

Club Convergence And Cluster Persistence: A Novel Approach To Multivariate Convergence Analysis

*Convergenza in club e persistenza nei gruppi:
una proposta per l'analisi multivariata della convergenza*

Raffaele Scuderi

Università degli studi di Palermo, Dipartimento di Contabilità Nazionale
ed Analisi dei Processi Sociali, email: scuderi@economia.unipa.it

Riassunto: Partendo dai concetti di crescita come *fenomeno multidimensionale* e di *club convergence* il lavoro verifica la presenza di convergenza tra i paesi aderenti al Partenariato Euro-Mediterraneo. Sono proposte due misure sintetiche dell'evoluzione delle coalizioni nel tempo, basate su algoritmi classici di analisi dei gruppi. La prima, che analizza il grado di persistenza dei paesi nella stessa coalizione, viene integrata dalla seconda, che propone un indice di polarizzazione basato sull'assiomatizzazione di Esteban e Ray (1994); ambedue gli approcci sono applicati ai paesi aderenti al Partenariato Euro-Mediterraneo. I risultati mostrano la presenza di persistenza nei gruppi ed una tendenza all'aumento della polarizzazione.

Keywords: Club convergence, multivariate analysis, persistency matrix, polarization.

1. Introduction

According to neoclassical growth theory, *convergence hypothesis* postulates that economies tend to the same level of per capita income in the long term. Consistently with the original Solow model, debate on statistical methods to test the evolution of disparities over time has been focused on univariate measures, typically per capita GDP. Statistical approaches have anyway gradually abandoned the original theoretical framework based on the production function. Nevertheless applied literature has put in evidence features of growth dynamics that theory lacked to highlight like *club convergence*, that is a “local” convergence to a group of countries rather than to the whole world. Some authors have also pointed out the *multidimensional* nature of growth, which is in contrast with earlier empirical approaches examining the sole per capita GDP as synthesis of the growth process. Both the assumptions of growth as multidimensional phenomenon and club convergence are the starting point for the convergence tests we present in this paper.

2. Measuring club convergence from a batch of indicators

Unlike recent contributions on living standards which perform multivariate analysis of convergence by simply examining single indicators, we investigate about the joint effect of a batch of indicators in making economies more or less “similar” over time. We

define the *coalition structure* $\mathcal{C} = \{C(1), \dots, C(r), \dots, C(|\mathcal{C}|)\}$ as a nonempty partition of a finite set N of N economies (called *grand coalition*), where C_r is a *group*, or *coalition*, or *cluster*, and $|\mathcal{C}| = N$ is a positive integer (Quah, 1999). Countries form coalitions according to a *similarity criterion*, so that groups are characterized by a high degree of heterogeneity, each owning members with a high degree of homogeneity. We also denote with $(\pi, \mathcal{C})_t = \{\pi_{1,t}, \dots, \pi_{|\mathcal{C}|,t}; C_{1,t}, \dots, C_{|\mathcal{C}|,t}\}$ a distribution, where $\pi_{r,t}$ is the population of $C_{r,t}$ at time $t \in [1, T]$, and $\sum_{i=1}^{|\mathcal{C}|} \pi_i = \pi = N$ is the total population associated with C . *Convergence* happens if $\lim_{t \rightarrow \infty} (\pi, \mathcal{C})_t = \{\pi; \mathcal{A}\}$; *club convergence* happens instead when $\lim_{t \rightarrow \infty} (\pi, \mathcal{C})_t = (\pi_1, \dots, \pi_s; C_1, \dots, C_s)$, $s = N$. The following methodologies aim then to observe the evolution of each $(\pi, \mathcal{C})_t$ by looking at cluster composition and membership. We determine groups through a hierarchical agglomerative method, which allows to form coalitions with no initial specification of the number of clusters. In particular we choose *Ward's* minimum variance algorithm applied to *Euclidean distances*, which minimizes the deviance between centroids of the pairs of objects aggregated at each step. A “classical” *z-score* standardisation can be employed at each t in order to avoid problems related to scale and variability of indicators to aggregate. At the same time we choose the same level of the cutting line of the dendrogram over the whole time interval in order to compare coalitions over time, and to see how they evolve from the same fixed “point of view”. Nevertheless $z d_{ij,k,t+1}^2 / z d_{ij,k,t}^2 = (d_{ij,k,t+1}^2 / d_{ij,k,t}^2) (s_{kt}^2 / s_{k,t+1}^2)$: dynamics of univariate Euclidean distance between two units, $d_{ij,k}^2$, is not the same as the correspondent z-scored distribution, $z d_{ij,k}^2$ (multivariate case is a mere extension of the univariate one). In order to avoid this bias, instead of z-score we propose to employ the transformation

$$\zeta_{n\bar{k}, \bar{t}+l} = \left[(x_{n\bar{k}, \bar{t}+l} - m_{\bar{k}\bar{t}}) / s_{\bar{k}\bar{t}} \right], \quad \bar{t} = 1, l = 0, \dots, T-1, \quad (1)$$

where $x_{n\bar{k}, \bar{t}+l}$ is the value the indicator $\bar{k}$ assumes over country n at time $\bar{t}+l$. Following Notarstefano and Scuderi (2003), information coming from each $(\pi, \mathcal{C})_t$ is reported into a squared $N \times N$ matrix $\mathbf{A}_t = \{a_{ij}\}_t$, where $a_{ij} = 1$, $i \neq j$, if i and j are in the same coalition at t , and $a_{ij} = 1$, $i = j$, if i forms a cluster by itself at t . The matrix $\mathbf{CA}\{c_{ij}\} = (1/T) \sum_t \mathbf{A}_t$ ($0 \leq c_{ij} \leq 1$) is called *persistence matrix*. A synthetic index to evaluate the *degree of presence* of two average units in the same cluster in t is

$$\psi_t = \sum_{i=1}^N \sum_{j \neq i=1}^N a_{ij,t} / N(N-1) \in [0, 1], \quad (2)$$

which is equal to 0 if each country forms a cluster by itself, and 1 if units belong to the same group.

Table 1: Cluster persistence matrix

	DEN	SWE	GER	BEL	FIN	NET	UKI	AUS	FRA	ISR	ITA	IRE	GRE	POR	SPA	CYP	HUN	POL	EGY	ALG	SYR	TUN	TUR	JOR
DEN	1.00	0.13	0.06	0.06	0.06	0.06	0.06																	
SWE	1.00		0.13	0.06	0.06	0.06	0.06																	
GER	0.13	0.13		0.94	0.94	0.94	0.94	0.87	0.87	0.74	0.58	0.26												
BEL	0.06	0.06	0.94		1.00	1.00	1.00	0.94	0.94	0.81	0.65	0.32												
FIN	0.06	0.06	0.94	1.00		1.00	1.00	0.94	0.94	0.81	0.65	0.32												
NET	0.06	0.06	0.94	1.00	1.00		1.00	0.94	0.94	0.81	0.65	0.32												
UKI	0.06	0.06	0.94	1.00	1.00	1.00		0.94	0.94	0.81	0.65	0.32												
AUS			0.87	0.94	0.94	0.94	0.94		1.00	0.87	0.71	0.39												
FRA			0.87	0.94	0.94	0.94	0.94	1.00		0.87	0.71	0.39												
ISR			0.74	0.81	0.81	0.81	0.81	0.87	0.87		0.71	0.39	0.13	0.13	0.13	0.13								
ITA			0.58	0.65	0.65	0.65	0.65	0.71	0.71	0.71		0.48	0.29	0.26	0.26	0.23								
IRE			0.26	0.32	0.32	0.32	0.32	0.39	0.39	0.39	0.48		0.61	0.61	0.61	0.55								
GRE										0.13	0.29	0.61		1.00	1.00	0.94								
POR										0.13	0.26	0.61	1.00		1.00	0.94								
SPA										0.13	0.26	0.61	1.00	1.00		0.94								
CYP										0.13	0.23	0.55	0.94	0.94	0.94			0.06	0.06					
HUN																0.06		1.00						
POL																0.06	1.00							
EGY																			1.00	1.00	1.00	1.00	1.00	1.00
ALG																			1.00	1.00	1.00	1.00	1.00	1.00
SYR																			1.00	1.00	1.00	1.00	1.00	1.00
TUN																			1.00	1.00	1.00	1.00	1.00	1.00
TUR																			1.00	1.00	1.00	1.00	1.00	1.00
JOR																			1.00	1.00	1.00	1.00	1.00	1.00

Empirical analysis concerns 24 of the 35 countries belonging to the enlarged European Union and the Mediterranean non European economies agreeing to the Euro-Mediterranean Partnership ⁽¹⁾, over 1970-2000. The three chosen indicators proxy each of the three aspects considered by Human Development Indicator ⁽²⁾, in order to detect the evolution of differences from a broader point of view than the sole GDP. The fixed level of the cutting line of dendrograms has been chosen empirically, and it is equal to 3.7. Countries substantially show persistence in the same coalition (figure 1, left). From persistency matrix (table 1), the most evident feature is the separation between European Countries and MPC, and the emerging of persistent clusters within Europe. Club convergence and stratification dynamics are also confirmed by stochastic kernel applied over each simple indicator, not reported here.

Useful indications to integrate persistency analysis emerge from Esteban and Ray's (1994 - ER, from now on) theory about income *polarization*. A distribution is said to be *polarized* if it tends to form clusters with very similar members, which are simultaneously "dissimilar" from other groups' units. An asymmetric ⁽³⁾ polarization measure between income classes formulated by the authors is

$$P^*(\pi, \mathbf{y}) = K \sum_{i=1}^n \sum_{j=1}^n \pi_i^{1+\alpha} \pi_j y_i^\beta |y_i - y_j|, \text{ where } K \text{ is an integer; } \mathbf{y} \in \mathbb{I}^n \text{ is a finite}$$

distribution of incomes, $y_i \neq y_j \quad \forall (i, j)$; $|y_i - y_j|$ is the absolute distance between two income classes; α and β weight the alienation of two income classes (specifically, $\alpha = 1.6$ is said to be an optimal value for symmetric measures, but no indications are given about the asymmetric ones; our empirical analysis considers this value).

Following ER's framework we propose the following *coalitions polarization index*:

¹([□]) Considered countries are Algeria, Austria, Belgium, Cyprus, Denmark, Egypt, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Jordan, Netherlands, Poland, Portugal, Spain, Sweden, Syria, Tunisia, Turkey, United Kingdom,

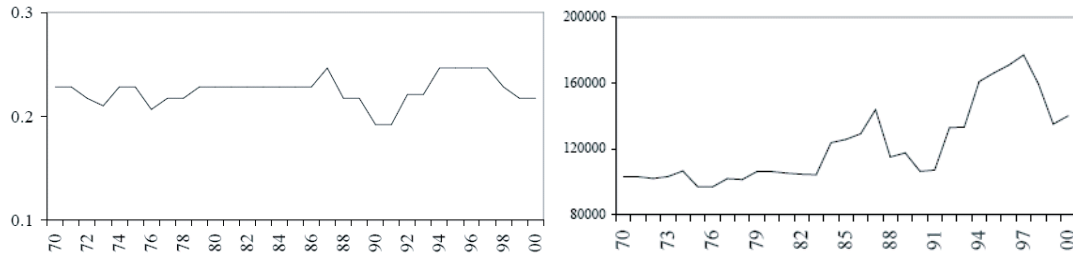
²([□]) Chosen indicators are *Life expectancy at birth* and *GDP per capita (constant 2000 US\$)* from The World Bank dataset; *Average schooling years in the total population* from Barro and Lee dataset. Data are available under request.

³([□]) The term "asymmetric" refers to the different perception by each of two groups of the distance that separates them. For instance, rich does not want to be poor, but the contrary is not always true.

$$P(\mathcal{C}) = \sum_{i=1}^{|\mathcal{C}|} \sum_{j \neq i=1}^{|\mathcal{C}|} \pi_i^{1+\alpha} \pi_j \delta(C_i C_j). \quad (3)$$

Specifically, $\delta(C_i C_j) = \lambda(C_i C_j) - \lambda(C_i)$; $\lambda(C_i)$ is the linkage distance of coalition C_i ; $\lambda(C_i C_j)$ is the linkage distance at which coalitions C_i and C_j merge into a bigger one. As regards to the original formulation, we put $y_i^\beta |y_i - y_j| \equiv \delta(C_i C_j)$: since $\delta(C_i C_j) \neq \delta(C_j C_i)$, the measure of dissimilarity $\delta(C_i C_j)$ aims to weight the different distance at which two groups are far from the level at which they merge. The index then increases as units tend to form bigger coalitions, and as coalitions get further from each other over time. The measure implements persistence analysis in the sense that it indicates whether groups tend to be “closer” over time, or if instead they tend to form “poles”. Polarization index (figure 1, right) substantially shows stationarity in the first years of the interval. Since 1984 we also note an increasing trend and a wider range of its cyclic behavior.

Figure 1: Ψ_t index (left) and $P(\mathcal{C})$ index (right)



Information coming from both cluster persistence algorithm and polarization index point out a substantial trend of units to persist within the same clusters, which also increase their distance over time. Specifically, while Ψ_t 's stationarity indicates that numerousness of clusters has not varied significantly, polarization index suggests that distance between clusters has tendentially widened. Results reject then the convergence hypothesis within countries agreeing to the Euro-Mediterranean Partnership, and the tendency to form quite well-defined coalitions.

References

- Esteban J.-M., Ray D. (1994) On the Measurement of Polarization, *Econometrica*, 62, 819-851.
- Notarstefano G., Scuderi R. (2003) Persistenza e appartenenza sfocata: un confronto su unità territoriali, *Rivista Italiana di Economia, Demografia e Statistica*, 3/4, 103-111.
- Pacini B., Pellegrini G. (2002) Metodi non parametrici multivariati: un'applicazione al caso della crescita, *Statistica*, 3, 379-395.
- Quah D. T. (1997) Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs, *Journal of Economic Growth*, 2, 27-59.
- Quah D. T. (1999) Ideas determining convergence clubs, *mimeo*, from the website “<http://econ.lse.ac.uk/staff/dquah/p/9907ccb.pdf>”.