



The Classical Approach to Convergence Analysis

Author(s): Xavier X. Sala-i-Martin

Source: *The Economic Journal*, Vol. 106, No. 437 (Jul., 1996), pp. 1019-1036

Published by: Blackwell Publishing for the Royal Economic Society

Stable URL: <http://www.jstor.org/stable/2235375>

Accessed: 15/06/2010 10:31

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=black>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Royal Economic Society and Blackwell Publishing are collaborating with JSTOR to digitize, preserve and extend access to *The Economic Journal*.

<http://www.jstor.org>

THE CLASSICAL APPROACH TO CONVERGENCE ANALYSIS*

Xavier X. Sala-i-Martin

The concepts of σ -convergence, absolute β -convergence and conditional β -convergence are discussed in this paper. The concepts are applied to a variety of data sets that include a large cross-section of 110 countries, the sub-sample of OECD countries, the states within the United States, the prefectures of Japan, and regions within several European countries. Except for the large cross-section of countries, all data sets display strong evidence of σ -convergence and absolute β -convergence. The cross-section of countries exhibits σ -divergence and conditional β -convergence. The speed of conditional convergence, which is very similar across data sets, is close to 2% per year.

Will relatively poor economies remain poor for many generations? Will the rich countries of the year 2100 be the same countries that are relatively rich today? Is the degree of income inequality across economies increasing or falling over time? The importance of these questions, which lie at the heart of the convergence debate, is obvious to anyone interested in general worldwide welfare.

Even though economists have been interested in these issues for many decades, it was not until the end of the 1980s that the convergence debate captured the attention of mainstream macroeconomic theorists and econometricians. In addition to the inherent importance of the questions dealt with, the reason for this sudden increase in interest was two-fold. First, the existence of convergence across economies was proposed as the main test of the validity of modern theories of economic growth. Moreover, estimates of the speeds of convergence across economies (to be more precisely defined in a later section) were thought to provide information on one of the key parameters of growth theory: the share of capital in the production function (see the discussion in Section III). For this reason, growth theorists started paying close attention to the evolution of the convergence debate. Second, and perhaps more importantly, a data set on internationally comparable GDP levels for a large number of countries became available in the mid 1980s.¹ This new data set allowed empirical economists to compare GDP levels across a large number of countries, and to look at the evolution of these levels over time, a necessary feature for the study of the convergence hypothesis.

In this paper I shall discuss the classical approach to convergence analysis. I call this the classical approach because it was the first methodology to be used in the literature and because it uses the traditional techniques of classical econometrics, a feature that is shared by some, but not all, of the alternative approaches. Perhaps more importantly, like classical art, it is the basis of

* I thank Paul Cashin, Michelle Connolly and Tornarà A. Serrica-i-Plena for useful comments.

¹ The University of Pennsylvania project was originally started in the 1960s by A. Kravis and was finished by Allan Heston and Robert Summers (see Summers and Heston (1991)).

reference and target of criticism of all other methodologies, and it has survived the challenges of more modern and 'surrealist' movements.²

The rest of the paper is organised as follows. Section I introduces two useful definitions of convergence and highlights their similarities and differences. Section II analyses some evidence on convergence using a sample of 110 economies. In section III I interpret the above evidence in light of the neoclassical model and introduce the concept of conditional convergence. Section IV applies the concept of conditional convergence to the sample of 110 countries. Section V provides evidence on convergence for a number of regional data sets. I conclude in section VI.

I. DEFINITIONS OF CONVERGENCE

Two main concepts of convergence appear in the classical literature. They are called β -convergence and σ -convergence.³ We say that *there is absolute β -convergence if poor economies tend to grow faster than rich ones.*⁴ Imagine that we have data on real *per capita* GDP for a cross-section of economies. Let $\gamma_{i,t,t+T} \equiv \log(y_{i,t+T}/y_{i,t})/T$ be economy i 's annualised growth rate of GDP between t and $t+T$, and let $\log(y_{i,t})$ be the logarithm of economy i 's GDP *per capita* at time t . If we estimate the following regression

$$\gamma_{i,t,t+T} = \alpha - \beta \log(y_{i,t}) + \epsilon_{i,t}, \quad (1)$$

and we find $\beta > 0$, then we say that the data set exhibits absolute β -convergence.

The concept of σ -convergence can be defined as follows: *a group of economies are converging in the sense of σ if the dispersion of their real per capita GDP levels tends to decrease over time.* That is, if

$$\sigma_{t+T} < \sigma_t, \quad (2)$$

where σ_t is the time t standard deviation of $\log(y_{i,t})$ across i . The concepts of σ - and absolute β -convergence are, of course, related. If we take the sample variance of $\log(y_{i,t})$ from (1) (note that the growth rate is the difference between $\log(y_{i,t+T})$ and $\log(y_{i,t})$ divided by T), we will get a relation between σ_t and σ_{t+T} which depends on β . Intuitively, we can see that if the GDP levels of two economies become more similar over time, it must be the case that the poor economy is growing faster. As an illustration, Fig. 1 displays the behaviour of the log of GDP *per capita* ($\log(\text{GDP})$) for two economies over time. Imagine that we observe the data at two discrete intervals, t and $t+T$. Economy A starts

² Quah (1996) proposes a new way of telling a good empirical technique from a bad one, which is summarised by the following statement: 'if I can find a set of questions (no matter how unrelated to your initial query) that your research technique cannot address, then your technique is useless and mine becomes extremely useful... *even if my technique cannot answer those questions either*'. Notice that the strict application of this methodology would render all empirical techniques (with the possible exception of Quah's) both useless and extremely useful (see Salvador Dali (1965) for a description of a similar methodology applied to surrealist double images).

³ This terminology was first introduced in Sala-i-Martin (1990).

⁴ In Section III I will introduce the important concept of *conditional β -convergence*.

out being richer than economy *B*. there is an initial distance or dispersion between the two levels of income. In Panel *a*, the growth rate of economy *A* is smaller (actually negative) than the growth rate of economy *B* between times *t* and *t + T* and, therefore, we say that there is β -convergence. Moreover, because the dispersion of $\log(\text{GDP})$ at *t + T* is smaller than at time *t*, we also say that there is σ -convergence. Note that it is impossible for the two economies to be closer together at *t + T* without having the initially poor economy (in this case economy *B*) growing faster. In other words, *a necessary condition for the existence of σ -convergence is the existence of β -convergence.*

Moreover, it is natural to think that when an initially poor economy grows faster than a rich one, then the levels of GDP *per capita* of the two economies will become more similar over time. In other words, *the existence of β -convergence will tend to generate σ -convergence.* Panel *a* in Fig. 1 is an example where

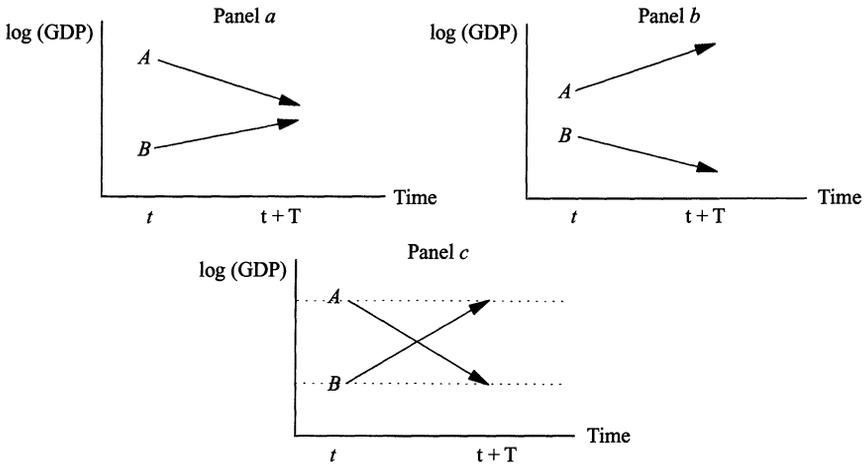


Fig. 1. The relation between σ and β -convergence.

β -convergence exists and is associated with σ -convergence. Panel *b* provides an example where the lack of β -convergence (the initially rich economy grows faster) is associated with the lack of σ -convergence (the distance between economies increases over time). Hence, it would appear that the two concepts are identical. However, at least theoretically, it is possible for initially poor countries to grow faster than initially rich ones, without observing that the cross-sectional dispersion fall over time. That is, we could find β -convergence without finding σ -convergence. In Panel *c*, I have constructed an example where the initially poor economy (*B*) grows faster than the initially rich (*A*), so there is β -convergence. However, the growth rate of *B* is so much larger than the growth rate of *A* that, at time $t + T$, *B* is richer than *A*. In fact, the example is such that, at time $t + T$, the distance between *A* and *B* is the same as it was at time t (except that now the rich economy is *B*). Hence, the dispersion between these two economies has not fallen, so there is no σ -convergence. In fact I could have constructed the example so that the dispersion at $t + T$ was larger than at

t. In that case there would have been σ -divergence despite there being β -convergence. It follows that β -convergence, although necessary, is not a sufficient condition for σ -convergence.

The reason why the two concepts of convergence may not always show up together is that they capture two different aspects of the world. σ -convergence relates to whether or not the cross-country distribution of world income shrinks over time. β -convergence, on the other hand, relates to the mobility of different individual economies within the given distribution of world income. Panels *a* and *b* are examples where the movements of the various economies over time change the final distribution of income. Panel *c*, on the other hand, is an example where there is mobility within the distribution, but the distribution itself remains unchanged.⁵

II. CROSS-COUNTRY EVIDENCE

Maddison (1991) provided data on GDP levels across 13 rich countries starting in 1870. These data were constructed following the methodology of the United Nations' International Comparison Project (ICP) so, in principle, the data across countries can be compared and are, therefore, suitable for use in the analysis of convergence. The main disadvantage of these data is that they are available for rich countries only, a problem that almost proved fatal to early students of convergence. Using Maddison's data, Baumol (1986) documented the existence of cross-country convergence. He found that convergence was especially strong after World War II. This evidence, however, was quickly downplayed by Romer (1986) and DeLong (1988), on the grounds of *ex-post* sample selection bias: by working with Maddison's data set of nations which were industrialised *ex-post* (that is, by 1979), those nations that did not converge were excluded from the sample, so convergence in Baumol's study was all but guaranteed. As soon as Maddison's data set was expanded to include countries that appeared rich *ex-ante* (that is, by 1870) the evidence for convergence quickly disappeared (DeLong, 1988 and Baumol and Wolff, 1988).

An obvious solution to the sample selection problem was to analyse a larger set of countries. This is where the newly created Summers–Heston data set came in handy, as it comprised GDP levels for more than one hundred countries. However, unlike Maddison's project, where the time series dimension of the data was quite large, 1960 is the first year for which the Summers–Heston

⁵ This possibility led some economists (most prominently Quah (1993)) to criticise the classical approach on three grounds. First they suggested that classical analysts were confusing the two concepts of convergence. Secondly, they argued that the only meaningful concept of convergence was that of σ . Finally, they said that the concept of β -convergence conveyed no interesting information about σ -convergence (or about anything else) so it should not be studied. Needless to say that the three points were not entirely correct. First, classical analysts were well aware of the distinction from the very beginning (see for example Easterlin (1960), Sala-i-Martin (1990) and Barro and Sala-i-Martin (1992)). In fact, that is why they made the distinction in the first place! Secondly, the intra-distributional mobility (reflected in β) is at least as interesting as the behaviour of the distribution itself (reflected in σ). In fact, it could even be argued that if mobility was very high, the evolution of σ would be uninteresting. Surprisingly, Quah (1994, 1996) highlights the importance of intra-distributional mobility in the context of stochastic Kernel estimators. Finally, β provides information about σ to the extent that any necessary condition does – the fact that the two phenomena tend to appear together in most data sets seems to support this view.

data are available. Hence, by using the Summers–Heston data set analysts could study a broader set of countries, but at the cost of a much shorter time span. In Fig. 2 I display the behaviour of the dispersion of GDP *per capita* for

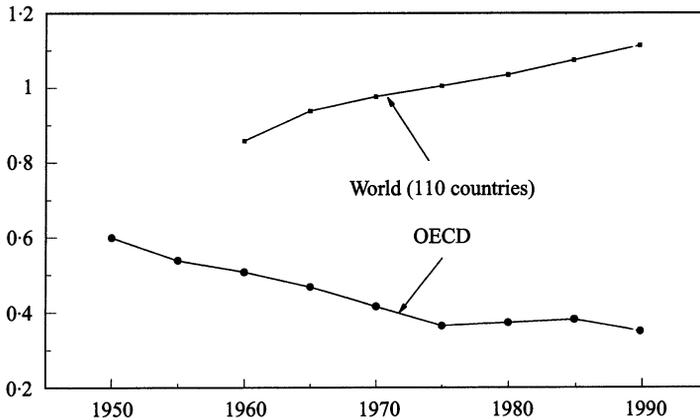


Fig. 2. Dispersion of GDP across 110 countries.

the set of 110 countries for which I have data in all years between 1960 and 1990. Note that the dispersion, σ , increases steadily from $\sigma = 0.89$ in 1960 to $\sigma = 1.12$ in 1980. The cross-country distribution of world income has become increasingly unequal: we live in a world where economies have diverged in the sense of σ over the last 30 years.

Fig. 3 analyses the existence of β -convergence across the same set of 110 economies. On the horizontal axis, I display the log of GDP *per capita* in 1960. On the vertical axis, I depict the growth rate between 1960 and 1990. The

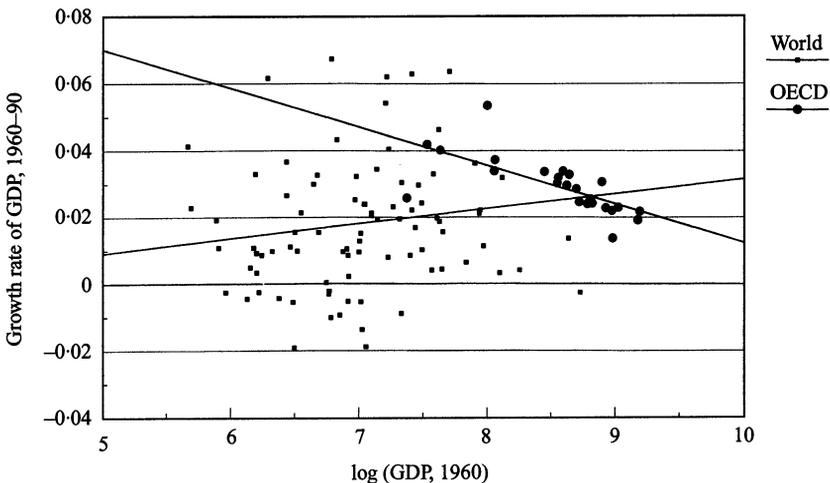


Fig. 3. Convergence across countries, 1960–90.

Table 1
Estimates of the speeds of β -convergence for a variety of data sets

Data set	Long-run single regression*				Panel estimates*
	Absolute convergence (1)		Conditional convergence (2)		Conditional convergence (3)
World 110 countries (1960–90)	–0.004 (0.002)	0.04 [0.0176]	0.013 (0.004)	0.46 [0.0134]	0.025 (0.0028)
OECD countries (1960–90)	0.014 (0.003)	0.48 [0.0062]	0.029 (0.008)	0.78 [0.0050]	—
United States 48 states (1880–1990)	0.021 (0.0003)	0.89 [0.0015]	0.017 (0.002)	0.89 [0.0015]	0.022 (0.002)
Japan 47 prefectures (1955–90)	0.019 (0.003)	0.59 [0.0027]	0.019 (0.004)	0.59 [0.0027]	0.031 (0.004)
Europe total 90 regions (1950–90)†	0.015 (0.002)	0.51 [0.0030]	0.015 (0.002)	0.52 [0.0030]	0.018 (0.003)
Germany (11 regions)	0.014 (0.006)	0.56 [0.0028]	0.014 (0.005)	0.55 [0.0027]	0.016 (0.006)
United Kingdom (11 regions)	0.020 (0.008)	0.62 [0.0021]	0.030 (0.007)	0.61 [0.0021]	0.029 (0.009)
France (21 regions)	0.016 (0.005)	0.55 [0.0023]	0.016 (0.004)	0.55 [0.0022]	0.015 (0.003)
Italy (20 regions)	0.010 (0.003)	0.46 [0.0033]	0.010 (0.003)	0.46 [0.0031]	0.016 (0.003)
Spain (17 regions) (1955–87)	0.021 (0.005)	0.63 [0.0042]	0.023 (0.007)	0.63 [0.0040]	0.019 (0.005)

* The regressions use nonlinear least squares to estimate equations of the form:

$$(1/T) \ln (y_{it}/y_{i,t-T}) = a - [\ln (y_{i,t-T})] (1 - e^{-\beta T}) (1, T) + \text{'other variables'}$$

where $y_{i,t-T}$ is the *per capita* income in country or region i at the beginning of the interval divided by the overall CPI. T is the length of the interval; 'other variables' are regional dummies and sectoral variables that hold constant temporary shocks that may affect the performance of a region in a manner that is correlated with the initial level of income.

Each column contains four numbers. The first one is the estimate of β . Underneath it, in parentheses, its standard error. To its right, the adjusted R^2 of the regression and below the R^2 , the standard error of the regression. The constant, regional dummies and/or structural variables are not reported in the Table.

The coefficients for Europe total include one dummy for each of the eight countries.

Columns 1 and 2 report the value of β estimated from a single cross section using the longest available data. Column 1 reports the coefficient when the only variable held constant is the initial level of income. Column 2 reports the value of β estimated when additional variables are held constant.

Column 3 reports the panel estimates when all the subperiods are assumed to have the same coefficient β . This estimation allows for time effects. For most countries, the restriction of β being constant over the subperiods cannot be rejected (see Barro and Sala-i-Martin (1995)).

† The regressions for Europe total allow each country to have its own constant term.

figure shows that the relation between growth and the initial level of GDP is not negative. In fact, the slope of the regression line (also shown in the fig.) is positive, although the fit is far from impressive.

In order to quantify the lack of convergence across these 110 countries, I estimate the following equation

$$\gamma_{i,t,t+T} = a - b \log (y_{i,t}) + \epsilon_{i,t,t+T} \quad (3)$$

Table 1 reports the estimated speed of convergence, β , for a variety of data

sets under three different setups.⁶ The first row relates to the large sample of 110 countries. The first column refers to the estimate of β when a single equation is estimated for the whole time period and no other explanatory variable is included. Each box in this table contains four numbers. The first one is the estimate of β . The number just below (in parentheses) is its standard error. To its right, we have the adjusted R^2 of the regression and below the R^2 , the standard error of the regression. The estimated speed of convergence for the cross-section of 110 countries is negative, $\beta = -0.004$ (s.e. = 0.002), so the relation between growth and initial income is positive as shown in Fig. 3. The R^2 is 0.04 and the standard error of the regression is 0.0176. During this period of 30 years, therefore, poor economies did not grow faster than rich ones. The set of 110 countries in the world did not converge in the sense of β .

III. INTERPRETATION OF THESE FINDINGS IN THE LIGHT OF MODELS OF ECONOMIC GROWTH: ABSOLUTE VERSUS CONDITIONAL CONVERGENCE

The lack of convergence across countries is an interesting finding on various grounds. It says that, in our world, the degree of cross-country income inequality not only fails to disappear, but rather tends to increase over time (σ -divergence). It also suggests that countries which are predicted to be richer a few decades from now are the same countries that are rich today (β -divergence). These findings may be used by economists or politicians to devise international institutions which may work to overturn this sombre tendency.

These findings were also seen by growth theorists in the mid-1980s as evidence against the neoclassical model of Ramsey (1928), Solow (1956), Swan (1956), Cass (1965), and Koopmans (1965), and as support for their new models of endogenous growth. The intuition behind this conclusion is the following: the assumption of diminishing returns to capital implicit in the neoclassical production function has the prediction that the rate of return to capital (and therefore its growth rate) is very large when the stock of capital is small and *vice versa*. If the only difference across countries is their initial levels of capital, then the prediction of the neoclassical growth model is that countries with little capital will be poor and will grow faster than rich countries with large capital stocks, so there will be cross-country β -convergence. Since the model does not predict the type of overshooting displayed in Panel *c* of Fig. 1, the prediction of β -convergence will tend to be associated with a reduction in cross-economy dispersion of income over time, i.e. σ -convergence.

More precisely, consider a neoclassical model with a Cobb–Douglas production function

$$Y_{i,t} = A_i K_{i,t}^\alpha L_{i,t}^{1-\alpha},$$

where $Y_{i,t}$ is economy i 's aggregate output at time t , $K_{i,t}$ and $L_{i,t}$ are the stock

⁶ If one estimates the linear equation (3), then the speed of convergence can be computed by using the formula $b = (1 - e^{-\beta T})/T$, where T is the length of time between two observations (that is, if we regress the growth rate of GDP between 1960 and 1985, then T is 25). Instead of estimating the linear relation (3) and using this formula to compute the speed, β , we can estimate a nonlinear least squares (NLS) relation directly. The direct estimation of NLS delivers standard errors for the speed of convergence directly. In this paper, I report the estimates of β (and not b) and its standard errors.

of capital and labour in that economy respectively, and A_i is the level of technology. Following Solow (1956) and Swan (1956), suppose that the saving rate in this economy is constant (the key results do not depend on this assumption) and that the rate of depreciation of K is δ , the rate of population growth is n and the rate of productivity growth is x . The dynamic equation that characterises the behaviour of economy i over time says that capital accumulation is the difference between overall savings and effective depreciation. If we log-linearise this dynamic equation around the steady state, we find that the growth rate of economy i between periods t and $t + T$ is given by (3). Moreover, given this production function, the parameter β is exactly equal to

$$\beta = (1 - \alpha)(\delta + n + x), \quad (4)$$

where α is again the capital share in the production function.⁷ Since, according to the neoclassical model, $0 < \alpha < 1$, the prediction is that $\beta > 0$. In other words, the neoclassical model predicts convergence.

This prediction contrasts with the implications of the first generation of models of endogenous growth (see, for example, Romer (1986) and Rebelo (1990)). These models rely on the existence of externalities, increasing returns and the lack of inputs that cannot be accumulated.⁸ The key point of these new models is the absence of diminishing returns to capital (the concept of capital should be understood in a broad sense that includes human capital), so these models do not exhibit the neoclassical model's convergence property. In terms of equation (4), the one sector models of endogenous growth are similar to the neoclassical model except that, in the former, α has a value of 1. Note that if $\alpha = 1$, equation (4) then says that the speed of convergence should be $\beta = 0$. For this reason, in the mid-1980s, the lack of convergence across countries was seen as evidence against the neoclassical model and in favour of the new models of endogenous growth.⁹

The Absolute Convergence Fallacy and the Concept of Conditional Convergence

The argument that the neoclassical model predicts that initially poor countries will grow faster than initially rich ones relies heavily on the key assumption that *the only difference across countries lies in their initial levels of capital*. In the real world, however, economies may differ in other things such as their levels of technology, A_i , their propensities to save, or their population growth rates. If

⁷ See Barro and Sala-i-Martin (1995, Chapter 1) for a derivation of this result. See also Chapter 2 of the same book for the extension of this result to the optimising version of the neoclassical model first developed in Ramsey (1928).

⁸ Labour, which is not purposely accumulated in the neoclassical model, is often substituted with human capital, the stock of which increases in accordance with the investment decisions of private agents.

⁹ Equation (4) is interesting for another reason. The parameters δ , n and x can be estimated fairly accurately. Hence, if we have an estimate of β , we will indirectly have an estimate of the capital share, α . This particular parameter is very important because the first generation of models of endogenous growth highlighted the importance of physical capital externalities and the existence of human capital. This meant that the traditional way to compute the capital share by using income shares was incorrect. Since the exact size of the externalities was unknown and the fraction of labour that could be accumulated in the form of human capital was also unknown, the relevant capital share (whose size was seen as crucial from a theoretical point of view) remained unknown. Equation (4) says that the convergence literature can provide an indirect way to say something about the size of α .

different economies have different technological and behavioural parameters, then they will have different steady state and the above argument (developed by the early theorists of endogenous growth) will be flawed. The intuition can be captured by a simple two-economy example. Imagine that the first economy is poor but is in the steady state. Accordingly, its growth rate is zero. The second economy is richer, but has a capital stock below its steady-state level. The model predicts that its growth rate is positive and, therefore, will be larger than the growth rate of the first economy, even though the first economy is poorer! What the model says is that, as the capital stock of the growing economy increases, its growth rate will decline and go to zero as the economy reaches its steady state. Hence, the prediction of the neoclassical model is that the growth rate of an economy will be positively related to the distance that separates it from its own steady state. This is the concept known in the classical literature¹⁰ as *conditional β -convergence*. To facilitate the distinction, the concept of β -convergence discussed above is sometimes called *absolute convergence*. Only if all economies converge to the *same steady state* does the prediction that poor economies should grow faster than rich ones holds true. This is because with common steady states, initially poorer economies will be unambiguously farther away from their steady state. In other words, the conditional convergence and the absolute convergence hypotheses coincide, only if all the economies have the same steady state. Since the neoclassical model predicts conditional convergence, the evidence on absolute convergence discussed in the previous section says little about the validity of the model in the real world.

To test the hypothesis of conditional convergence one has to, somehow, hold constant the steady state of each economy. Classical analysts have tried to hold the steady state constant in two different ways. The first one is by introducing variables that proxy for the steady state in a regression like (1) or (3). In other words, instead of estimating (1) or (3) one estimates

$$\gamma_{i,t,t+T} = a - b \log(y_{i,t}) + \psi \mathbf{X}_{i,t} + \epsilon_{i,t,t+T}, \quad (5)$$

where $\mathbf{X}_{i,t}$ is a vector of variables that hold constant the steady state of economy i , and $b = (1 - e^{-\beta T})/T$. If the estimate of β is positive once $\mathbf{X}_{i,t}$ is held constant, then we say that the data set exhibits *conditional β -convergence*. In Section IV I will use this first approach to condition the data.

The second way to hold constant the steady state is to restrict the convergence study to sets of economies for which the assumption of similar steady states is not unrealistic. For example, because we think that the technology, institutions, and tastes of most African economies are very different from those of Japan or the United States, the assumption that these economies converge to a common steady state is not realistic. However, the technological and institutional differences across regions within a country or across 'similar' countries (for example, those of the OECD) are probably smaller. Hence, we may want to look for absolute convergence within these sets of 'more similar' economies. This second approach is used in Section V.

¹⁰ See Sala-i-Martin (1990), Barro and Sala-i-Martin (1992) and Mankiw *et al.* (1992).

IV. CONDITIONAL CONVERGENCE (I): MULTIPLE REGRESSION
ANALYSIS

The concept of conditional β -convergence defined above suggests the estimation of a multiple regression like (5). If the neoclassical model is correct and the vector \mathbf{X} successfully holds constant the steady state, we should find a positive b (and therefore a positive β). The key, therefore, is to find variables that proxy for the steady state, and economic theory should guide our search for such variables. Different growth models suggest different variables. The strict version of the Solow–Swan model, for example, says that the steady state depends on the level of technology, A , the saving rate, and the parameters δ , n , and x . A broad interpretation of ‘technology’ would allow A to capture various types of distortions (public or otherwise), political variables, etc. Following Barro (1991), a large literature has estimated equations like (5). In this literature, more than 50 variables have been used in this type of analysis (and found to be significant in at least one regression).¹¹ The key point for our purposes here is that, once some variables that can proxy for the steady state are held constant, the estimate of β becomes significantly positive, as predicted by the neoclassical theory. This finding is robust to the exact choice of \mathbf{X} .¹² For example, columns 2 and 3 of Table 1 report the estimate of β when additional variables are held constant. In this particular case, the primary and secondary school enrolments, the saving rate, and some political variables are used as the vector \mathbf{X} . Note that, unlike column 1, the estimate of β for the 110 countries (row one of Table 1) is now positive and significant, $\beta = 0.013$ (s.e. = 0.004). Column 3 divides the sample period 1960–90 into two subperiods and estimates β by restricting it to be the same across sub-periods. The estimated β is now 0.025 (s.e. = 0.003) which is, again, significantly positive.

The conclusion is that the sample of 110 countries in the world displays conditional β -convergence. Furthermore, the estimated speed of conditional convergence is close to 2% per year. I should re-emphasise, however, that this does not mean that poor economies grow faster or that the world distribution of income is shrinking. These are phenomena captured by the concepts of *absolute* β -convergence and σ -convergence and, in this sense, the set of economies diverges unambiguously. What this evidence says is that economies seem to approach some long-run level of income which is captured by the vector of variables \mathbf{X} , and the growth rate falls as the economy approaches this long-run level.

V. CONDITIONAL CONVERGENCE (II): REGIONAL EVIDENCE

The second method for ‘holding constant the steady state’ is to analyse sets of economies that appear similar to the researcher, so that the assumption of a common steady state is reasonable. For example, the OECD economies and

¹¹ See for example, Mankiw *et al.* (1992), Levine and Renelt (1992) and Barro and Sala-i-Martin (1995, chapter 12).

¹² The existence of convergence is less strong during the 1980s, a phenomenon that holds true for almost all data sets analysed in this paper. See Levine and Renelt (1992) for some evidence on this point.

regions within countries could be considered as similar *ex-ante*. The neoclassical theory that guided our analysis suggests that, if it is true that these sets of economies are similar, we should find that these data sets display absolute β -convergence as well as σ -convergence. If evidence of absolute β -convergence is to be found anywhere, it will be in these data sets.¹³

OECD Economies

Figs. 2 and 3 also display the convergence behaviour of a subset of the world sample: the OECD countries. In Fig. 2, I plot the cross-sectional dispersion of GDP *per capita* for these economies starting in 1950.¹⁴ The key message of this Fig. is that there was an overall downward trend in dispersion between 1950 and 1990, interrupted only during the period from 1975 to 1985.

Fig. 3 also highlights the differential behaviour of OECD economies (which are denoted by black dots). For these countries, the relation between growth and the initial level of income is significantly negative as depicted, in Fig. 3, by the downward-sloping regression line. OECD economies have thus also converged in the sense of β . The exact estimate of the speed of absolute β -convergence can be found in the second row of Table 1. When no other variables are held constant, the estimated speed of convergence is $\beta = 0.014$ (s.e. = 0.003). Hence, OECD economies exhibit absolute β -convergence. When additional conditioning variables are held constant (column 2 in Table 1), the estimated β is 0.029 (s.e. = 0.008). Hence, the estimated speed of convergence across OECD economies is close to 2% per year.¹⁵

Dowrick and N'guyen (1989) add to this evidence by using various measures of productivity. They show that, not only do GDP levels *per capita* converge across OECD economies, but so do the levels of productivity.

The States of the United States

The third row of Table 1 shows estimates of equation (3) for 48 U.S. states for the period 1880 to 1990.¹⁶ The first and second columns report the estimate for a single long sample (1880–1990). Column 1 uses the initial level of income as the only explanatory variable. The estimated speed of convergence is $\beta = 0.021$ (s.e. = 0.0003). The good fit ($R^2 = 0.89$) can be appreciated by looking at Fig. 4, which is a scatter plot of the average growth rate of income *per capita* between 1880 and 1990 versus the log of income *per capita* in 1880. Column 2 includes some additional explanatory variables such as the share of agriculture and

¹³ These data sets include economies that are open in the sense that capital flows across regional or national boundaries. Thus, evidence on convergence cannot be directly interpreted in the light of the closed economy neoclassical growth model. Barro *et al.* (1995) amend the neoclassical model to allow for *partial* capital mobility. They show that this version of the neoclassical model predicts the same type of transitional dynamics as the strict closed economy version. Hence, it is satisfactory to look at these data sets through the lenses of the closed-economy neoclassical growth model.

¹⁴ The reason for starting in 1950 is that data are available for all 24 OECD economies in 1950. This is not true for the 110 countries that constitute the world data set.

¹⁵ The existence of classical measurement error could deliver a negative relation between growth and the initial level of income. Barro and Sala-i-Martin (1992) show that this is an unlikely explanation for the finding of β -convergence. Measurement error, on the other hand, cannot explain the existence of σ -convergence, unless one argues that the size of the error falls over time.

¹⁶ See Easterlin (1960) and Barro and Sala-i-Martin (1992, 1995, chapter 11).

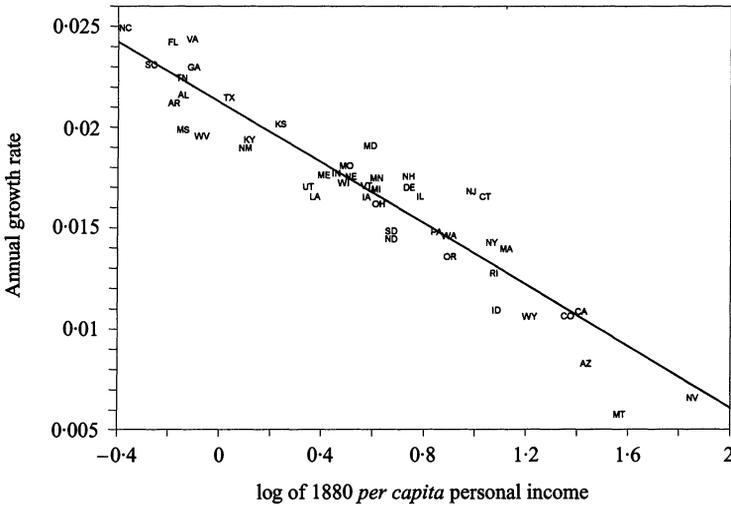


Fig. 4. Convergence of personal income across U.S. states. 1880 personal income and income growth from 1880 to 1990.

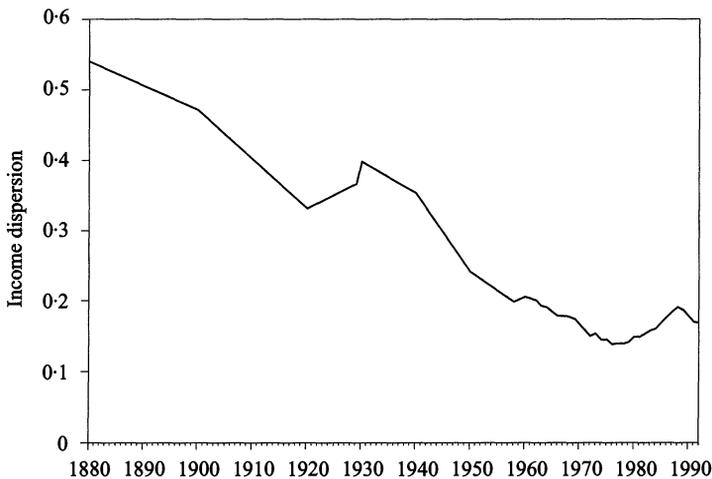


Fig. 5. Dispersion of personal income across U.S. states, 1880–1992.

mining in total income as well as some regional dummies. These variables have little effect on the estimates of β over the long run. The point estimate in this case is 0.017 (s.e. = 0.0026).

The third column reports the estimated β when the overall sample period is divided into sets of 10 year pieces (20 years for 1880 to 1900 and 1900 to 1920). I restrict the estimate of β over time, but allow for time fixed-effects. I also hold constant the shares of income originating in agriculture and industry to proxy for sectoral shocks that affect growth in the short run. The restricted point estimate of β is 0.022 (s.e. = 0.002). Thus, the estimated speed of convergence across states of the United States is similar to that of OECD economies: about 2% per year. Fig. 5 shows an overall downward trend in the cross-sectional

standard deviation for the log of *per capita* personal income net of transfers for 48 U.S. states from 1880 to 1992. The fall in dispersion was interrupted during the 1920s (a period of large declines in agricultural prices which hurt the poorer agricultural states) and the 1980s.

Japanese Prefectures

The fourth row of Table 1 reports similar estimates for 47 Japanese prefectures for the period 1955–87.¹⁷ As for the United States, the first and second columns correspond to a single regression for the entire sample period which, in the case of Japan, is 1955–90. Column 1 estimates β with the initial level of income as the sole explanatory variable. The estimated coefficient is 0.019 (s.e. = 0.003). Column 2 adds some sectoral variables (such as the fraction of income originating in agriculture or industry), but they do not affect the estimated β . The estimates reported in Table 1 use data starting in 1955 because income data by sector are not available before that date. Income data, however, are available from 1930. If we use the 1930 data, the estimated speed of convergence would be 0.027 (s.e. = 0.003). The good fit can be appreciated in Fig. 6. The evidently strong negative correlation between the growth rate from

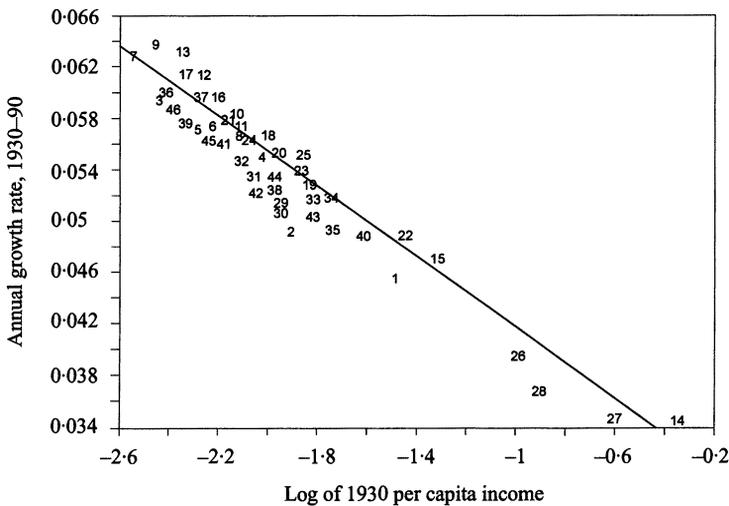


Fig. 6. Convergence of personal income across Japanese prefectures 1930 income and income growth 1930–90.

1930 to 1990 and the log of *per capita* income in 1930 confirms the existence of absolute β -convergence across the Japanese prefectures.

To assess the extent to which there has been σ -convergence across prefectures in Japan, Fig. 7 shows that the dispersion of personal income increased from 0.47 in 1930 to 0.63 in 1940. After 1940, the cross-prefectural dispersion decreased dramatically until 1978 and it increased slightly afterwards. Note

¹⁷ See Barro and Sala-i-Martin (1995, chapter 11) and Shioji (1995).

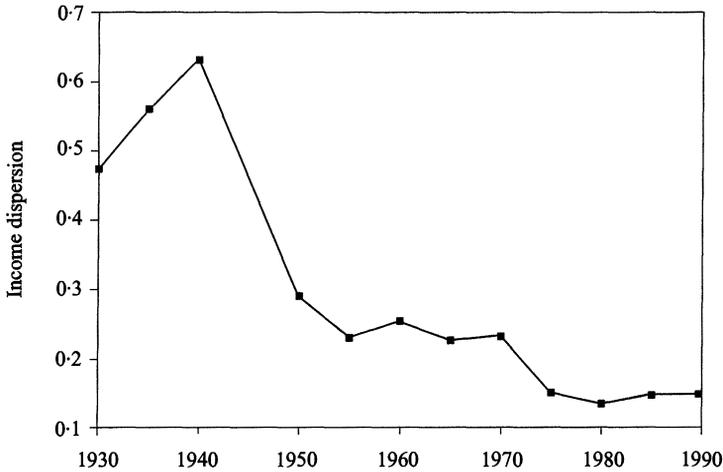


Fig. 7. Dispersion of personal income across Japanese prefectures 1930-90.

that Japan shares with the United States and the set of OECD economies the phenomenon of σ -divergence during a decade that starts somewhere in the mid-1970s.

European Regions

Rows 5 to 10 in Table 1 refer to β -convergence across regions within five European countries (Germany, France, the United Kingdom, Italy, and Spain).¹⁸ The fifth row relates to the estimate of β for a sample of 40 years, 1950-90, when the speeds of convergence are restricted across the 90 regions of the 5 countries and over time. The estimate does, however, allow for country fixed-effects. The estimated β is 0.015 (s.e. = 0.002). The third column shows that the panel estimates of β is 0.031 (s.e. = 0.004). Again, these estimates lie in the neighbourhood of the 2% per year found in the other data sets.

Fig. 8 shows the results on β -convergence visually. The values shown are all measured relative to the means of the respective countries. The figure shows the type of negative relation that is familiar from our earlier discussion of the U.S. states and Japanese prefectures. The correlation between the growth rate and the log of initial *per capita* GDP in Fig. 5 is -0.72. Because the underlying numbers are expressed relative to own-country means, the relation in Fig. 8 pertains to β -convergence within countries rather than between countries, and corresponds to the estimates reported in column 1 of Table 1.

Separate estimates for the long sample for each of the five countries are reported in columns 1 and 2 for the next five rows. The estimates range from 0.010 (s.e. = 0.003) for Italy to 0.030 (s.e. = 0.007) for the United Kingdom. The restricted panel estimates for the individual countries are reported in Column 3. It is interesting to note that the individual point estimates are all close to 0.020 or 2% per year. They range from 0.015 for France to 0.029 for the United Kingdom.

¹⁸ See Barro and Sala-i-Martin (1995, chapter 11). See also Dolado *et al.* (1994) for a careful study of the Spanish case.

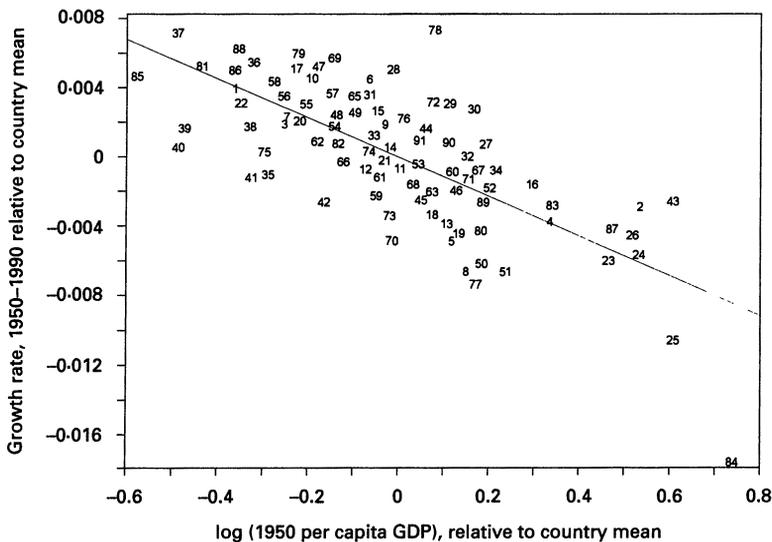


Fig. 8. Growth rate from 1950 to 1990 versus 1950 per capita GDP for 90 regions in Europe.

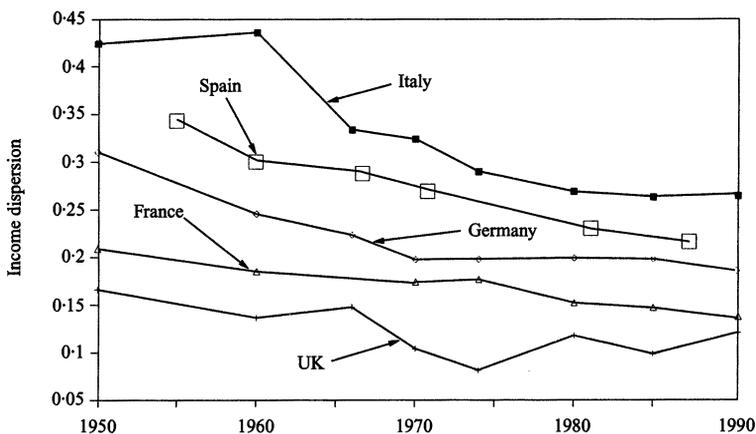


Fig. 9. Dispersion of GDP *per capita* within five European countries.

Fig. 9 shows the behaviour of σ_t for the regions within each country. The countries are always ranked, from highest to lowest, as Italy, Spain, Germany, France, and the United Kingdom. The overall pattern shows declines in σ_t over time for each country, although little net change has occurred since 1970 for Germany and the United Kingdom. In particular, the rise in σ_t from 1974 to 1980 for the United Kingdom – the only oil producer in the European sample – likely reflects the effects of oil shocks.

Other Countries

The empirical research on convergence across regions within a country is now substantial. The main conclusion of most of the studies is that there is regional convergence, and that the speed of convergence is close to 2% per year (some

countries display faster and some countries display slower speed of convergence, but it always lies close to the 2%). Among others, the countries studied are Canada (see Coulombe and Lee (1993)), Australia (see Cashin (1995)), India (see Cashin and Sahay (1995)), Sweden (see Persson (1994)), Austria and Germany (see Keller (1994)).

VI. CONCLUSIONS

There are four main lessons to be gained from the classical approach to convergence analysis. First, the cross-country distribution of world GDP between 1960 and 1990 did not shrink, and poor countries have not grown faster than rich ones. Using the classical terminology, in our world there is no σ -convergence and there is no *absolute* β -convergence. Secondly, holding constant variables that could proxy for the steady state of the various economies, the same sample of 110 economies displays a negative partial correlation between growth and the initial level of GDP, a phenomenon called *conditional* β -convergence. The estimated speed of conditional convergence is close to 2% per year. Thirdly, the sample of OECD economies converge in an absolute sense at a speed which is also close to 2% per year. The sample of countries displays σ -convergence over the same period. However, the process of σ -convergence did seem to stop for about a decade somewhere in the mid-1970s. Fourth, the regions within the United States, Japan, Germany, the United Kingdom, France, Italy, Spain, and other countries display absolute and conditional β -convergence, as well as σ -convergence. Interestingly, the estimated speed of convergence is, in all cases, close to 2% per year. As for the OECD economies, within most of these countries the process of σ -convergence also seemed to stop for about a decade somewhere in the mid-1970s.

I would like to finish this paper with four thoughts about these results. First, we have seen that something strange happened in the mid-1970s all over the world: the process of σ -convergence in most data sets that displayed σ -convergence, stopped for about a decade. In other words, income inequality within the countries studied, increased for a while. Secondly, the speed of convergence, β , has been estimated to be within a narrow range centring on 2% per year ($\beta = 0.02$). Although this is a very robust and strongly significant finding, I would like to emphasise that a speed of 2% per year is very small. For example, it suggests that it will take 35 years for half of the distance between the initial level of income and the steady state level to vanish. This is quite slow. Thirdly, the estimate of $\beta = 0.02$ and equation (4) can be used to provide estimates of the relevant capital share, α . If we let $x = 0.02$ (the rate of productivity growth must be equal to the long-run growth rate of an economy, which is close to 0.02), $n = 0.01$ (the estimated rate of population growth in recent decades), and $\delta = 0.05$ (this rate of depreciation is more controversial; 0.05 corresponds to the rate of depreciation for the overall stock of structures and equipment for the United States), then the capital share implied by the estimated $\beta = 0.02$ is $\alpha = 0.75$. This capital share is larger than the traditional $\alpha = 0.30$ estimated under the assumptions of perfectly

competitive economies with no externalities and no human capital. A value of $\alpha = 0.75$ suggests that, even though the neoclassical model is qualitatively consistent with the data, from a quantitative point of view, it tends to predict too high a speed of conditional convergence. For the model to be consistent with the slow speed of 2% per year, it needs to be amended so that the relevant capital share is larger. Finally, in this paper I followed the classical convergence literature and analysed the empirical results in the light of the neoclassical model. The reason is that, as I said in the text, early theorists of endogenous growth proposed the absence of absolute β -convergence as the main evidence in favour of their models and against neoclassical growth. The introduction of the concept of conditional convergence showed that the neoclassical model is consistent with the data, so it can be a useful framework to guide the convergence literature. However, this is not to say that no other models may be consistent with the existence of convergence. As an example, it can be shown that a model of endogenous growth and technological diffusion (which, in Quah's terminology, make the distinction between growth and convergence effects) can predict an equation exactly like (4) (see Barro and Sala-i-Martin (1995, chapter 8)).

Yale University and Universitat Pompeu Fabra

REFERENCES

- Barro, Robert J. (1991). 'Economic growth in a cross section of countries.' *Quarterly Journal of Economics*, vol. 106, no. 2 (May), pp. 407-43.
- Barro, Robert J., Mankiw, N. Gregory and Sala-i-Martin, Xavier (1995). 'Capital mobility in neoclassical models of growth.' *American Economic Review*, vol. 85, no. 5, pp. 103-15.
- Barro, Robert J. and Sala-i-Martin, Xavier (1992). 'Convergence.' *Journal of Political Economy*, vol. 100, no. 2 (April), pp. 223-51.
- Barro, Robert J. and Sala-i-Martin, Xavier (1995). *Economic Growth*. Boston MA: McGraw Hill.
- Baumol, William J. (1986). 'Productivity growth, convergence, and welfare: what the long-run data show.' *American Economic Review*, vol. 76, no. 5 (December), pp. 1072-85.
- Baumol, William J. and Wolff, E. (1986). 'Productivity growth, convergence, and welfare: reply.' *American Economic Review*, vol. 78, no. 5 (December), pp. 1195-9.
- Cashin, Paul (1995). 'Economic growth and convergence across the seven colonies of Australasia: 1861-1991.' *Economic Record*, vol. 71, pp. 128-40.
- Cashin, Paul and Sahay, Ratna (1995). 'Internal migration, center-state grants, and economic growth in the states of India.' *International Monetary Fund Working Paper 95/75*, Washington, DC.
- Cass, David (1965). 'Optimum growth in an aggregative model of capital accumulation.' *Review of Economic Studies*, vol. 32, no. 91 (July), pp. 233-40.
- Coulombe, S. and Lee, F. C. (1993). 'Regional economic disparities in Canada.' Unpublished paper, University of Ottawa, July.
- Dali, Salvador (1965). 'The critical-paranoic method and the diary of a genius.' *Departament de Publicacions de la Generalitat de Catalunya* (translated to English by Richard Howard).
- DeLong, J. Bradford (1988). 'Productivity growth, convergence, and welfare: comment.' *American Economic Review*, vol. 78, no. 5 (December), pp. 1138-54.
- Dolado, Juan, Gonzalez-Paramo, J. M. and Roldan, J. M. (1994). 'Convergencia economica entre provincias Españolas.' *Moneda y Credito*, no. 1998.
- Dowrick, Steve and Nguyen, Duc Tho (1989). 'OECD comparative economic growth 1950-85: catch-up and convergence.' *American Economic Review*, vol. 79, no. 5 (December), pp. 1010-30.
- Easterlin, Richard A. (1960). 'Regional growth of income: long-run tendencies.' In *Population Redistribution and Economic Growth, United States, 1870-1950*. (eds. Simon Kuznets, Ann Ratner Miller, and Richard A. Easterlin). II: *Analyses of Economic Change*. Philadelphia: The American Philosophical Society.
- Keller, Wolfgang (1994). 'On the relevance of conditional convergence under diverging growth paths. The case of east and west German regions, 1955-1988.' Mimeographed, Yale University, November.

- Koopmans, Tjalling C. (1965). 'On the concept of optimal economic growth.' In *The Econometric Approach to Development Planning*, Amsterdam: North Holland.
- Levine, Ross and Renelt, David (1992). 'A sensitivity analysis of cross-country growth regressions.' *American Economic Review*, vol. 82, no. 4 (September), pp. 942-63.
- Maddison, Angus (1991). *Dynamic Forces in Capitalist Development*. Oxford: Oxford University Press.
- Mankiw, N. Gregory, Romer, David and Weil, David N. (1992). 'A contribution to the empirics of economic growth.' *Quarterly Journal of Economics*, vol. 107, no. 2 (May), pp. 407-37.
- Quah, Danny (1993). 'Galton's fallacy and tests of the convergence hypothesis.' *Scandinavian Journal of Economics*, vol. 95, no. 4, pp. 427-43.
- Quah, Danny (1994). 'Empirics for economic growth and convergence.' Unpublished manuscript, London School of Economics, September.
- Quah, Danny (1996). 'Twin Peaks: growth and convergence in models of distribution dynamics.' *ECONOMIC JOURNAL* (this issue).
- Persson, Joakim (1994). 'Convergence in per capita income and migration across the Swedish counties 1906-1990' (mimeo). University of Stockholm.
- Ramsey, Frank P. (1928). 'A mathematical theory of saving.' *ECONOMIC JOURNAL*, vol. 38, December, pp. 543-59.
- Rebelo, Sergio (1991). 'Long run policy analysis and long run growth.' *Journal of Political Economy*, October, vol. 94, pp. 1002-37.
- Romer, Paul (1986). 'Increasing returns and long run growth.' *Journal of Political Economy*, June, vol. 99, pp. 500-21.
- Sala-i-Martin, Xavier (1990). 'On growth and states' Ph.D. Dissertation, Harvard University.
- Shioji, Etsuro (1995). 'Regional growth in Japan.' Working paper no. 138. Universitat Pompeu Fabra, October.
- Solow, Robert M. (1956). 'A contribution to the theory of economic growth.' *Quarterly Journal of Economics*, vol. 70, pp. 65-94.
- Summers, Robert and Heston, Alan (1991). 'The Penn World Table (Mark 5): an expanded set of international comparisons, 1950-1988.' *Quarterly Journal of Economics*, vol. 106, no. 2 (May), pp. 327-68.
- Swan, Trevor W. (1956). 'Economic growth and capital accumulation.' *Economic Record*, vol. 32, no. 63 (November), pp. 334-61.