



UNU-CRIS e-Working Papers

W-2004/2

Indicators of Real Economic Convergence.

A Primer

José Villaverde Castro*

*Professor of Economics, University of Cantabria (Spain). E-mail: villavej@unican.es
Paper prepared for the UNU-CRIS Virtual Workshop on Indicators of Regional Integration, April-October 2003. For more information on the Workshop, please contact: pdelombaerde@cris.unu.edu.

Indicators of Real Economic Convergence. A Primer.

José Villaverde Castro

1. Introduction

Whether poor economies tend to converge towards rich ones or else to diverge over time is an issue that has attracted the attention of policy-makers and academics alike for some decades. Economic convergence or divergence is a topic of considerable interest and debate, not only for validating or otherwise the two leading and competing growth models (the neoclassical and the endogenous growth approaches) but also for its policy-oriented implications. Generally speaking, the presence of convergence is considered as a valid test in favour of the neoclassical growth model as opposed to the endogenous models that predict divergence in most cases.

The convergence issue has been revived in the last two decades thanks to the seminal works by Abramovitz (1986), Baumol (1986) and Barro and Sala-i-Martin (1991,1992), to name just a few. Convergence analysis has even gained momentum by the development of endogenous growth models that allow a major role for economic policy. Unfortunately, the available empirical evidence does not provide unambiguous support for either of the two aforementioned growth frameworks, although it does point to the existence of conditional convergence.

At the regional level, the issue of economic convergence is also much debated. A look at growth performances of member countries of integration schemes in Europe, Asia and Latin America over a relatively long period of time does not show unambiguous empirical support for the convergence hypothesis neither (De Lombaerde, 2002). This is in sharp contrast with the fact that, like in the case of the EU, for the majority of commentators, economists and politicians economic convergence is an expected (if not necessary) outcome of regional economic

integration processes (Meeusen and Villaverde, 2002b). The empirical assessments of the effects of the elimination of barriers to flows of goods, services and production factors on growth (and its spatial distribution) and the debate on cohesion policies in the EU have considerably contributed to the development of the conceptual framework and the methodological toolbox for studying economic convergence.¹

According to the Oxford Dictionary, convergence is a tendency to become similar or identical. Although this definition is somehow illuminating, when economists talk about convergence they usually refer to what is known either as nominal or real convergence². Nominal convergence relates to the process of nominal variables approaching stability levels. It is mainly concerned with such things as, for instance, convergence in interest rates and inflation rates, but also stability in the exchange rate, the government deficit and/or the government debt, both as a percentage of GDP. Broadly speaking, real convergence is understood to mean the approximation in the levels of economic welfare or development across economies. Thus, real convergence relates mainly to the time performance of variables such as per capita income, productivity, unemployment rate and so on³. In its simplest way, real convergence implies a long-run tendency towards the equalization of per capita income levels across economies (Abramovitz, 1986). Nevertheless, the relevant point is that real convergence is a multifaceted concept that, as it has been posed by Quah (1997) *‘reflects on –among other things– polarization, income distribution and inequality’* (p. 27).

The focus of this paper is with real convergence⁴. That is with the long-term process of reducing inequalities across economies. One difficulty is that because

¹ For recent overviews, see for example Meeusen and Villaverde (2002a), Funck and Pizzati (2003), Barry and Begg (2003) and the other papers in the Special Issue on EMU and Cohesión of the Journal of Common Market Studies (vol. 41, no. 5, 2003).

² The relationship between nominal and real convergence has also been the subject of considerable debate and attention. For a summary, see Viñals (1994).

³ For simplicity of notation and exposition we will always refer to per capita income convergence, although it should be evident that the same principles carry over to other relevant economic variables.

⁴ From now on we will always refer to real convergence.

there are different concepts of convergence, there are also different ways to measure it. The two most popular approaches, according to the terminology coined by Barro and Sala-i-Martin (1991, 1992), are the so-called s -convergence and b -convergence. Albeit different, these two concepts are not only complementary (in that they respond to different questions), but, as Sala-i-Martin (1996) has shown, they also are closely related: some type of b -convergence is a necessary but not sufficient condition for s -convergence.

This paper will not only explore s -convergence and b -convergence approaches, but will also examine other less standard concepts like convergence as reduction of inequality, stochastic (or time-series) convergence, dynamics distribution convergence and spatial analysis convergence⁵.

The paper is organized as follows. Section 2 reviews convergence from the point of view of the inequality literature; this includes the s -convergence approach. Section 3 portrays a simple description of the well-known b -convergence approach. Section 4 contains the main traits of the time-series approach to convergence. Section 5 offers a general view of the spatial analysis convergence. In all these four sections, the objective is how best to measure real economic convergence. Finally, section 6 gives a summary and presents the conclusions.

2. Convergence as a reduction of inequality

The simplest concept of economic convergence refers to the reduction of per capita income inequality across a sample of economies (countries, regions, states, provinces, ...). In order to measure it, a whole array of inequality indicators has been proposed, the three most popular being some summary measure of dispersion, the Gini index and the Theil index.

⁵ A straightforward application of some of these approaches can be seen in Villaverde (2003), Villaverde and Sánchez-Robles (2002) and Villaverde and Maza (2003).

In particular, it is said that there exists s -convergence if the dispersion of the income per capita in a cross section of economies diminishes over time. In other words, if $s_{t+T} < s_t$, where s is the indicator of dispersion, there exist s -convergence. To test for s -convergence either the variance, the standard deviation or the coefficient of variation⁶ are conventionally used. By far the most frequently employed summary statistics for measuring dispersion are the variance and the standard deviation. However, these two indicators are unsatisfactory descriptive measures of dispersion in that their value is related to the units of measurement. Instead, the coefficient of variation is independent of the units of measurement, and this is why it is used here as an indicator of s -convergence. As is well known, the conventional coefficient of variation, CV_1 , is defined as

$$CV_1 = \frac{\sqrt{V_1}}{\bar{Y}} \quad (1)$$

where V_1 stands for the variance $[V_1 = \frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^2]$, Y_i the per capita income of the i -th economy, \bar{Y} the mean per capita income and N the size of the population sample. CV_1 is a non-weighted measure of inequality as it does not incorporate the population size of each economy. This index can be transformed in a new weighted one (CV_2) by simply employing a weighted variance V_2 given by the expression $V_2 = \frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^2 (p_i - p)$, where p_i refers to the population's share of the i -th economy ($p=1$). Thus

$$CV_2 = \frac{\sqrt{V_2}}{\bar{Y}} \quad (2)$$

⁶ For an interesting review of the different conclusions that may be obtained when using different indicators of dispersion see, for instance, Dalgaard and Vastrup (2001).

A very popular inequality indicator is the Gini coefficient. This is derived from the Lorenz curve (Figure 1), a cumulative frequency curve which plots the cumulative share of population of the economies in the sample on the X-axis and the cumulative share of total income of these same economies on the Y-axis. In both cases the economies are ranked according to their per capita income from bottom to top. The Gini index (G) corresponds to twice the area between the Lorenz curve and the 45° line. Its general expression is given by

$$G = \frac{1}{2\bar{Y}} \sum_{i=1}^N \sum_{j=1}^N p_i p_j |Y_i - Y_j| \quad (3)$$

where all symbols have already been defined. The Gini coefficient ranges from 0 (complete equality) to 1 (complete inequality).

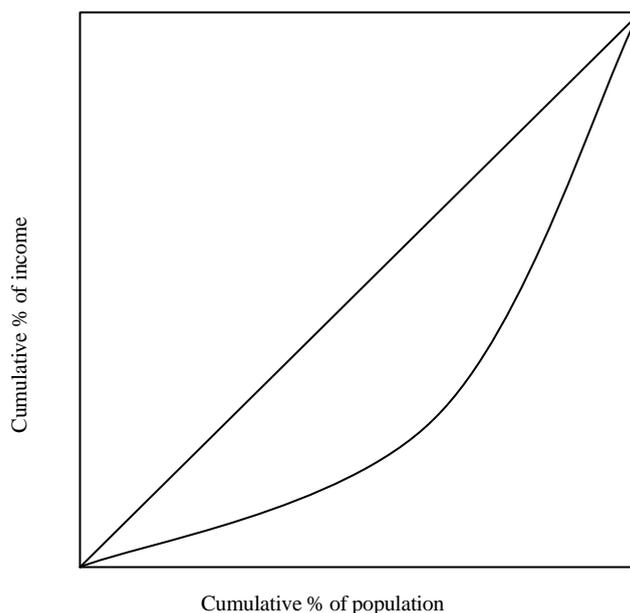


Figure 1.- Lorenz curve

The Theil inequality index, on the other hand, is derived from the notion of entropy in information theory. The $T(I)$ version is given by the expression

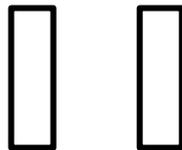
$$T(1) = \frac{1}{N} \sum_1^N \left(\frac{Y_i}{\bar{Y}} \right) \log \left(\frac{Y_i}{\bar{Y}} \right) \quad (4)$$

Among other interesting properties⁷, the Theil index $T(I)$ is additively decomposable⁸, meaning that –among other possibilities- it can be decomposed in a way such that

$$T(1) = \sum_{k=1}^K s_k T(1)_k + \sum_{k=1}^K s_k \log \left(\frac{Y_k}{\bar{Y}} \right) \quad (5)$$

where the first term describes inequality within each of the K population groups in the sample and the second term measures inequality between these groups; s_k is the income share of group k in total income.

Finally, it is interesting to consider the notion of polarization. Although this notion is closely related to that of inequality, it is different from it. According to Esteban and Ray (1994) the concept of polarisation emerges because “*the axioms of inequality measurement (...) fail to adequately distinguish between “convergence” to the global mean and “clustering” around local means*” (p. 821). Polarisation –a concept used to compare the homogeneity of a group with the overall heterogeneity of the sample population (across groups)- refers to the formation of clusters around local poles. The difference between inequality and polarisation can be easily applied to the case of per capita income distribution. For example, assume that this distribution is uniform over income levels 1 to 6 (Figure 2). Now, consider a transformation that causes the income of all



⁷ The $T(I)$ index, as a member of the generalised entropy family, satisfies the axioms of symmetry (or anonymity), population replication, mean independence (homogeneity) and Dalton-Pigou principle of transfers. In order to normalize $T(I)$ between 0 and 1, it suffices to divide $T(I)$ by $\log(N)$.

⁸ Although it is less commonly used, the Gini coefficient can also be decomposed in three terms: a within group term, a between group term and an interaction or residual term. For a recent reference, see Dickey (2001)

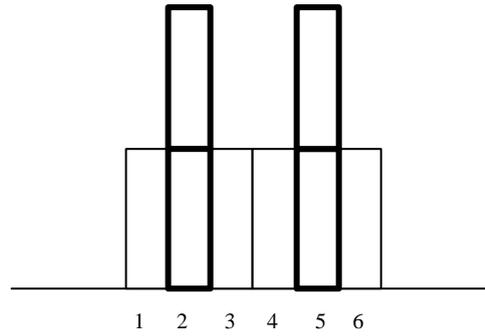


Figure 2.- Per capita income distribution. Polarisation versus inequality

the economies with an income level between 1 and 3 to converge to 2 and the income of all the economies with an income level between 4 and 6 to collapse to 5. Although the Theil index will show a decline of inequality, polarisation (clustering around two poles, the poor economies at an income level of 2 and the rich economies at income level 5) will increase. A simple index of polarization (PI) is given by

$$PI = \sum_{i=1}^N \sum_{j=1}^N p_i^{1+a} p_j |Y_i - Y_j| \quad (6)$$

where $p_{i(j)}$ stands for the weight of population of economy $i(j)$ in the sample and a is an index between 1 and 1.6 that measures the polarization sensitivity. The smaller the sensitivity to polarisation, the closer the notion of polarisation is to that of inequality.

3. Cross-section convergence (*b*-convergence)

The concept of *b*-convergence implies that poor economies grow faster than rich economies. Accordingly, a statistical negative relationship between the growth rate of per capita income and its initial level is expected to be found in a cross-section analysis of these economies. Neoclassical growth theory (Solow, 1956), assuming that technology exhibits diminishing returns to physical and human

capital, predicts **b**-convergence⁹: this is because the further an economy is away from its steady-state level of capital per capita, the faster is the growth of capital and income levels. The neoclassical model not only predicts convergence of each economy to its steady-state level but also that the speed of convergence is inversely related to the distance from its steady-state position or long-run equilibrium.

b-convergence may be absolute or conditional. The difference is related to the existence of key parameters in the economies that compose the sample: technology, savings rate, population growth rate and depreciation rate. If these parameters are identical for all the economies, neoclassical growth theory predicts absolute (or unconditional) **b**-convergence. This is often referred to as catching up theory.

If, on the contrary, the aforementioned parameters differ across economies, neoclassical growth theory predicts conditional **b**-convergence. In the first case, all the economies tend to have the same per capita income in the steady-state whereas in the second, each economy tends to its own steady-state or long-run equilibrium¹⁰. When some economies do not converge to the same steady-state but certain groups of countries converge to a particular steady-state, it is said that they form a convergence club. Within each club, economies converge to each other, but convergence does not happen across different clubs. Chatterji (1993) and Chatterji and Dewhurst (1996) have proposed a simple methodology to estimate the existence of convergence clubs (see below).

⁹ For a critical view of **b**-convergence as being considered either as a useful test of the neoclassical growth theory or a robust measure of convergence (or divergence), see Cheshire and Carbonaro (1995). Endogenous growth models (Romer, 1986 and Rebelo, 1991) consider that income convergence is just but one possible outcome. The main point of these models is that, although individual inputs can be subject to decreasing returns, the positive externalities associated to some of these inputs (technology, human capital, ...) that have public good characteristics can result in constant or increasing returns at the aggregate level.

¹⁰ Although plausible and illustrative, conditional **b**-convergence is also a weaker and less robust concept than absolute **b**-convergence.

As has been shown by Barro and Sala-i-Martin (1995) the transitional growth process in the neoclassical model can be represented as

$$(1/T) \log (Y_{iT}/Y_{i0}) = \mathbf{a} + \mathbf{b} \log(Y_{i0}) + \mathbf{e}_{iT} \quad (7)$$

where i indexes the economy, Y is, once again, per capita income, 0 and T are the initial and final year of the sample and \mathbf{e} is the error term. Thus, typically, the existence of absolute \mathbf{b} -convergence is tested by regressing the growth in per capita income on its initial level for the cross-section of economies under consideration¹¹. If $\mathbf{b} < 0$ and is statistically significant, it is inferred that there is \mathbf{b} -convergence¹²; if, on the contrary, $\mathbf{b} > 0$, then it is said there is \mathbf{b} -divergence. It can be shown that $\mathbf{b} = [(1-e^{-bT})/T]$, where b represents the rate of convergence. That is, b indicates the speed at which per capita income approaches its steady-state level. The number of years needed to fill in half the gap to the steady-state is then given by the expression $0.5 = \exp^{-bT}$

The regression approach generally used to test the existence of \mathbf{b} -convergence has been criticized by some economists -the seminal papers are Friedman (1992) and Quah (1993-a). They point out that the aforementioned technique can produce biased estimates of \mathbf{b} -convergence because it is subject to the Galton's fallacy. To solve the problem, Friedman (1992) proposes the use of the coefficient of variation, suggesting that it provides an unbiased estimate of \mathbf{b} -convergence; this is the notion of convergence that has been previously referred to as \mathbf{s} -convergence. Quah (1993-a) proposes a completely different measure of

¹¹ The model can be modified to include in the RHS the term $\mathbf{y}X_{i0}$, where X_{i0} denotes a set of control variables at the initial year, with associated coefficients \mathbf{y} , determining the individual equilibrium rates of the economies. A negative value for \mathbf{b} is considered to imply conditional convergence.

¹² Chatterji (1993) maintains that this result is representative of weak convergence, the existence of strong convergence requiring both the existence of a steady state in which per capita income is equalized and the presence of dynamic forces driving the economy, in the long run, to this steady state. For this to happen, Chatterji (1993) shows that $-2 < \mathbf{b} < 0$.

convergence (the analysis of the dynamics of evolving cross-country distributions of per capita income), which will be discussed below.

Another proposal to solve the bias of the regression approach has been put forward by Boyle and McCarthy in two papers (1997, 1999). These authors propose an indicator of rank concordance (Kendall's index of rank concordance), labeled as *g*-convergence, which, in addition to *s*-convergence, offers a good test for *b*-convergence. This new indicator has also the virtue of capturing the changes in the ordinal ranking of the economies in the per capita income distribution. The index can be computed either in a multi-annual version or in a binary version. In the second case, the index is given by the expression

$$g = \text{Variance}(RPCY_{iT} + RPCY_{i0}) / \text{Variance}(2 * RPCY_{i0}) \quad (8)$$

where $RPCY_i$ refers to the ranks of per capita income of the *i*-th economy. This rank concordance index ranges from 0 to 1, the closer the index is to 0 the greater the mobility within the income distribution.

Simple regression models of the type shown in equation (7) rule out the possibility that some particular groups of economies may form a convergence club. As was pointed out above, there exists a convergence club when a specific group of economies tend to converge among themselves but diverge from economies belonging to other group(s). Fortunately, as Chatterji and Dewhurst (1996) have shown, equation (7) can be modified in order to incorporate the possibility of the existence of convergence clubs. This implies rewriting equation (7) so that the key variable is the natural log of the relative per capita income¹³.

That is:

¹³ Chatterji (1993) and Chatterji and Dewhurst (1996) take as the relevant variable the log of the ratio between the per capita income of the leader economy to the per capita income of any other economy. This procedure is accurate if the leader economy does not change over time; because this may not be the case, we prefer to take as the relevant variable the log of the ratio of any per capita income economy to the average (subscript A) per capita income.

$$(1/T) [\log (Y_{iT}/Y_{i0}) - \log (Y_{AT}/Y_{A0})] = \mathbf{a} + \mathbf{b} [\log(Y_{i0}) - \log(Y_{A0})] + \mathbf{e}_{iT} \quad (9)$$

If, for simplicity, we drop \mathbf{a} and \mathbf{e}_{it} in (7) and denote by $z_T = [\log (Y_{iT}/Y_{i0}) - \log (Y_{AT}/Y_{A0})]$ and by $z_0 = [\log (Y_{i0}) - \log (Y_{A0})]$, equation (9) can be rewritten as

$$(1/T) z_{iT} = \mathbf{b} z_{i0} \quad (10)$$

In order to allow for the existence of multiple (k) convergence clubs, equation (10) can be reformulated as

$$(1/T) z_{iT} = \sum_{k=1}^K \mathbf{b}_k (z_{i,t-T})^k = \mathbf{f}(z_{i,t-T}) \quad (11)$$

where the functional form \mathbf{f} depends on the data. A graph of z_{iT} against z_{i0} , shows the gap of each economy to the average at the final and initial years. If this line is then compared to a 45° line (Figure 3), four different situations become apparent.

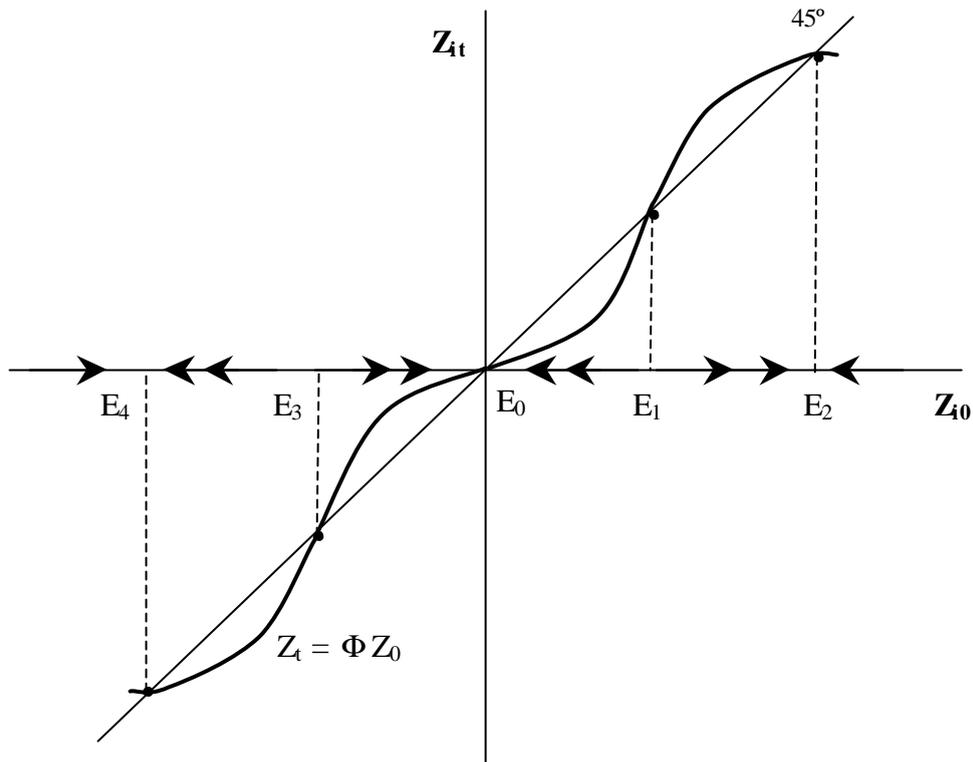


Figure 3.- Convergence clubs

If the initial situation is between E_1 and E_3 the gap to the average will tend to decline and the economy will converge to the equilibrium point E_0 . On the other hand, if the initial gap is greater than E_1 (E_3), then the gap will tend to increase over time and will converge to E_2 (E_4). This implies the existence of multiple steady-states or locally stables. In other words, this implies the existence of different convergence clubs.

4. Time-series convergence (Stochastic convergence)

Standard tests of b -convergence are carried out by using cross-sectional techniques. Bernard and Durlauf (1995, 1996), for example, consider the so-called unit root or stochastic convergence. According to these authors, stochastic convergence applies if per capita income disparities between two economies follow a stationary process, with a zero mean. This implies that the two economies have reached their own steady state and that shocks are not persistent but short-lived. Obviously, when per capita income differences between these two economies contain a unit root, stochastic convergence fails to hold.

There are at least three interrelated concepts of stochastic (or times-series) convergence. The concept of strong convergence –or asymptotically perfect convergence according to Bernard and Durlauf (1995)- holds when the difference between two time series Y_{it} and Y_{jt} , where i and j are any pair of economies within the sample, contain neither a unit root nor a time trend, either deterministic or stochastic. In formal terms

$$\lim_{T \rightarrow \infty} E(Y_{iT} - Y_{jT} | \mathbf{x}_0) = 0 \quad (12)$$

where \mathbf{x}_0 captures all relevant information at time T . This implies that between the two series there exists a cointegration vector (1, -1). This concept of convergence has been criticized because it is rather strict: for the strong convergence to exist it is necessary that the long-run expected value (forecast) of the per capita income differences between the two economies is equal to zero. Thus, this concept of convergence is not very useful when the objective is to determine whether convergence has taken place in the past or not.

An alternative, weaker concept of stochastic convergence (sometimes called asymptotically relative convergence) implies that the aforementioned disparities between any two economies within the sample do not need to converge to zero but to a finite constant. Following Bernard and Durlauf (1996), this definition “considers the behavior of the output differences between two economies over a fixed time interval and equates convergence with the tendency of the difference to narrow” (p. 165). This can be written as

$$E(Y_{i,T} - Y_{j,T} | \mathbf{x}_t) < (Y_{i,0} - Y_{j,0}) \quad (13)$$

where 0 refers to the present and T to some year in the future. According to this definition, the difference between the two time series should also be stationary, but now the time trend can be deterministic. Once again, the only cointegration vector between these two series can be (1, -1).

Finally, the less strict concept of stochastic convergence holds when, although the two time series have different trends, there exist a functional relationship between them in such a way that

$$\lim_{T \rightarrow \infty} E(Y_{iT} - \mathbf{b}Y_{jT} | \mathbf{x}_0) = 0; \mathbf{b} > 0 \quad (12)$$

In this case, both series are also cointegrated, but the cointegration vector is now (1, - \mathbf{b}).

Although the stochastic approach to convergence solves some of the problems of the standard, cross-section convergence approach, it also has some important shortcomings. In particular to our case, and as is widely known, unit root tests (augmented Dickey-Fuller¹⁴ and/or Phillips-Perron) suffer from low statistical power in finite samples, which implies that they might lead to failures in rejecting the null-hypothesis of non-stationarity. This is closely related to the existence of gaps (or structural breaks) in the series, in which case it is possible to accept the existence of unit roots when they do not exist. Under these circumstances, it is possible to wrongly admit or reject the existence of convergence. Recently, more powerful tests (panel unit root tests) have been proposed to address the issue of low statistical power of univariate unit root tests.

4. Distribution dynamics

As has been previously mentioned, both standard s and b -convergence approaches, but mainly the second one, have been heavily criticised on the grounds that they incur in the Galton's fallacy. This implies that negative b coefficients are compatible with the absence of convergence (Quah, 1993-a). Although Sala-i-Martin (1994) argues that b -convergence offers interesting insights about the mobility of income within a given distribution and that s -convergence shows how the distribution evolves over time, it is true that standard convergence analysis is, as a general rule, uninformative about the dynamics of per capita income distribution across countries. A cross-section

¹⁴ To run a Dickey-Fuller test on a time series Y_t is necessary to estimate the equation

$$Y_t = \mathbf{a} + \mathbf{b}Y_{t-1} + \mathbf{e}_t \quad t = 1, \dots, T$$

and run a one-sided t -test on the hypothesis that $\mathbf{b} = 1$. Thus $H_0: \mathbf{b} = 1$; $H_A: \mathbf{b} < 1$, so if the estimated \mathbf{b} is significantly less than 1, then the null hypothesis of non-stationarity can be rejected; this is tantamount to saying that the time series Y_t has a unit root or is integrated of order 1: I(1). To test whether $\mathbf{b} = 1$ is the same than to test whether $\mathbf{q} = 0$ in equation $\Delta Y_t = Y_t - Y_{t-1} = \mathbf{a} + \mathbf{q}Y_{t-1} + \mathbf{e}_t$, where $\mathbf{q} = \mathbf{b} - 1$.

regression cannot capture the essence of a convergence process in that it has implicitly a dynamic component; at the best, the cross-section regression approach can just approximate this process.

Quah's approach is, in this regard, much more promising in that it offers additional insights on the phenomenon of convergence by emphasizing not only a particular snapshot of the income distribution but to the evolution of the entire income distribution over time. In particular, the distribution dynamics approach involves estimating both the density functions associated with the entire income distribution at each point in time and tracking how it evolves over time.

A density function $f(y)$ is the mathematical counterpart of a smooth curve that represents the probability distribution of a continuous random variable Y , where y refers to specific value of Y . The density function $f(y)$ informs about the location, shape and other external characteristics of the income distribution Y at a given point in time. The external shape of the density function may be unimodal, bimodal or multimodal. This shows whether there is convergence, polarisation or stratification. In the typical case of bimodality, or twin-peakedness studied by Quah (1996) there is a clear polarisation between a group of rich economies and a group of poor economies, whereas the middle-income class economies tends to disappear.

Generally speaking, density functions are estimated non-parametrically by using the kernel method, this implies the substitution of the boxes in a histogram by smooth "bumps". The kernel function determines the shape of the "bumps", a crucial element in it being the bandwidth that controls the smoothness of the density. Silverman (1986) has proposed a method for selecting the bandwidth that is currently the most widely used in density function estimations.

Regarding intra-distribution dynamics, a standard methodology currently employed to trace movements within a distribution is based on the computation of transition probability matrices. Thus, if F_0 and F_T refer to the initial and final

distribution, the link between them can be defined as $F_T = M^T * F_0$ where M^T represents the transition probability matrix. Operator M^T is approximated by dividing the income distribution into intervals or “income states”. Due to the lack of sound theoretical methods to obtain an appropriate partition of the income distribution, the selection of “income states” is somewhat arbitrary. After having divided the distribution into “income states”, it is necessary to observe how many of the economies which are in a given income state in the initial period end up in that very state or elsewhere. Thus, each entry in the transition matrix refers to the probability that an economy in a given state income transits to another different state or stays in the initial one. Interpreting the transition matrix is, then, as follows: the elements (probabilities: row probabilities add up to 1) on the diagonal indicate persistence while all the other elements in the matrix represent mobility. If we assume that $t \rightarrow \infty$, the transition matrix allows to take a long run view of the evolution of the entire distribution, if nothing structural were to change in the system being analysed: the result is the so-called “ergodic” distribution¹⁵ or long-run steady state.

The main drawback of this approach is the arbitrariness when choosing the number and size of the “income states”. The stochastic kernel approach solves this problem by replacing the discrete income states by a continuum of states: thus, a stochastic kernel can be interpreted as the counterpart of a transition probability matrix with an infinite number of rows and columns. The stochastic kernels are represented as three dimensional diagrams and contour plots (Overman and Puga, 2002), as shown in Figure 4. To read the three dimensional diagram is useful to think of period 0 (T) axis as the rows (columns) of the transition matrix. Then, starting from any point on the 0 axis any slice parallel to the T axis traces out a probability density, describing the likelihood of transition, over T years, into different parts of the income space, conditional on beginning at a specific income (point) at time 0 .

¹⁵ As it has been pointed out by Quah (1993-b) *nothing ...enforces the existence or uniqueness of an ergodic distribution* (p. 431)

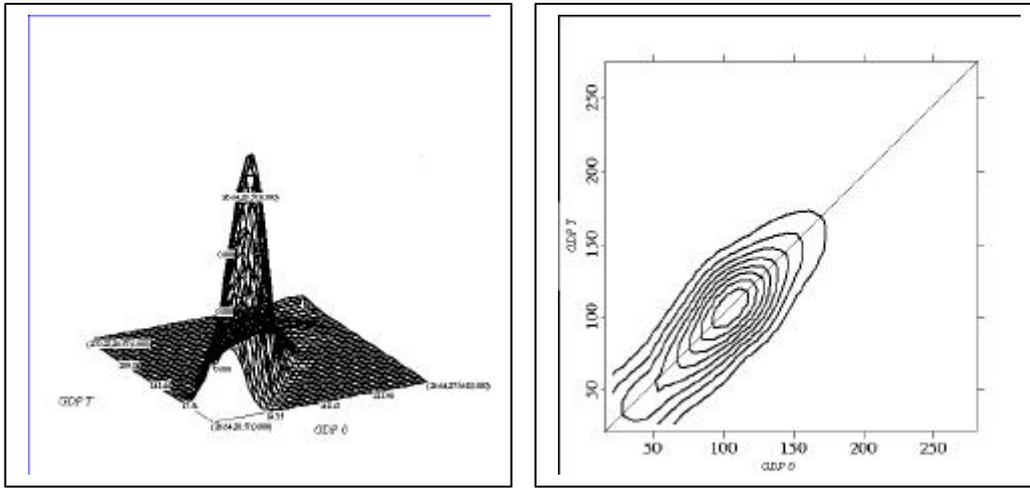
On the other side, the lines on the contour plots connect points of the same height (density) on the corresponding three dimensional diagrams. Then, on the contour plots is easy to understand that:

- 1) If the probability mass concentrates along the positive slope diagonal, this indicates persistence;
- 2) If the probability mass concentrates along the negative slope diagonal, this indicates overtaking of the economies in the ranking;
- 3) If the probability mass runs parallel to the T axis, this indicates that the probability of being in any state at period T is independent of their position in the initial (0) period; and
- 4) If the probability mass runs parallel to axis t , this indicates convergence.

The main shortcoming of this approach is that its interpretation is not as direct and clear-cut as that of the transition probability matrix because it does not offer quantitative information about the degree of mobility (or persistence); it only offers qualitative information.

Figure 4.- Stochastic kernel

5. Spatial analysis convergence



In all the aforementioned notions of convergence, the spatial dimension of the data under consideration is completely neglected. Because these spatial effects are largely ignored, problems of model misspecification in the previous convergence analysis may arise; in particular, OLS estimates ignoring spatial effects –as in the standard **b**-convergence approach- will be inefficient and/or biased (Anselin, 1988).

Spatial effects refer to spatial dependence (autocorrelation) and/or spatial heterogeneity, not being an easy task to differentiate between them from a practical point of view. Spatial autocorrelation implies that the observations (economies) in cross-sectional data are not independent. Following Anselin (1988), it means “the existence of a functional relationship between what happens at one point in space and what happens elsewhere” (p. 11). Spatial dependence can originate either as a result of a true spatial interaction among the economies (substantive spatial dependence) or as a measurement error problem (nuisance spatial dependence). Spatial heterogeneity comes from the lack of homogeneity of the economies under consideration.

The statistical device usually employed to test for the presence of spatial effects (spatial dependence) is the Moran’s I statistic expressed as

$$I_t = \left(\frac{n}{s_0} \right) \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} y_{it} y_{jt}}{\sum_{i=1}^n \sum_{j=1}^n y_{it} y_{jt}} \quad (13)$$

where w_{ij} is an element of a weight matrix W , n is the number of economies, y_{it} is the log of per capita income of the economy i in year t , and s_0 is equal to the sum of all elements of the matrix W ¹⁶. W is a non-stochastic square matrix which elements represent the intensity of the interdependence between each two economies; in the simplest case $w_{ij}=1$ if the economies i and j are neighbours and 0 otherwise.

When the Moran's coefficients are statistically significant, they provide support for the hypothesis of global spatial effects or global spatial dependence¹⁷. This being the case, the β -convergence approach needs to be reformulated in order to consider the effects of this spatial dependence. There are mainly three alternative specifications of the β -convergence unconditional equation in order to consider spatial dependence¹⁸: the spatial error model, the spatial lag model and the spatial cross-regressive model.

The first alternative, takes place when the error term follows a spatial autoregressive process expressed as $\mathbf{e}_t = \mathbf{I}W\mathbf{e}_t + u_t$, where \mathbf{I} is a scalar spatial error coefficient and u_t is normally distributed with a zero mean and a constant variance. The β -convergence equation becomes

$$(1/T) \log (Y_{iT}/Y_{i0}) = \mathbf{a} + \mathbf{b} \log(Y_{i0}) + (\mathbf{I} - \mathbf{I}W)u_{iT} \quad (14)$$

In the second specification of spatial dependence –coming from actual interaction among economies- the regression equation is

$$(1/T) \log (Y_{iT}/Y_{i0}) = \mathbf{a} + \mathbf{b} \log(Y_{i0}) + (\mathbf{r}W/T) \log (Y_{iT}/Y_{i0}) + \mathbf{e}_T \quad (15)$$

¹⁶ Usually, the weight matrix is row standardised so the elements of the rows sum to 1. For other possible specifications of the weight matrix see, for instance, Cliff and Ord (1981).

¹⁷ When the values of I are larger (smaller) than the expected value $E[I_t(k)]=-1/(n-1)$, this indicates the existence of a positive (negative) spatial autocorrelation.

¹⁸ - Anselin (2003) offers a new taxonomy of formal models of spatial effects in cross-sectional data.

where \mathbf{r} is the scalar lag parameter. Finally, in the third model specification – where the spatial variable is the independent variable- the spatial lag of the starting incomes is added to the original \mathbf{b} -convergence unconditional equation, which becomes

$$(1/T) \log (Y_{iT}/Y_{i0}) = \mathbf{a} + \mathbf{b} \log(Y_{i0}) + \mathbf{t}W \log(Y_{i0}) + \mathbf{e}_{iT} \quad (16)$$

SUMMARY AND CONCLUSIONS

This paper has addressed the topic of how to measure real economic convergence. Importantly, it has shown that there are different concepts of convergence and that each one is in need of its own type of measurement technique. Notwithstanding the concepts of \mathbf{s} -convergence and \mathbf{b} -convergence continue to be the most popular methods of measurement. The former method evaluates convergence through the time evolution of a summary measure of dispersion and the latter method through the estimate of cross-section regression.

As it has been shown, there are other ways to evaluate the convergence process that are also illuminating. Among these are the Gini and Theil inequality measures. These indices permit the disaggregation of the convergence process (or divergence) in different components. The extension by Chatterji and Dewhurst to deal with the potential existence of convergence clubs is also particularly interesting. This same topic is addressed, although from a different perspective, in the works pioneered by Quah. Density functions not only allow the estimation of the external form of the distribution but also the potential existence of convergence clubs. This same approach, via the estimate of either transition matrices or stochastic kernels, also gives insights into the degree of mobility within the distribution. Finally, and in order to take into consideration the spatial dimension of the income distribution, spatial dependence can also be included in the analysis, rendering in this way a more accurate estimate of the convergence process.

REFERENCES

- Abramovitz, M. (1986), "Catching Up, Forging Ahead and Falling Behind", *Journal of Economic History*, 46, pp. 385-406.
- Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Kluwer, Dordrecht.
- Anselin, L. (2003), "Spatial externalities, spatial multipliers and spatial econometrics", *International Regional Science Review*, no. 2, pp. 153-166.
- Barro, R. and X. Sala-i-Martin (1991), "Convergence across States and Regions", *Brooking Papers on Economic Activity*, no. 1, pp 107-158.
- Barro, R. and X. Sala-i-Martin (1992), "Convergence", *Journal of Political Economy*, no. 2, pp. 223-251.
- Barro, R. and X. Sala-i-Martin (1995), *Economic growth*, McGraw-Hill, New York.
- Baumol, W. (1986), "Productivity Growth, Convergence and Welfare: What the Long-Run Data Show", *American Economic Review*, no. 5, pp 1072-1085.
- Bernard, A. and S. Durlauf (1995), "Convergence in International Outputs", *Journal of Applied Econometrics*, 10, pp 97-108.
- Bernard, A. and S. Durlauf (1996), "Interpreting tests of the convergence hypothesis", *Journal of Econometrics*, 71, pp. 161-173.

Boyle, G.E. and T.G. McCarthy (1997), "A simple measure of β -convergence", *Oxford Bulletin of Economics and Statistics*, no. 2, pp 257-264.

Boyle, G.E. and T.G. McCarthy (1999), "Simple measures of convergence in per capita GDP: a note on some further international evidence", *Applied Economic Letters*, no. 6, pp 343-347.

Chatterjee, M. (1993), "Convergence clubs and endogenous growth", *Oxford Review of Economic Policy*, no. 4, pp. 57-69.

Chatterjee, M. and Dewhurst (1996), "Convergence clubs and relative economic performance in Great Britain: 1977-1991", *Regional Studies*, no. 1, pp. 31-40

Cheshire, P. and G. Carbonaro (1995) "Convergence-Divergence in Regional Growth Rates: an Empty Black Box?", in: H.W. Armstrong and R.W. Vickerman (eds.), *Convergence and Divergence Among European Regions*, Pion Limited, London.

Cilff, A. and J.K. Ord (1981), *Spatial processes: Models and applications*, Pion, London.

Dalgaard, C.J. and J. Vastrup (2001), "On the measurement of σ -convergence", *Economic Letters*, no. 70, pp. 283-287.

De Lombaerde, P. (2002), "Integración asimétrica", in: P. De Lombaerde (ed.), *Integración asimétrica y convergencia económica en las Américas*, Universidad Nacional de Colombia – Antropos, pp. 16-32.

Dickey, H. (2001), "Regional earnings inequality in Great Britain: A decomposition analysis", *Regional Studies*, no. 7 pp.605-612.

Esteban, J.M. and D. Ray (1994), "On the measurement of polarization", *Econometrica*, no. 4, pp. 819-851.

Friedman, M.J. (1992), "Do old fallacies ever die?", *Journal of Economic Literature*, 30, pp. 2129-2132.

Funck, B. and L. Pizzati (eds.) (2003), *European Integration, Regional Policy, and Growth*, The World Bank, Washington D.C.

Meeusen, W. and J. Villaverde (eds.) (2002a), *Convergence Issues in the European Union*, Edward Elgar, Cheltenham.

Meeusen, W. and J. Villaverde (2002b), "Introduction and Outline", in W. Meeusen and J. Villaverde (eds.), *Convergence Issues in the European Union*, Edward Elgar, Cheltenham, pp. 1-5.

Overman, H. and D. Puga (2002), "Unemployment clusters across European regions and countries", *Economic Policy*, no. 34, pp. 115-147.

Quah, D. (1993a), "Galton's fallacy and test of the convergence hypothesis", *The Scandinavian Journal of Economics*, 95, pp.427-443.

Quah, D. (1993b), "Empirical cross-section dynamics in economic growth", *European Economic Review*, no. 2-3, pp. 426-434.

Quah, D. (1996), "Twin peaks: Growth and convergence in models of distribution dynamics", *Economic Journal*, 106, pp. 1045-1055.

Quah, D. (1997), "Empirics for economic growth and distribution. Polarization, stratification and convergence clubs", *Journal of Economic Growth*, no. 1, pp. 27-59.

- Rebelo, S. (1991), “Long-run policy analysis and long-run growth”, *Journal of Political Economy*, no. 3, pp. 500-521.
- Romer, P. (1986), “Increasing returns and long run growth”, *Journal of Political Economy*, no. 5, pp. 1002-1037.
- Sala-i-Martin, X. (1994), “Cross Sectional Regressions and the Empirics of Economic Growth”, *European Economic Review*, no. 3-4, pp. 739-747.
- Sala-i-Martin, X. (1996), “Regional Cohesion: Evidence and Theories of Regional Growth and Convergence”, *European Economic Review*, no. 6, pp. 1325-1352.
- Silverman, B.W. (1986), *Density estimation for statistics and data analysis*, Chapman and Hall, New York.
- Solow, R. (1956), “A contribution to the theory of economic growth”, *Quarterly Journal of Economics*, vol. 70, pp. 65-94
- Villaverde, J. (2003), “Regional convergence, polarisation and mobility in the European Union, 1980-1996”, *Journal of European Integration*, No. 1, pp. 73-86.
- Villaverde, J. and B. Sánchez-Robles (2002), “Convergence or Twin Peaks? The Spanish Case”, in: W. Meeusen and J. Villaverde (eds.), *Convergence Issues in the European Union*, Edward Elgar, Cheltenham, pp. 41-60.
- Villaverde, J. and A. Maza (2003), “Desigualdades regionales y dependencia espacial en la Unión Europea”, *CLM-Economía*, no. 2, pp. 109-128.
- Viñals, J. (1994), “Es posible la convergencia en España? En busca del tiempo perdido”, Banco de España, *Documento de Trabajo*, no. 9430.