuat C., M. Ege Yazgan, and Thanasis Stengos. 2020. "Detecting convergence clubs". *Macroeconomic Dynamics* 24 (3): 629–669. doi:10.1017/S1365100518000391.

Bianchi, Marco. 1997. "Testing for convergence: Evidence from non-parametric multimodality tests". *Journal of Applied Econometrics* 12 (4): 393–409. doi:10.2307/2284960.

Bliss, Christopher. 1999. "Galton's fallacy and economic convergence". *Oxford Economic Papers* 51 (1): 4–14. doi:10.1093/oep/51.1.4.

Cannon, Edmund S., and Nigel W. Duck. 2000. "Galton's fallacy and economic convergence". *Oxford Economic Papers* 52 (2): 415–419. doi:10.1093/oep/52.2.415.

Carvalho, Vasco M., and Andrew C. Harvey. 2005. "Growth, cycles and convergence in U.S. regional time series". *International Journal of Forecasting* 21 (4): 667–686. doi:10.1016/j.ijforecast.2005.04.017.

Durlauf, Steven N., and Paul A. Johnson. 1995. "Multiple regimes and cross-country growth behaviour". *Journal of Applied Econometrics* 10 (4): 365–384. doi:10.1002/jae.3950100404.

Durlauf, Steven N., Paul A. Johnson, and Jonathan R. W. Temple. 2005. "Growth econometrics". In *Handbook of Economic Growth*, ed. by Philippe Aghion and Steven Durlauf, vol. 1A, 555–677. Amsterdam, North-Holland: Elsevier. ISBN: 9780444520418. doi:10.1016/S1574-0684(05)01008-7.

— . 2009. "The econometrics of convergence". In *Palgrave Handbook of Econometrics*, ed. by Terence C. Mills and Kerry Patterson, 2:1087–1118. London, United Kingdom: Palgrave Macmillan. ISBN: 9780230244405. doi:10.1057/9780230244405_23.

Durlauf, Steven N., David Romer, and Christopher A. Sims. 1989. "Output persistence, economic structure, and the choice of stabilization policy". *Brookings Papers on Economic Activity* 1989 (2): 69–136. doi:10.2307/2534462.

Engle, Robert F., and Clive W. J. Granger. 1987. "Co-integration and error correction: Representation, estimation, and testing". *Econometrica* 55 (2): 251–276. doi:10.2307/1913236.

Friedman, Milton. 1992. "Do old fallacies ever die?" *Journal of Economic Literature* 30 (4): 2129–2132. doi:10.2307/2727976.

Galor, Oded. 1996. "Convergence? Inferences from theoretical models". *Economic Journal* 106 (437): 1056–1069. doi:10.2307/2235378.

Garcia-Hiernaux, Alfredo, and David E. Guerrero. 2021. "Price convergence: Representation and testing". *Economic Modelling* 104:105641. doi:10.1016/j.econmod.2021.105641.

Granger, Clive W. J. 1981. "Some properties of time series data and their use in econometric model specification". *Journal of Econometrics* 16 (1): 121–130. doi:10.1016/0304-4076(81)90079-8.

Harvey, Andrew C., and Vasco M. Carvalho. 2005. "Convergence in the trends and cycles of Euro-zone income". *Journal of Applied Econometrics* 20 (2): 275–289. doi:10.1002/jae.820.

Hobijn, Bart, and Philip H. Franses. 2000. "Asymptotically perfect and relative convergence of productivity". *Journal of Applied Econometrics* 15 (1): 59–81. doi:10.2307/2678580.

Johansen, Søren. 1988. "Statistical analysis of cointegration vectors". *Journal of Economic Dynamics and Control* 12 (2): 231–254. doi:10.1016/0165-1889(88)90041-3.

— . 1991. "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models". *Econometrica* 59 (6): 1551–1580. doi:10.2307/2938278.

Johansen, Søren, and Katarina Juselius. 1990. "Maximum likelihood estimation and inference on cointegration - with applications to the demand for money". *Oxford Bulletin of Economics and Statistics* 52 (2): 169–210. doi:10.1111/j.1468-0084.1990.mp52002003.x.

Kong, Jianning, Peter C. B. Phillips, and Donggyu Sul. 2019. "Weak σ-convergence: Theory and applications". *Journal of Econometrics* 209 (2): 185–207. doi:10.1016/j.jeconom.2018.12.022.

Mankiw, N. Gregory, David Romer, and David N. Weil. 1992. "A contribution to the empirics of economic growth". *Quarterly Journal of Economics* 107 (2): 407–437. doi:10.2307/2118477.

Ng, Serena, and Pierre Perron. 2001. "Lag length selection and the construction of unit root tests with good size and power". *Econometrica* 69 (2): 1519–1554. doi:10.1111/1468-0262.00256.

Olmo, Jose. 2022. *A simple cointegration test robust to serial correlation*. SSRN Working Paper No 4203205. doi:10.2139/ssrn.4203205.

Perron, Pierre, and Serena Ng. 1996. "Useful modifications to some unit root tests with dependent errors and their local asymptotic properties". *The Review of Economic Studies* 63 (2): 435–463. doi:10.2307/2297890.

Pesaran, M. Hashem. 2007. "A pair-wise approach to testing for output and growth convergence". *Journal of Econometrics* 138 (1): 312–355. doi:10.1016/j.jeconom.2006.05.024.

Phillips, Peter C. B. 1986. "Understanding spurious regressions in econometrics". *Journal of Econometrics* 33 (2): 311–340. doi:10.1016/0304-4076(86)90001-1.

— . 1987. "Time series regression with a unit root". *Econometrica* 55 (2): 277–301. doi:10.2307/1913237.

Phillips, Peter C. B., and Donggyu Sul. 2007. "Transition modeling and econometric convergence tests". *Econometrica* 75 (6): 1771–1855. doi:10.1111/j.1468-0262.2007.00811.x.

— . 2009. "Economic transition and growth". *Journal of Applied Econometrics* 24 (7): 1153–1185. doi:10.1002/jae.1080.

Phillips, Peter C.B., and Pierre Perron. 1988. "Testing for a unit root in time series regression". *Biometrika* 75 (2): 335–346. doi:10.2307/2336182.

Quah, D. 1997. "Empirics for growth and distribution: polarization, stratification and convergence club". *Journal of Economic Growth* 2:27–59.

Quah, Danny T. 1993a. "Empirical cross-section dynamics in economic growth". *European Economic Review* 37 (2): 426–434. doi:10.1016/0014-2921(93)90031-5.

— . 1993b. "Galton's fallacy and tests of the convergence hypothesis". *Scandinavian Journal of Economics* 95 (4): 427–443. doi:10.2307/3440905.

— . 1996a. "Convergence empirics across economies with (some) capital mobility". *Journal of Economic Growth* 1 (1): 95–124. doi:10.2307/40215883.

— . 1996b. "Twin peaks: Growth and convergence in models of distribution dynamics". *Economic Journal* 106 (437): 1045–1055. doi:10.2307/2235377.

Said, Said E., and David A. Dickey. 1984. "Testing for unit roots in autoregressive-moving average models of unknown order". *Biometrika* 71:599–607. doi:10.2307/2336570.

Schwert, G. William. 1989. "Tests for unit roots: A Monte Carlo investigation". *Journal of Business and Economic Statistics* 7 (2): 147–159. doi:10.2307/1391432.

Silva Lopes, Artur. 2016. "A simple proposal to improve the power of income convergence tests". *Economics Letters* 138 (C): 92–95. doi:10.1016/j.econlet.2015.11.

Young, Andrew T., Matthew J. Higgins, and Daniel Levy. 2008. "Sigma convergence versus beta convergence: Evidence from U.S. county-level data". *Journal of Money, Credit and Banking* 40 (5): 1083–1093. doi:10.2307/25096293.

Mathematical Appendix

The following auxiliary lemmas are useful for proving the main results in the paper. Additional technical details for the general test of cointegration under the presence of different deterministic trends can be found in Olmo (2022).

Lemma A.1: Let $x_{it} = \alpha_{x_i} + \pi_i t + \theta_i f_t^o + u_{it}$, with $f_t^o = \sum_{s=1}^{t} v_s$ and u_{1t} as in (2). Then, $\widehat{\pi}_i - \pi_i = O_p(T^{-1/2})$, with $\widehat{\pi}_i$ the corresponding OLS estimator. The asymptotic distribution is

$$\sqrt{T}\left(\widehat{\pi}_i - \pi_i\right) \xrightarrow{d} 12\theta_i \sigma_v \int_0^1 \left(r - \frac{1}{2}\right) W_i(r) dr, \text{ as } T \to \infty, \tag{10}$$

with σ_v^2 the long-run variance of the innovation v_t of the common factor f_t^o and $W_i(r)$ a Brownian motion.

Proof. Let x_{it} be defined as in process (1) for $i = 1, \ldots, n$. The OLS estimator of the drift component satisfies that

$$\widehat{\pi}_i = \frac{\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right)\left(x_{it} - \overline{x}_i\right)}{\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right)^2} = \pi_i + \frac{\frac{1}{T^3}\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right) x_{it}^0}{\frac{1}{T^3}\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right)^2} = \pi_i + 12\frac{1}{T^3}\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right) x_{it}^0 + o_p(1),$$

with $x_{it}^0 = \theta_i(f_t^o - \overline{f^o}) + u_{it} - \overline{u}_i$ a demeaned unit root process without drift, given that $\frac{1}{T^3}\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right)^2 \to 1/12$, as $T \to \infty$. Note also that $\frac{1}{T^{5/2}}\sum_{t=1}^{T} t f_t^o \xrightarrow{d} \sigma_v \int_0^1 r W_i(r) dr$ and $\frac{1}{T^{5/2}}\sum_{t=1}^{T} t \overline{f^o} \xrightarrow{d} \frac{1}{2}\sigma_v \int_0^1 W_i(r) dr$. Furthermore, under weak dependence of u_{it}, it also holds that $\frac{1}{T^{3/2}}\sum_{t=1}^{T} t u_{it} = O_p(1)$ and $\frac{1}{T^{1/2}}\sum_{t=1}^{T} u_{it} = O_p(1)$. Then,

$$\begin{aligned}
\widehat{\pi}_i - \pi_i &= \frac{12\theta_i}{T^3}\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right)\left(f_t^o - \overline{f^o}\right) + \frac{12}{T^3}\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right)\left(u_{it} - \overline{u}_i\right) + o_p(1) \\
&= O_p(T^{-1/2}) + O_p(T^{-3/2}) + o_p(1).
\end{aligned}$$

It also follows that $\frac{1}{T^{5/2}}\sum_{t=1}^{T} t(f_t^o - \overline{f^o}) \xrightarrow{d} \sigma_v\left(\int_0^1 rW_i(r)dr - \frac{1}{2}\int_0^1 W_i(r)dr\right)$. Then,

$$\sqrt{T}\left(\widehat{\pi}_i - \pi_i\right) \xrightarrow{d} 12\theta_i\sigma_v\left(\int_0^1 rW_i(r)dr - \frac{1}{2}\int_0^1 W_i(r)dr\right), \tag{11}$$

as $T \to \infty$. ■

Lemma A.2: If the processes x_{it} and x_{jt} are cointegrated as in (3),

$$\widetilde{x}_{it} - \beta_{ij}\widetilde{x}_{jt} = O_p(T^{-3/2})\left(t - \frac{T+1}{2}\right) + \varepsilon_{0t}, \tag{12}$$

with $\beta_{ij} = \theta_i/\theta_j$; $\widetilde{x}_{it}$ and $\widetilde{x}_{jt}$ are the sample detrended processes associated to the unit roots x_{it} and x_{jt}, and $\varepsilon_{0t} = u_{it} - \overline{u}_i - \beta_{ij}(u_{jt} - \overline{u}_j)$.

Proof. Let $\widetilde{x}_{it} = x_{it} - \overline{x}_i - \widehat{\pi}_i\left(t - \frac{T+1}{2}\right)$ be the sample detrended process for $i = 1, \ldots, n$. If the processes x_{it} and x_{jt} are cointegrated as in (3), we have

$$x_{it} - \overline{x}_i = \pi_i\left(t - \frac{T+1}{2}\right) + \theta_i(f_t^o - \overline{f^o}) + u_{it} - \overline{u}_i,$$

such that the detrended processes satisfy that

$$\widetilde{x}_{it} = \theta_i(f_t^o - \overline{f^o}) - (\widehat{\pi}_i - \pi_i)\left(t - \frac{T+1}{2}\right) + u_{it} - \overline{u}_i. \tag{13}$$

Thus,

$$\widetilde{x}_{it} - \beta_{ij}\widetilde{x}_{jt} = \left[\beta_{ij}(\widehat{\pi}_j - \pi_j) - (\widehat{\pi}_i - \pi_i)\right]\left(t - \frac{T+1}{2}\right) + \varepsilon_{0t}, \tag{14}$$

with $\varepsilon_{0t} = u_{it} - \overline{u}_i - \beta_{ij}(u_{jt} - \overline{u}_j)$.

Using the algebra in the proof of Lemma A.1, it follows that $\widehat{\pi}_i - \pi_i = \theta_i \frac{\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right)(f_t^o - \overline{f^o})}{\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right)^2} + O_p(T^{-3/2})$. Then,

$$\widetilde{x}_{it} - \beta_{ij}\widetilde{x}_{jt} = O_p(T^{-3/2})\left(t - \frac{T+1}{2}\right) + \varepsilon_{0t}. \tag{15}$$

■

Lemma A.3: Let x_{it} and x_{jt} be as in (3). Then,

$$\frac{\widehat{\sigma}_{\widetilde{x}_i}}{\widehat{\sigma}_{\widetilde{x}_j}} = |\beta_{ij}| + O_p(T^{-1/2}) \quad \text{as } T \to \infty. \tag{16}$$

Proof. The proof of this result builds on Lemma A.2. Thus, using expression (15), we obtain

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{it}^2 &= \beta_{ij}^2\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{jt}^2 + O_p(T^{-3})\frac{1}{T}\sum_{t=1}^{T}\left(t-\frac{T+1}{2}\right)^2 + O_p(T^{-3/2})\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{jt}\left(t-\frac{T+1}{2}\right)\\ &+ 2\beta_{ij}\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{jt}(u_{it}-\overline{u}_i) - 2\beta_{ij}^2\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{jt}(u_{jt}-\overline{u}_j) + \frac{1}{T}\sum_{t=1}^{T}(u_{it}-\overline{u}_i)^2 + \beta_{ij}^2\frac{1}{T}\sum_{t=1}^{T}(u_{jt}-\overline{u}_j)^2\\ &+ O_p(T^{-3/2})\frac{1}{T}\sum_{t=1}^{T}(u_{it}-\overline{u}_i)\left(t-\frac{T+1}{2}\right) + O_p(T^{-3/2})\frac{1}{T}\sum_{t=1}^{T}(u_{jt}-\overline{u}_j)\left(t-\frac{T+1}{2}\right).\end{aligned}$$

The time series u_{it} and u_{jt} are two stationary processes exhibiting weak dependence as described in (3). The following asymptotic conditions hold in this case: $\frac{1}{T}\sum_{t=1}^{T}\left(t-\frac{T+1}{2}\right)^2 = O(T^2)$ and $\frac{1}{T}\sum_{t=1}^{T}(u_{it}-\overline{u}_i)\left(t-\frac{T+1}{2}\right) = O_p(T^{1/2})$. Furthermore, we also require the asymptotic conditions $\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{jt}\left(t-\frac{T+1}{2}\right) = O_p(T^{3/2})$ and $\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{jt}(u_{it}-\overline{u}_i) = O_p(1)$. The validity of the latter two conditions is shown below:

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{jt}\left(t-\frac{T+1}{2}\right) &= \frac{\theta_j}{T}\sum_{t=1}^{T}(f_t^o-\overline{f^o})\left(t-\frac{T+1}{2}\right) - (\widehat{\pi}_j-\pi_j)\frac{1}{T}\sum_{t=1}^{T}\left(t-\frac{T+1}{2}\right)^2\\ &+ \frac{1}{T}\sum_{t=1}^{T}(u_{jt}-\overline{u}_j)\left(t-\frac{T+1}{2}\right)\\ &= O_p(T^{3/2}) + O_p(T^{-1/2})O(T^2) + O_p(T^{1/2}) = O_p(T^{3/2}).\end{aligned} \tag{17}$$

Similarly,

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{jt}(u_{jt}-\overline{u}_j) &= \frac{\theta_j}{T}\sum_{t=1}^{T}(f_t^o-\overline{f^o})(u_{jt}-\overline{u}_j) - (\widehat{\pi}_j-\pi_j)\frac{1}{T}\sum_{t=1}^{T}(u_{jt}-\overline{u}_j)\left(t-\frac{T+1}{2}\right)\\ &+ \frac{1}{T}\sum_{t=1}^{T}(u_{jt}-\overline{u}_j)^2 = O_p(1) + O_p(T^{-1/2})O_p(T^{1/2}) + O_p(1) = O_p(1),\end{aligned} \tag{18}$$

if the long run variance of u_{jt} is finite as imposed in model (3). Thus,

$$\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{it}^2 = \beta_{ij}^2\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{jt}^2 + O_p(1), \tag{19}$$

given that $\frac{1}{T}\sum_{t=1}^{T}(u_{it}-\overline{u}_i)^2 = O_p(1)$, under weak dependence.

We now prove that $\frac{1}{T^2}\sum_{t=1}^{T}\widetilde{x}_{jt}^2 = O_p(1)$. The dynamics of this process are derived in (13).

$$\begin{aligned}\frac{1}{T^2}\sum_{t=1}^{T}\widetilde{x}_{jt}^2 &= \frac{\theta_j^2}{T^2}\sum_{t=1}^{T}(f_t^o-\overline{f^o})^2 + (\widehat{\pi}_j-\pi_j)^2\frac{1}{T^2}\sum_{t=1}^{T}\left(t-\frac{T+1}{2}\right)^2 + \frac{1}{T^2}\sum_{t=1}^{T}(u_{jt}-\overline{u}_j)^2\\ &\quad -2(\widehat{\pi}_j-\pi_j)\theta_j\frac{1}{T^2}\sum_{t=1}^{T}(f_t^o-\overline{f^o})\left(t-\frac{T+1}{2}\right) + 2\theta_j\frac{1}{T^2}\sum_{t=1}^{T}(f_t^o-\overline{f^o})(u_{1t}-\overline{u}_1)\\ &\quad +\frac{1}{T^2}\sum_{t=1}^{T}(u_{1t}-\overline{u}_1)\left(t-\frac{T+1}{2}\right).\end{aligned}$$

Using the limiting results for unit root processes in Phillips (1986, 1987) and Lemma A.1, we obtain

$$\frac{1}{T^2}\sum_{t=1}^{T}\widetilde{x}_{jt}^2 = O_p(1)+O_p(T^{-1})O(T)+o_p(1)+O_p(T^{-1/2})O_p(T^{1/2})+o_p(1)+o_p(1) = O_p(1).$$

Dividing by $\widehat{\sigma}^2_{\widetilde{x}_j}$ in expression (19), we obtain

$$\frac{\widehat{\sigma}^2_{\widetilde{x}_i}}{\widehat{\sigma}^2_{\widetilde{x}_j}} = \beta_{ij}^2 + O_p(T^{-1}) = \beta_{ij}^2 + o_p(1) \quad \text{as } T\to\infty. \tag{20}$$

■

Proof of Proposition 1. Let $y_{it} = \frac{\widetilde{x}_{it}}{\widehat{\sigma}_{\widetilde{x}_i}}$ and note from Lemma A.2 that $\widetilde{x}_{it} = \beta_{ij}\widetilde{x}_{jt} + O_p(T^{-3/2})\left(t-\frac{T+1}{2}\right)+\varepsilon_{0t}$, with $\varepsilon_{0t} = u_{it}-\overline{u}_i-\beta_{ij}(u_{jt}-\overline{u}_j)$. Combining this result with (20), we obtain

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}y_{it}y_{jt} &= sgn(\beta_{ij})\frac{1}{T}\sum_{t=1}^{T}\frac{\widetilde{x}_{jt}^2}{\widehat{\sigma}^2_{\widetilde{x}_j}} + O_p(T^{-3/2})\frac{1}{T}\sum_{t=1}^{T}\frac{\widetilde{x}_{jt}}{\widehat{\sigma}^2_{\widetilde{x}_j}}\left(t-\frac{T+1}{2}\right)+\frac{1}{T}\sum_{t=1}^{T}\frac{\widetilde{x}_{jt}}{\widehat{\sigma}^2_{\widetilde{x}_j}}\varepsilon_{0t}\\ &= sgn(\beta_{ij}) + O_p(T^{-3/2})\frac{1}{T}\sum_{t=1}^{T}\frac{\widetilde{x}_{jt}}{\widehat{\sigma}^2_{\widetilde{x}_j}}\left(t-\frac{T+1}{2}\right)+O_p(T^{-1})\\ &= sgn(\beta_{ij}) + O_p(T^{-3/2})O_p(T^{1/2})+O_p(T^{-1}) = sgn(\beta_{ij})+O_p(T^{-1}),\end{aligned}$$

noting that $\frac{1}{T^{5/2}}\sum_{t=1}^{T}\widetilde{x}_{jt}\left(t-\frac{T+1}{2}\right) = O_p(1)$, $\frac{\widehat{\sigma}^2_{\widetilde{x}_j}}{T} = O_p(1)$ and $\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{jt}(u_{jt}-\overline{u}_j) = O_p(1)$ as shown in the proof of Lemma A.3.

Note also that the test statistic $\widehat{D}_{ij,T}$ can be expressed as $2\left(1 - \frac{1}{T}\sum_{t=1}^{T} y_{it}y_{jt}\right)$ given that $\frac{1}{T}\sum_{t=1}^{T} y_{rt}^2 = 1$, for $r = i, j$, due to the standardization. Therefore, $\widehat{D}_T = 2\left(1 - sgn(\beta_{ij}) + O_p(T^{-1})\right)$, implying that for $\beta_{ij} > 0$, $sgn(\beta_{ij}) = 1$ such that $\widehat{D}_T = O_p(T^{-1})$. ■

Proof of Corollary 1. The proof of this result is immediate from the proof of Proposition 1 by noting that $sgn(\beta_{ij}) = -1$ for $\beta_{ij} < 0$, such that $\widehat{D}_{ij,T} = 4 + O_p(T^{-1})$ as $T \to \infty$. ■

Proof of Proposition 2. Let the time series x_{it} and x_{jt} be generated from two independent unit root processes without drift (f_{it}^o and f_{jt}^o) such that $x_{it} = \alpha_{x_i} + \pi_i t + \theta_i f_{it}^o + u_{it}$ and $x_{jt} = \alpha_{x_j} + \pi_j t + \theta_j f_{jt}^o + u_{jt}$, with u_{it} and u_{jt} exhibiting weak dependence as described in (3). The detrended versions of these processes are $\widetilde{x}_{it} = x_{it}^0 - (\widehat{\pi}_i - \pi_i)\left(t - \frac{T+1}{2}\right)$, with $x_{it}^0 = \theta_i(f_{it}^o - \overline{f_i^o}) + u_{it} - \overline{u}_i$, and $\widetilde{x}_{jt} = x_{jt}^0 - (\widehat{\pi}_j - \pi_j)\left(t - \frac{T+1}{2}\right)$, with $x_{jt}^0 = \theta_j(f_{jt}^o - \overline{f_j^o}) + u_{jt} - \overline{u}_j$. Then, the asymptotic distribution of the test statistic $\widehat{D}_{ij,T}$ is determined by the asymptotic distribution of $\frac{1}{T}\sum_{t=1}^{T} y_{it}y_{jt}$, with $y_{it} = \frac{\widetilde{x}_{it}}{\widehat{\sigma}_{\widetilde{x}_i}}$. Thus,

$$\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{it}\widetilde{x}_{jt} = \frac{1}{T}\sum_{t=1}^{T}x_{it}^0x_{jt}^0 - (\widehat{\pi}_i - \pi_i)\frac{1}{T}\sum_{t=1}^{T}x_{jt}^0\left(t - \frac{T+1}{2}\right) - (\widehat{\pi}_j - \pi_j)\frac{1}{T}\sum_{t=1}^{T}x_{it}^0\left(t - \frac{T+1}{2}\right)$$
$$+ (\widehat{\pi}_i - \pi_i)(\widehat{\pi}_j - \pi_j)\frac{1}{T}\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right)^2.$$

Note also from Lemma A.1 that $\widehat{\pi}_r - \pi_r = \frac{\frac{1}{T}\sum_{t=1}^{T} x_{rt}^0\left(t - \frac{T+1}{2}\right)}{\frac{1}{T}\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right)^2}$, for $r = i, j$, such that

$$\frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{it}\widetilde{x}_{jt} = \frac{1}{T}\sum_{t=1}^{T}x_{it}^0x_{jt}^0 - \frac{T^2}{12}(\widehat{\pi}_i - \pi_i)(\widehat{\pi}_j - \pi_j) + o_p(1), \tag{21}$$

given that $\frac{1}{T^3}\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right)^2 \to \frac{1}{12}$ as $T \to \infty$. Similarly, we study the sample variance terms. Thus,

$$\widehat{\sigma}_{\widetilde{x}_i}^2 = \frac{1}{T}\sum_{t=1}^{T}\widetilde{x}_{it}^2 = \frac{1}{T}\sum_{t=1}^{T}(x_{it}^0)^2 - (\widehat{\pi}_i - \pi_i)^2\frac{1}{T}\sum_{t=1}^{T}\left(t - \frac{T+1}{2}\right)^2,$$

such that $\widehat{\sigma}^2_{\widetilde{x}_i} = \widehat{\sigma}^2_{x^0_i} - \frac{T^2}{12}(\widehat{\pi}_i - \pi_i)^2 + o_p(1)$. Then, simple algebra shows that $\widehat{\sigma}_{\widetilde{x}_i} = \widehat{\sigma}_{x^0_i}\left(1 - \widehat{A}^2_{\pi_i}\right)^{1/2} + o_p(1)$, with $\widehat{A}_{\pi_i} = \frac{1}{\sqrt{12}}\left(\frac{\sqrt{T}(\widehat{\pi}_i - \pi_i)}{\widehat{\sigma}_{x^0_i}/\sqrt{T}}\right)$, and $\widehat{\sigma}_{\widetilde{x}_i}\widehat{\sigma}_{\widetilde{x}_j} = \widehat{\sigma}_{x^0_i}\widehat{\sigma}_{x^0_j}\left(1 - \widehat{A}^2_{\pi_i}\right)^{1/2}\left(1 - \widehat{A}^2_{\pi_j}\right)^{1/2} + o_p(1)$.

Using this algebra, the test statistic can be expressed as

$$\frac{1}{T}\sum_{t=1}^{T} y_{it}y_{jt} = \frac{\widehat{\sigma}_{x^0_i}\widehat{\sigma}_{x^0_j}}{\widehat{\sigma}_{\widetilde{x}_i}\widehat{\sigma}_{\widetilde{x}_j}}\left(\frac{1}{T}\sum_{t=1}^{T} y^0_{it}y^0_{jt} - \widehat{A}_{\pi_i}\widehat{A}_{\pi_j}\right),$$

with $y^0_{rt} = \frac{x^0_{rt}}{\widehat{\sigma}_{x^0_r}}$, for $r = i, j$, the standardized version of the unit root processes without drift defined above.

Olmo (2022) shows that $\frac{1}{T}\sum_{t=1}^{T} y^0_{1t}y^0_{2t}$ can be interpreted as the OLS estimator of the linear regression of y^0_{it} on y^0_{jt}, given that both unit root processes have unit variance, by construction. Phillips (1986) derives the asymptotic distribution of this quantity in spurious regressions between two independent unit root processes. Thus, we obtain

$$\frac{1}{T}\sum_{t=1}^{T} y^0_{it}y^0_{jt} \xrightarrow{d} Z_{ij} = \frac{\int_0^1 W_i(r)W_j(r)dr - \int_0^1 W_i(r)dr\int_0^1 W_j(\tau)d\tau}{\left[\int_0^1 W_i(r)^2dr - \left(\int_0^1 W_i(r)dr\right)^2\right]^{1/2}\left[\int_0^1 W_j(r)^2dr - \left(\int_0^1 W_j(r)dr\right)^2\right]^{1/2}}.$$

Additionally, $\frac{\widehat{\sigma}_{x^0_i}\widehat{\sigma}_{x^0_j}}{\widehat{\sigma}_{\widetilde{x}_i}\widehat{\sigma}_{\widetilde{x}_j}} = (1 - \widehat{A}^2_{\pi_i})^{-1/2}(1 - \widehat{A}^2_{\pi_j})^{-1/2} + o_p(1)$, with $\widehat{A}_{\pi_i} = \frac{1}{\sqrt{12}}\left(\frac{\sqrt{T}(\widehat{\pi}_i - \pi_i)}{\widehat{\sigma}_{x^0_i}/\sqrt{T}}\right)$, such that

$$\widehat{A}_{\pi_i} \xrightarrow{d} A_{\pi_i} = \frac{\sqrt{12}\int_0^1\left(r - \frac{1}{2}\right)W_i(r)dr}{\left[\int_0^1 W_i(r)^2dr - \left(\int_0^1 W_i(r)dr\right)^2\right]^{1/2}} \tag{22}$$

using the asymptotic result in Lemma A.1 and the results in Phillips (1986) that imply that $\frac{1}{T}\widehat{\sigma}^2_{x^0_i} \xrightarrow{d} \theta^2_i\sigma^2_{v_1}\left[\int_0^1 W_i(r)^2dr - \left(\int_0^1 W_i(r)dr\right)^2\right]$. Therefore, $\frac{1}{T}\sum_{t=1}^{T} y_{it}y_{jt} \xrightarrow{d} B_\pi(Z_{ij} - A_{\pi_i}A_{\pi_j})$, with $B_\pi = (1 - A_{\pi_i})^{-1/2}(1 - A_{\pi_j})^{-1/2}$, and

$$\widehat{D}_{ij,T} \xrightarrow{d} 2\left(1 + B_\pi A_{\pi_i}A_{\pi_j} - B_\pi Z_{ij}\right).$$

■