

ECONOMIC GROWTH CENTER

YALE UNIVERSITY

P.O. Box 208269
27 Hillhouse Avenue
New Haven, CT 06520-8269

CENTER DISCUSSION PAPER NO. 734

THE CLASSICAL APPROACH TO CONVERGENCE ANALYSIS

Xavier Sala-i-Martin

Yale University

June 1995

Note: Center Discussion Papers are preliminary materials circulated to stimulate discussions and critical comments.

I thank Michelle Connolly and Haanat dePels for their useful comments.

The Classical Approach to Convergence Analysis

Abstract: 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, the 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 two percent per year.

Key Words: Convergence, Regional Economic Growth, Neoclassical Growth, Endogenous Growth.

JEL Classification: O40, O41, O51, O52, O53.

1. Introduction

The existence of convergence across economic units is an important economic question. Labor and public finance economists want to know whether relatively poor families will remain poor for many generations and whether the dynasties that will be rich in a hundred years are the same ones that are rich today. They also want to know whether the degree of income inequality across families increases or falls over time. The reason for finding these questions interesting are obvious to anyone interested in general welfare and to policy-makers who want to engage in redistributive policies and efforts to achieve social peace.

Macroeconomists and theorists of economic growth are interested in exactly the same questions. For them, however, the relevant unit of analysis is not the family but rather the country or the region within a country. For example, they want to know whether, in our world, rich countries will remain rich and poor countries will remain poor for many decades. They are also interested in knowing whether the distribution of world income and output across countries is becoming increasingly equal over time.

These important questions lie at the heart of *the convergence debate*. 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 way to test 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 4). For this reason, growth theorists started paying close attention to the evolution of the convergence debate. Second, and perhaps more importantly, the data set on internationally comparable GDP levels for a large number of countries became ready for use in the mid 1980s¹. This new data set allowed empirical economists to compare GDP levels across a large number of countries economies 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 music or classical paintings, it is the basis of reference and target of criticism of all other methodologies. And also like classical art, it has survived and will keep surviving the challenges of more modern and “surrealist” movements.

The rest of the paper is organized as follows. Section 2 introduces two useful definitions of convergence and highlights their similarities and differences. Section 3 analyzes some evidence on convergence using a sample of 110 economies. In section 4 I interpret the above evidence in light of the neoclassical model and introduce the concept of conditional convergence. Section 5 applies the concept of conditional convergence to the sample of 110 countries. Section 6 provides evidence on convergence for number of regional data sets. I conclude in section 7.

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].)

2. Definitions of Convergence.

Two main concepts of convergence appear in the classical literature. They are called β -convergence and σ -convergence.² We say that *there is absolute β -convergence if poor economies tend to grow faster than rich ones.*³ Imagine that we have data on real per capita GDP for a cross-section of economies between years t and $t+T$. If we estimate the following regression

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

where $\gamma_{i,t,t+T} \equiv \log(y_{i,t+T}/y_{i,t})/T$ is economy i 's growth rate of GDP between t and $t+T$,

and $\log(y_{i,t})$ is the logarithm of economy i 's GDP per capita at time t 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

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

where σ_t is the time t standard deviation of $\log(y_{i,t})$ across i .⁴ The concepts of σ - and absolute

² This terminology was first introduced in Sala-i-Martin [1990].

³ In a later section I will introduce the important concept of conditional β -convergence.

⁴ The standard deviation of the *logarithm* of GDP per capita is invariant with the mean. In this regard, it is similar to the *coefficient of variation* of the *level* of GDP per capita, which is

β -convergence are, of course, related. If we take the sample variance of $\log(y_{i,t})$ from (1), 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 grows faster. As an illustration, Figure 1 displays the behavior of the log of GDP per capita for two economies over time. Economy A starts 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 than the growth rate of economy B between times t and $t+T$ and, therefore, we say that there is β -convergence. Since dispersion 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 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 a 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 Figure 1 is an example where β -convergence exists and is associated with σ -convergence. Panel B provides an example where the lack of β -convergence (the 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 the theoretical level, it is possible for poor countries to grow faster than rich ones, without observing that the cross-sectional dispersion fall over time. That is, we could in principle find β -convergence without

equal to the standard deviation of the level divided by the mean.

finding σ -convergence. In Panel C, for example, I have constructed an example where the poor economy grows faster 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 the fact of 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 do not always show up together is that they capture two different aspects of the world. σ -convergence relates to whether the *cross-country distribution of world income shrinks over time* or not. β -convergence, on the other hand, relates to the *mobility of the different economies within the given distribution of world income*. Panels A and B are examples where the movements of the various economies over time changes 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

⁵ This possibility led some economists (most prominently Quah [1993]) to criticize the classical approach on three grounds. First they suggested that the classical analysts were confusing the two concepts of convergence. Second, they suggested that the only meaningful concept of convergence was that of σ . And 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. The three points were wrong. 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's why they made the distinction in the first place! Second, the intra-distributional mobility (reflected in β) is at least as interesting as the behavior of the distribution itself (reflected in σ). Surprisingly, Quah [1994] highlights the importance of intra-distributional mobility in the context of stochastic Kernel estimators. And finally, β provides

Having made the theoretical distinction between σ and β convergence, I would like to mention at this point that, in practice, this distinction is not as important. The reason is that, when it comes to real world data, whenever we observe σ -convergence, we also observe β -convergence. However, since we can only say this after we have analyzed the data, throughout the rest of the paper I will follow the classical literature, and I will analyze the two concepts of convergence for every data set.

3. Cross-Country Evidence

Maddison [1991] provided data on GDP levels across a cross-section of 13 rich countries starting in 1870. These data were constructed following the methodology of the U.N.'s 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 proved fatal for the study 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 industrialized 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 the data set was expanded to include countries that appeared rich ex-ante (that is, by

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.

1870)⁶ the evidence for convergence quickly disappeared.

The solution to the sample selection problem was to analyze a larger set of countries. This is where the newly created Summers-Heston data set came in handy. This data set involved GDP levels for more than one hundred countries. Unlike Maddison's project, however, where the time series dimension of the data was quite large, the first year for which the Summers and Heston data is available is 1960. Hence, by using the Summers-Heston data set analysts could study a broader set of countries, but the cost was a much shorter time span. In Figure 2 I display the behavior of the dispersion of GDP per capita for 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.

Figure 3 analyzes 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 figure shows that the relation between growth and the initial level of GDP is not negative. In fact, the slope of the regression (also shown in the figure) 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 non-linear equation

⁶ This sample would include countries that appeared rich in 1870 but whose performance in the following century was not spectacular (Spain, Argentina, or Ireland would examples of such countries) and would exclude countries that were poor in 1870, but have become surprising successes ex post (like, for example, Japan). See DeLong [1988] and Baumol and Wolff [1988].

$$(3) \quad \gamma_{i,t,t+T} = a - \left(\frac{1-e^{-\beta \cdot T}}{T} \right) \cdot \log(y_{i,t}) + \epsilon_{i,t,t+T} .$$

The reason for estimating (3) instead of the linear version that appears in (1) is that I want to compare the speed of convergence across data sets that have different time lengths. The OLS coefficient in (1) would be inversely related to T (the length of the period over which I compute the growth rate.) The reason is that, if there is convergence, the growth rate should fall over time (because when the economy is richer, the growth rate is predicted to be smaller.) When we average long periods of time, we combine early periods with large growth rates with later periods with small growth rates. Hence, the growth rate predicted by the original (low) level of income is smaller the longer the time period of analysis. Note that as T goes to infinity, the coefficient on the linear regression (1) goes to zero. As T goes to zero (that is as the time period of analysis becomes short) the coefficients in the linear and non-linear equations coincide. Since the coefficient β in the non-linear regression is invariant to the length of the sample, I will estimate the non-linear equation 1 from here on.

A second reason for estimating the non-linear equation (3) instead of (1) is that, as I will argue in the next section, the convergence hypothesis has been discussed in the context of models of growth. The log-linearization of neoclassical model around its steady state yields an equation like (3). In that context, the parameter β can be interpreted as the speed of instantaneous convergence of an economy towards its steady-state position. I will therefore estimate (3) and I will call β the speed of convergence.

Table 1 reports the estimated speed of convergence, β , for a variety of data sets under three different setups. 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 Figure 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 .*

4. 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 does not tend to disappear, but rather tends to increase over time. It also suggests that the countries that are predicted to be richer a few decades from now are the same countries that are rich today. These findings may be used by economists or politicians to devise international institutions of cooperation or help that tend to overturn this somber tendency.

These findings were also seen by growth theorists in the middle of the 1980s as evidence against the neoclassical model of Ramsey [1928], Solow [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 the growth rate of capital) 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 poor countries with little amounts of capital will be poor and will grow faster than rich countries with large amounts of capital, so there will be cross-country β -convergence. Since the model does not predict the type of overshooting displayed by the economies in Figure 1's Panel C, the prediction of β -convergence will tend to be associated with a reduction of cross-economy dispersion over time, i.e σ -convergence.

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

$$Y_{i,t} = A_i \cdot K_{i,t}^\alpha \cdot L_{i,t}^{1-\alpha}, \text{ where } Y_{i,t} \text{ is economy } i\text{'s aggregate output at time } t, K_{i,t} \text{ and } L_{i,t} \text{ are the}$$

stock of capital and labor in that economy respectively, and A_i is the level of technology.

Following Solow [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 characterizes the behavior of economy i over time says that capital accumulation is the difference between overall savings and effective depreciation. If we log-linearize 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, the parameter β is exactly equal to

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

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 convergence property. In terms of equation (4), the one sector models of endogenous growth are similar to the neoclassical model except that 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 favor of the new models of endogenous growth (see for example, Romer [1986] and Rebelo [1990].)⁹

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 optimizing version of the neoclassical model.

⁸ Labor, which is not purposely accumulated in the neoclassical model is often substituted with human capital, whose stock 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 closely. 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

The Absolute Convergence Fallacy.

The argument that says that the neoclassical model predicts convergence relies heavily on the key assumption that *the only difference across countries was their initial levels of capital*. In the real world, however, economies may differ in other things such as their levels of A_i or their propensities to save. If different economies have different parameters, then they will have different steady states 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 the steady state. Hence, the prediction of the neoclassical model is that the growth rate of an economy will be negatively 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*

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 labor 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 .

¹⁰ See Sala-i-Martin [1990], Barro and Sala-i-Martin [1992] and Mankiw, Romer and Weil [1992].

(the concept of β -convergence discussed above is sometimes called absolute convergence to distinguish it from its new conditional counterpart.) Only if all the economies converge to the same steady state does the prediction that poor economies should grow faster than rich ones hold true. The reason is that, in that case, poor economies will be unambiguously farther away from their steady state. Put in another way, only if all the economies have the same steady state, do the *conditional convergence* and the *absolute convergence* hypotheses coincide. 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. Classical analysts have tried to hold the steady state constant in two different ways. The first one is the introduction of 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

$$(5) \quad \gamma_{i,t,t+T} = \alpha - \left(\frac{1 - e^{-\beta T}}{T} \right) \cdot \log(y_{i,t}) + \psi \cdot X_{i,t} + \epsilon_{i,t,t+T} \quad ,$$

where $X_{i,t}$ is a vector of variables that hold constant the steady state. If the estimate of β is positive, once $X_{i,t}$ is held constant, then we say that the data set exhibits *conditional β -convergence*. In section 5 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 state is not unrealistic. For example, because we think that the technology, institutions, and tastes of the African economies are very different from those of Japan or the United States, the assumption that these economies converge

to the same steady state is not realistic. However, the technological and institutional differences across regions within a country or across “similar” countries (like, for example, those of the OECD) are probably smaller. Hence, we may want to look for *absolute convergence* within these set of “more similar” economies. This second approach is used in section 6.

5. 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 X successfully holds constant the steady state, we should find 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 versions of the neoclassical model suggest different variables. The strict version of the Solow model, for example, says that 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 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 X .¹² For

¹¹ See for example, Mankiw, Romer and Weil [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 analyzed in this paper. See Levine and Renelt [1992] for some

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 enrollments, the saving rate, and some political variables are used as the vector X . Note that, unlike column 1, the estimate of β is now positive and significant, $\beta=0.013$ (s.e.=0.004). Column 3 divides the sample period 1960-1990 into two subperiods and estimates β by restricting it to be the same across subperiods. The estimated β is 0.025 (s.e.=0.003).

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 percent per year. I should emphasize, 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 X , and the growth rate falls as the economy approaches this long-run level.

6. Conditional Convergence (II): Regional Evidence.

The second method for "holding constant the steady state" is to analyze sets of economies that appear similar to the researcher so that the assumption of the same steady state is reasonable. For example, OECD economies, and 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

evidence on this point.

as σ -convergence. If evidence of absolute β -convergence is to be found anywhere, it will be in these data sets.¹³

OECD Economies

Figures 2 and 3 also display the convergence behavior of a subset of the world sample: the OECD countries. In Figure 2, I plot the cross-sectional dispersion of GDP per capita for OECD economies starting in 1950.¹⁴ In 1950, the dispersion was equal to $\sigma=0.60$. By 1960, the coefficient of dispersion for the OECD countries, $\sigma=0.51$, was much smaller than that of the world, $\sigma=0.89$. Contrary to what happened to the broad cross-section of countries, OECD dispersion falls steadily between 1950 and 1975. At that point, it reaches a value of $\sigma_{1975}=0.37$. Between 1975 and 1985, the dispersion of GDP across OECD economies increases slightly ($\sigma_{1980}=0.38$ and $\sigma_{1985}=0.39$). After 1985, the process of σ -convergence resumes so and the dispersion in 1990 becomes lower than it was in 1975, $\sigma_{1990}=0.36$. The conclusion is that the sample of OECD economies has converged in the sense of σ . Figure 3 also highlights the differential behavior 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 Figure 3 by the downward-sloping regression line. OECD economies have thus also

¹³ These data sets include economies that are open in the sense that capital flows across economies within the data set. Thus, evidence on convergence cannot be directly interpreted in the light of the closed economy neoclassical growth model. However, Barro, Mankiw and Sala-i-Martin [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 dynamics as the strict closed economy version.

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.

converged in the sense of β .¹⁵ 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 around 2 per cent per year.¹⁶

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 1990.¹⁷ 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

¹⁵ The OECD was founded in 1961 and its original membership included 20 countries. Today, the OECD has 24 members. Australia, Finland, Japan and New Zealand joined the Organization later on, and one could argue that they did so because they had become rich ex-post. Hence, one could argue that, to some extent, the sample of OECD economies is subject to the sample selection bias critique used by Romer [1986] and DeLong [1988] against Baumol [1986]. The elimination of these four countries from the sample, however, does not change the main conclusion: OECD countries have converged both in the sense of σ and in the sense of β .

¹⁶ 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.

¹⁷ See Easterlin [1960], Barro and Sala-i-Martin [1992, 1995, chapter 11].

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 Figure 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 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 in the United States is similar to that of OECD economies: about 2 percent per year.

Figure 5 shows 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. We observe that the dispersion declined from 0.54 in 1880 to 0.33 in 1920, but then rose to 0.40 in 1930. This rise reflects the adverse shock to agriculture during the 1920s: the agricultural states were relatively poor in 1920 and suffered a further reduction in income due to the fall in agricultural prices. After the 1920s shock, dispersion fell to 0.35 in 1940, 0.24 in 1950, 0.21 in 1960, 0.17 in 1970, and reached a low point of 0.14 in 1976. The long-run decline stopped in the mid-1970s, after the oil shock, and σ_t rose to 0.15 in 1980 and 0.19 in 1988. The rise in income dispersion was reversed in the last two years of the 1980s and continued to fall through 1992.

Japanese Prefectures.

The fourth row of Table 1 reports similar estimates for 47 Japanese prefectures for the period 1955-1987.¹⁸ 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 1955-1990. 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 Figure 6. The evidently strong negative correlation between the growth rate from 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, I compute σ_t for the 47 prefectures from 1930 to 1990. Figure 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 after: it fell to 0.29 by 1950, to 0.25 in 1960, to 0.23 in 1970 and hit a minimum of 0.125 in 1978. It increased slightly afterwards: σ_t rose to 0.13 in 1980, 0.14 in 1985 and 0.15 in 1987. Income dispersion has been relatively constant since then. Note 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.

¹⁸ See Barro and Sala-i-Martin [1995, chapter 11] and Shioji [1994].

European Regions

Rows 5 through 10 in Table 1 refer to β -convergence across regions within five European countries (Germany, France, the United Kingdom, Italy, and Spain).¹⁹ The fifth row relates to the estimate of β for a sample of 40 years, 1950-1990, when the speeds of convergence are restricted across the 90 regions 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 neighborhood of the 2 percent per year found in the previous data sets.

Figure 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 the U.S. states and Japanese prefectures. The correlation between the growth rate and the log of initial per capita GDP in Figure 5 is -0.72. Because the underlying numbers are expressed relative to own-country means, the relation in Figure 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 UK. 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 two percent per year. They range from 0.015 for France to 0.029 for the United

¹⁹ See Barro and Sala-i-Martin [1995, chapter 11]. See also Dolado, Gonzalez-Paramo and Roldan [1994] for a careful study of the Spanish case.

Kingdom.

Figure 9 shows the behavior 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. In 1990, the values of σ_t are 0.27 for Italy, 0.22 for Spain (this value corresponds to 1987), 0.186 for Germany, 0.139 for France, and 0.122 for the United Kingdom.

Other Countries

The evidence on convergence across regions within a country has been substantial. The main point of most of the studies is that there is regional convergence in almost any country analyzed and that the speed of convergence is close to 2 percent per year (some countries display faster and some countries display slower speed of convergence, but it always lies close to the two percent.) Among others, the countries studied are Canada (see Coulombe and Lee [1993]), Australia (see Cashin [1995a]), India (see Cashin [1995b]), China (see Rivera-Batiz [1994]), Sweden (see Persson [1994]), Austria, East Germany (see Keller [1994]) and Spain (see Dolado, Gonzalez-Paramo and Roldan [1994].).

7. Conclusions.

There are four main lessons to be gained from the classical convergence literature. First,

the cross-country distribution of world GDP between 1960 and 1990 did not shrink, and poor countries do not grow faster than rich ones. Using the classical terminology, in our world there is no σ -convergence and there is no absolute β -convergence. Second, 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. Interestingly, the estimated speed of conditional convergence is close to 2 percent per year. Third, the sample of OECD economies converge in an absolute sense at a speed which is also close to 2 percent 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. The estimated speed of convergence is, in all cases, close to 2 percent 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. While the evil policies of Ronald Reagan were blamed for this increase in income inequality in the United States, the President cannot be blamed for the existence of the same phenomenon within Japan, within European countries or across OECD economies. A more reasonable theory would point in the

direction of technological shocks that increased the productivity of the high-wage (educated) workers.

Second, the speed of convergence, β , has been estimated to be within a narrow range of two percent per year ($\beta=0.02$). Although this is a very robust and strongly significant finding, I would like to emphasize that a speed of 2 percent 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 to vanish. This is quite slow.

Third, 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 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 percent per year, it needs to be amended so that the relevant capital share is larger.²⁰

Finally, in this paper I followed the classical convergence literature and analyzed the empirical results in the light of the neoclassical model. As I said in the text, early theorists of

²⁰ See Barro and Sala-i-Martin [1992] and Mankiw, Romer and Weil [1992] who amend the neoclassical model to incorporate human capital. This amendment effectively increases the relevant capital share to a number consistent with the estimated speed of convergence.

endogenous growth proposed the absence of absolute β -convergence as the main piece of evidence in favor 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 also. For example, it can be shown that a model of endogenous growth and technological diffusion can predict equation exactly like (4) (see Barro and Sala-i-Martin [1995, chapter 8]) and that the two-sector models of endogenous growth may also be consistent with the empirical evidence (Barro and Sala-i-Martin [1995, chapter 5].)

TABLE 1: Estimates of the Speeds of β -Convergence for a Variety of Data Sets

	Long-Run Single Regression (*)	Long-Run Single Regression (*)	Panel Estimates (*)
DATA SET	(1) Absolute Convergence	(2) Conditional Convergence	(3) Conditional Convergence
World - 110 Countries (1960-1990)	-0.004 0.04 (0.002) [0.0176]	0.013 0.46 (0.004) [0.0134]	0.025 --- (0.0028)
- OECD Countries (1960-1990)	0.014 0.48 (0.003) [0.0062]	0.029 0.78 (0.008) [0.0050]	
United States - 48 States (1880-1990)	0.021 0.89 (0.0003) [0.0015]	0.017 0.89 (0.002) [0.0015]	0.022 --- (0.002)
Japan - 47 Prefectures (1955-1990)	0.019 0.59 (0.003) [0.0027]	0.019 0.59 (0.004) [0.0027]	0.031 --- (0.004)
Europe Total (**) - 90 Regions (1950-1990)	0.015 0.51 (0.002) [0.0030]	0.015 0.52 (0.002) [0.0030]	0.018 --- (0.003)
Germany (11 Regions)	0.014 0.56 (0.006) [0.0028]	0.014 0.55 (0.005) [0.0027]	0.016 --- (0.006)
United Kingdom (11 Regions)	0.020 0.62 (0.008) [0.0021]	0.030 0.61 (0.007) [0.0021]	0.029 --- (0.009)
France (21 regions)	0.016 0.55 (0.005) [0.0023]	0.016 0.55 (0.004) [0.0022]	0.015 --- (0.003)
Italy (20 Regions)	0.010 0.46 (0.003) [0.0033]	0.010 0.46 (0.003) [0.0031]	0.016 --- (0.003)
Spain (17 Regions) (1955-1987)	0.021 0.63 (0.005) [0.0042]	0.023 0.63 (0.007) [0.0040]	0.019 --- (0.005)

Notes to Table 1:

(*) The regressions use non-linear least squares to estimate equations of the form:

$$(1/T)\ln(y_{it}/y_{i,t-T}) = \alpha - [\ln(y_{i,t-T})](1-e^{-\beta T})(1/T) + \text{"other variables"},$$

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 (recall that when the error term is correlated with the explanatory variable, then the OLS estimate of β is biased).

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. Thus, 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 for each country to have its own constant term.

References

- Barro, Robert J. (1991). "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, 106, 2 (May), 407-443.
- Barro, Robert J., N. Gregory Mankiw, and Xavier Sala-i-Martin (1995). "Capital Mobility in Neoclassical Models of Growth," *American Economic Review*, vol. 85, n. 5, 103-115.
- Barro, Robert J. and Xavier Sala-i-Martin (1992). "Convergence," *Journal of Political Economy*, 100, 2 (April), 223-251.
- Barro, Robert J. and Xavier Sala-i-Martin (1995), "*Economic Growth*", McGraw Hill, Boston MA.
- Baumol, William J. (1986). "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show," *American Economic Review*, 76, 5 (December), 1072-1085.
- Baumol, William J. and E. Wolff (1986). "Productivity Growth, Convergence, and Welfare: Reply", *American Economic Review*, 78, 5 (December), 1195-1199.
- Cashin, Paul, (1995a), "Economic Growth and Convergence Across Australian Regions", mimeo *International Monetary Fund*.
- Cashin, Paul, (1995b), "Economic Growth and Convergence Across Indian Regions", mimeo *International Monetary Fund*.
- Cass, David (1965). "Optimum Growth in an Aggregative Model of Capital Accumulation," *Review of Economic Studies*, 32, 91 (July), 233-240.
- DeLong, J. Bradford (1988). "Productivity Growth, Convergence, and Welfare: Comment," *American Economic Review*, 78, 5 (December), 1138-1154.
- Dolado, Juan., J. M. Gonzalez-Paramo, and J. M. Roldan, (1994), "Convergencia Economica entre provincias Españolas", *Moneda y Credito*, 1998.
- Dowrick, Steve and Duc Tho Nguyen (1989). "OECD Comparative Economic Growth 1950-85: Catch-Up and Convergence," *American Economic Review*, 79, 5 (December), 1010-1030.
- Easterlin, Richard A. (1960). "Regional Growth of Income: Long-Run Tendencies," In Simon Kuznets, Ann Ratner Miller, and Richard A. Easterlin, eds., *Population Redistribution and Economic Growth, United States, 1870-1950. 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 1994.
- Koopmans, Tjalling C. (1965). "On the Concept of Optimal Economic Growth," in *The Econometric Approach to Development Planning*, Amsterdam, North Holland, 1965.
- Levine, Ross and David Renelt (1992). "A Sensitivity Analysis of Cross-Country Growth Regressions," *American Economic Review*, 82, 4 (September), 942-963.
- Maddison, Angus (1991). *Dynamic Forces in Capitalist Development*, Oxford, Oxford University Press.
- Quah, Danny (1993). "Galton's Fallacy and Tests of the Convergence Hypothesis," *Scandinavian Journal of Economics*, 95, 4, 427-443.
- Quah, Danny, (1994), "Empirics for Economic Growth and Convergence", unpublished manuscript, London School of Economics, September 1994.
- 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*, 38, December 1928, 543-559.
- Rebelo, Sergio, (1991), "Long Run Policy Analysis and Long Run Growth", *Journal of Political Economy*, October, 94, 1002-37.
- Rivera-Batiz, Luis (1994), "Convergence Across Chinese Provinces", Mimeo Universitat Pompeu Fabra
- Romer, Paul, (1986), "Increasing Returns and Long Run Growth", *Journal of Political Economy*, June, 94, 500-21.
- Sala-i-Martin, Xavier, (1990), "On Growth and States" Ph.D. Dissertation, Harvard University.
- Shioji, Etsuro, (1992), "Regional Growth in Japan", mimeo Yale University.
- Solow, Robert M. (1956), "A Contribution to the Theory of Economic Growth". *Quarterly Journal Of Economics*, 70: 65-94.
- Summers, Robert and Alan Heston (1991). "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988," *Quarterly Journal of Economics*, 106, 2 (May), 327-368.

Figure 1: The Relation Between σ and β -convergence

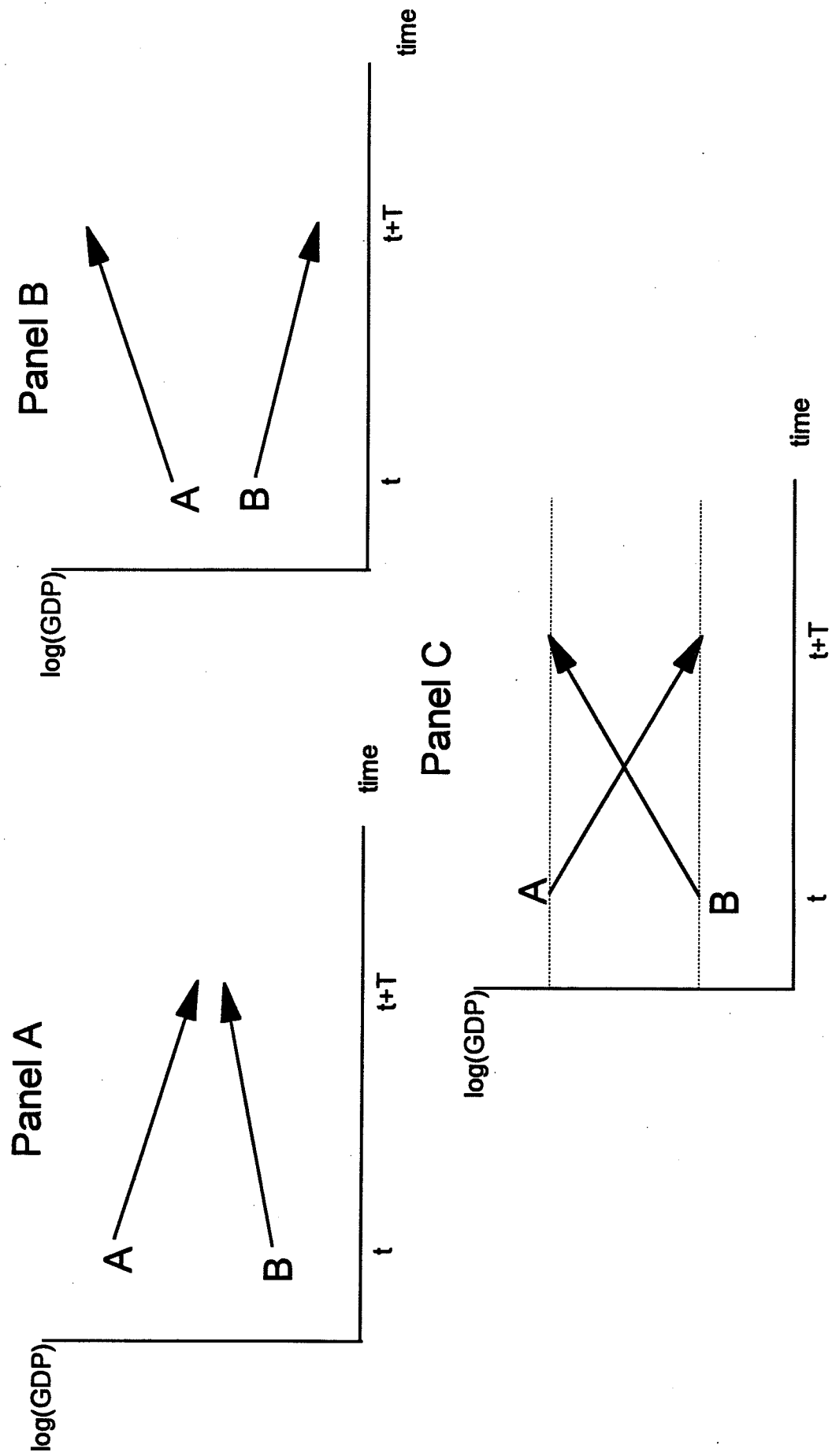


Figure 2: Dispersion of GDP Across 110 Countries

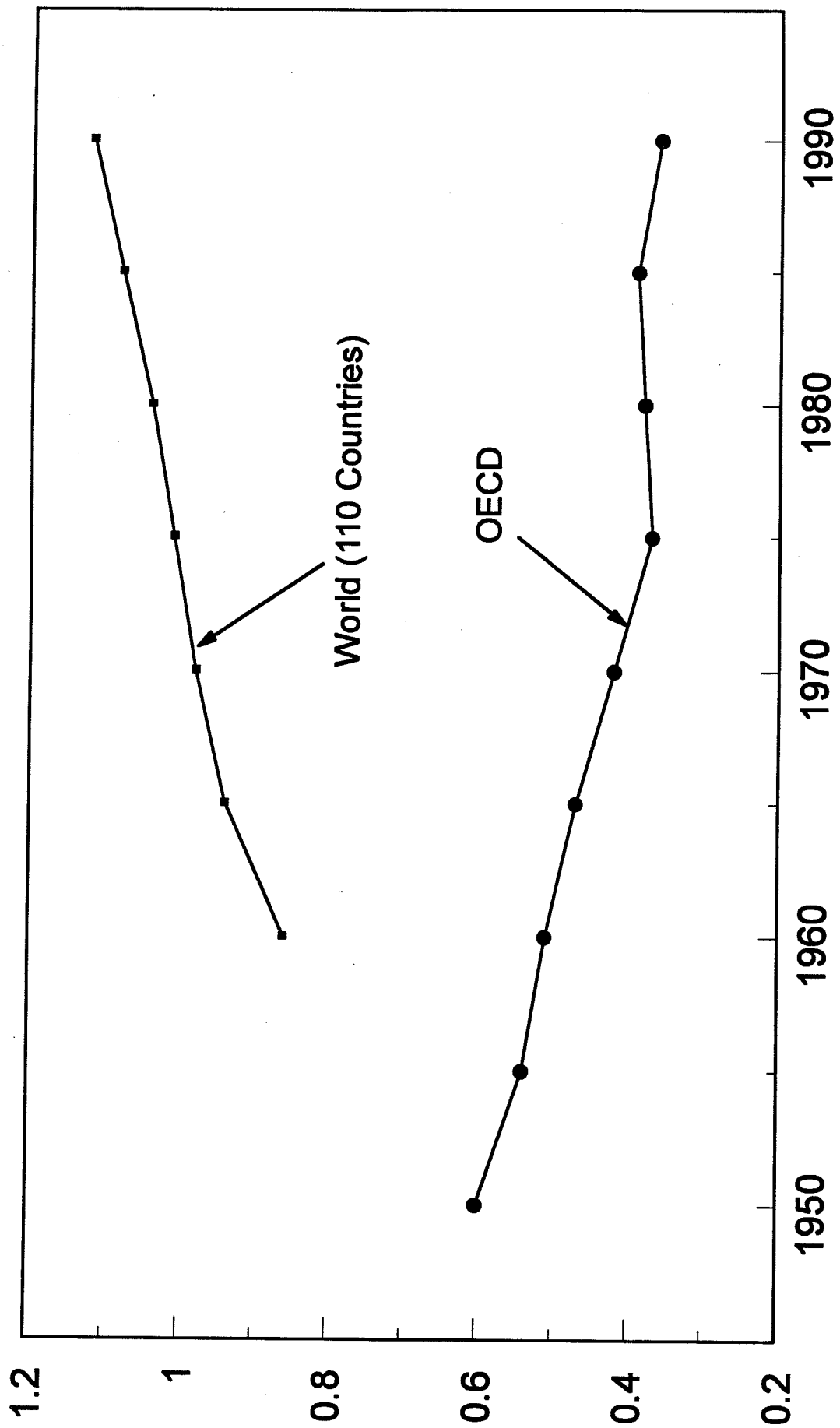


Figure 3: Convergence Across Countries, 1960-1990

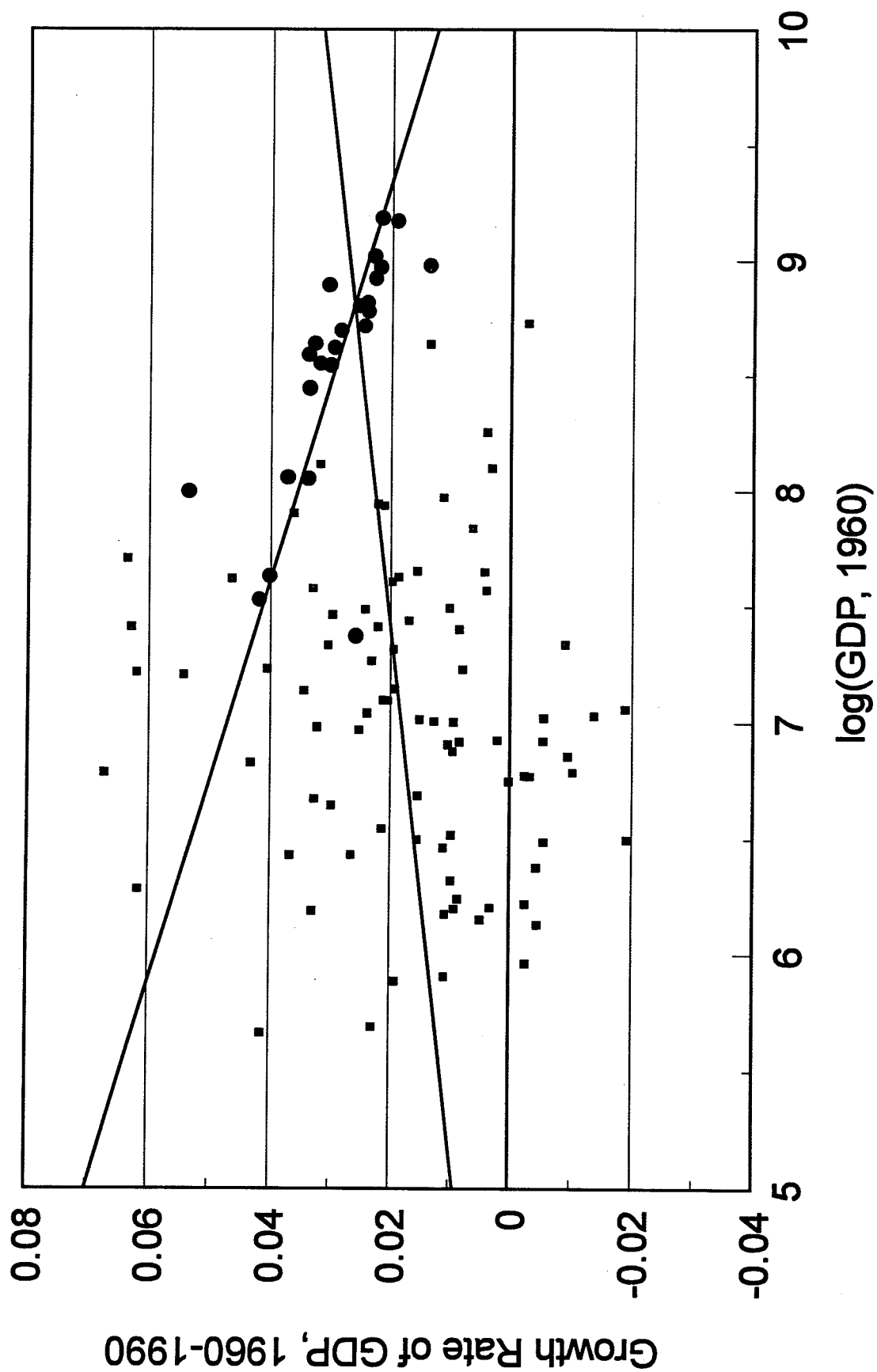


Figure 4: Convergence of Personal Income across U.S. States.
1880 Personal Income and Income Growth from 1880 to 1990

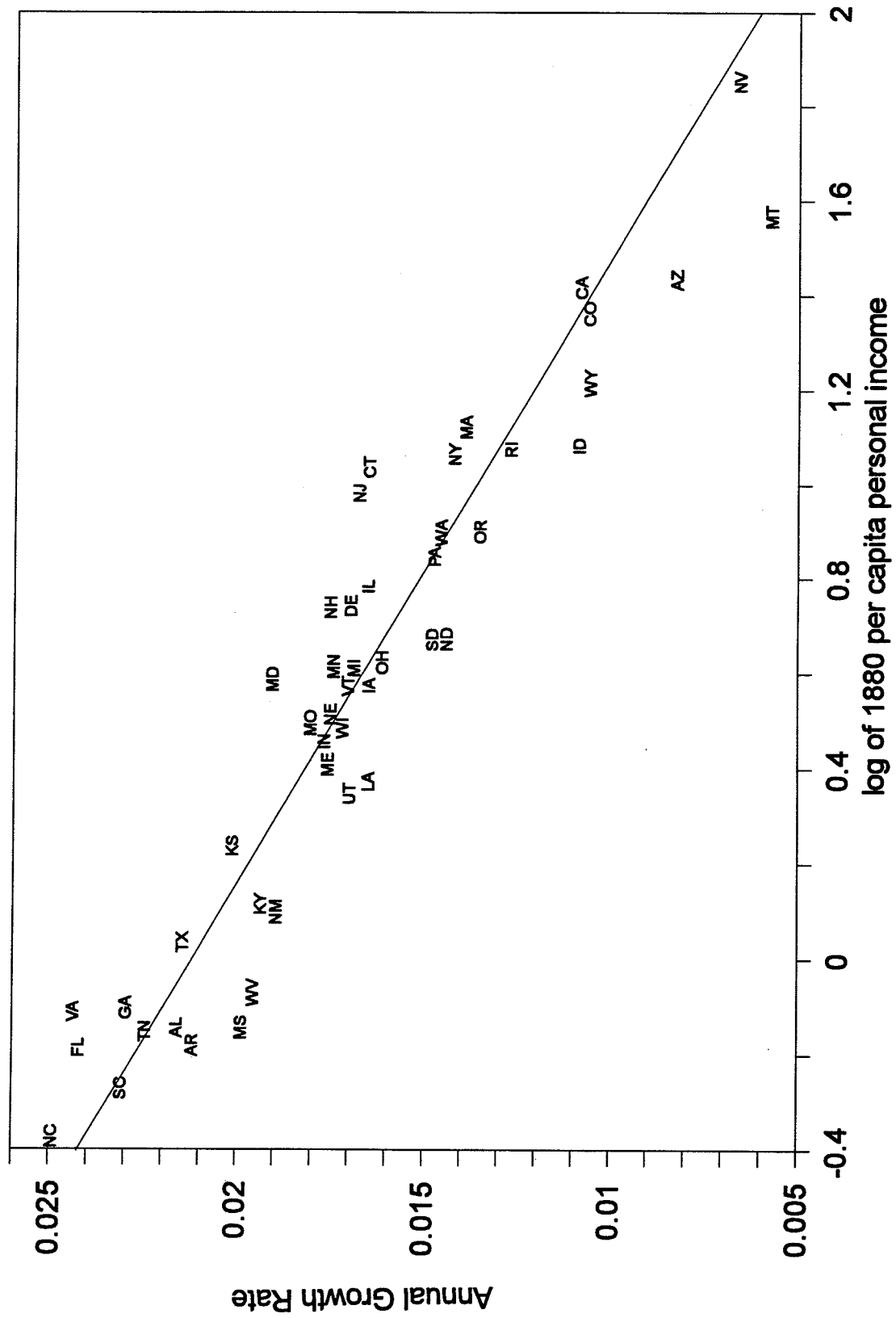


Figure 5:
Dispersion of Personal Income across U.S. States, 1880-1992

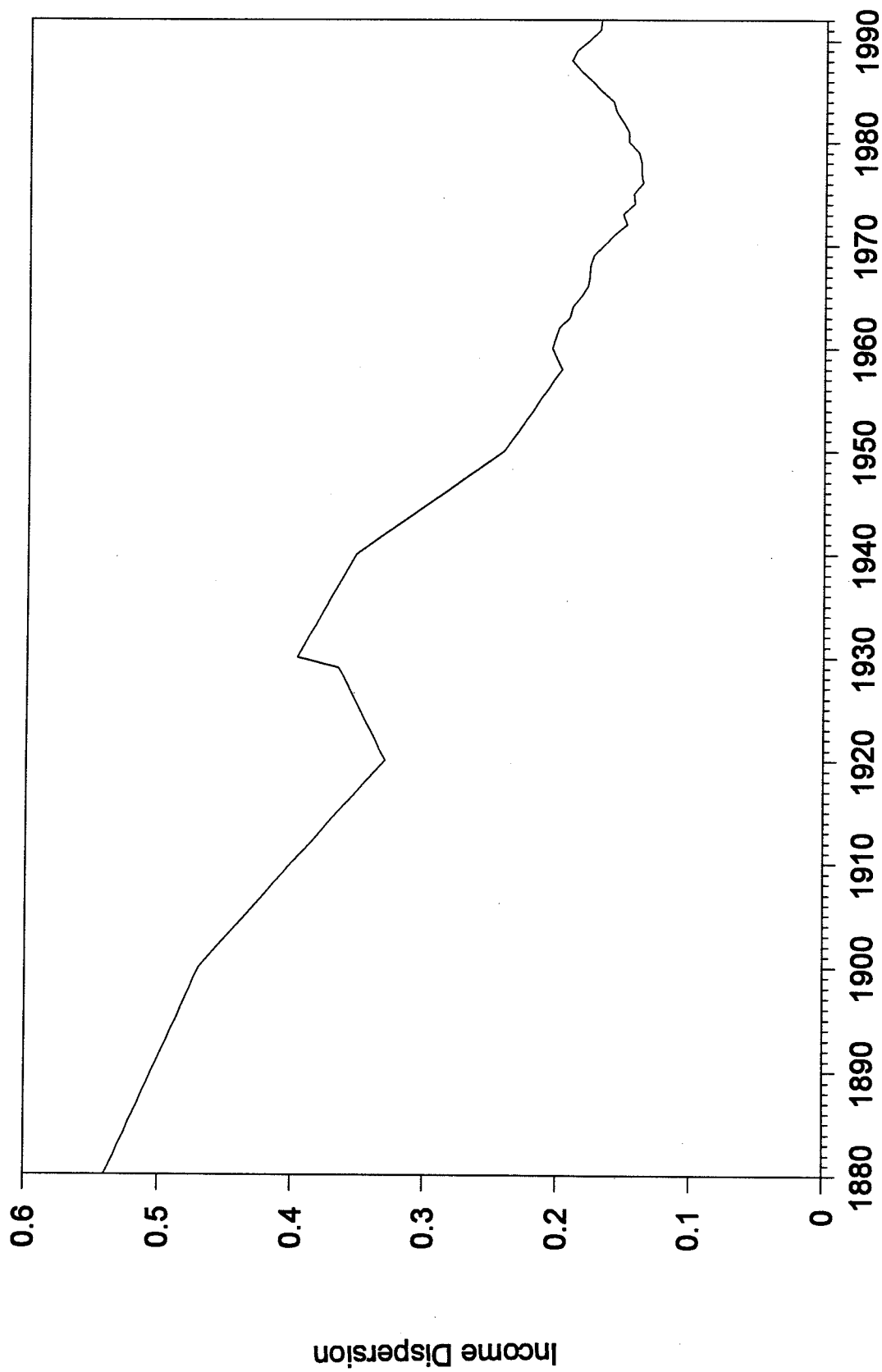


Figure 6:
Convergence of Personal Income Across Japanese Prefectures
1930 Income and Income Growth 1930-1990

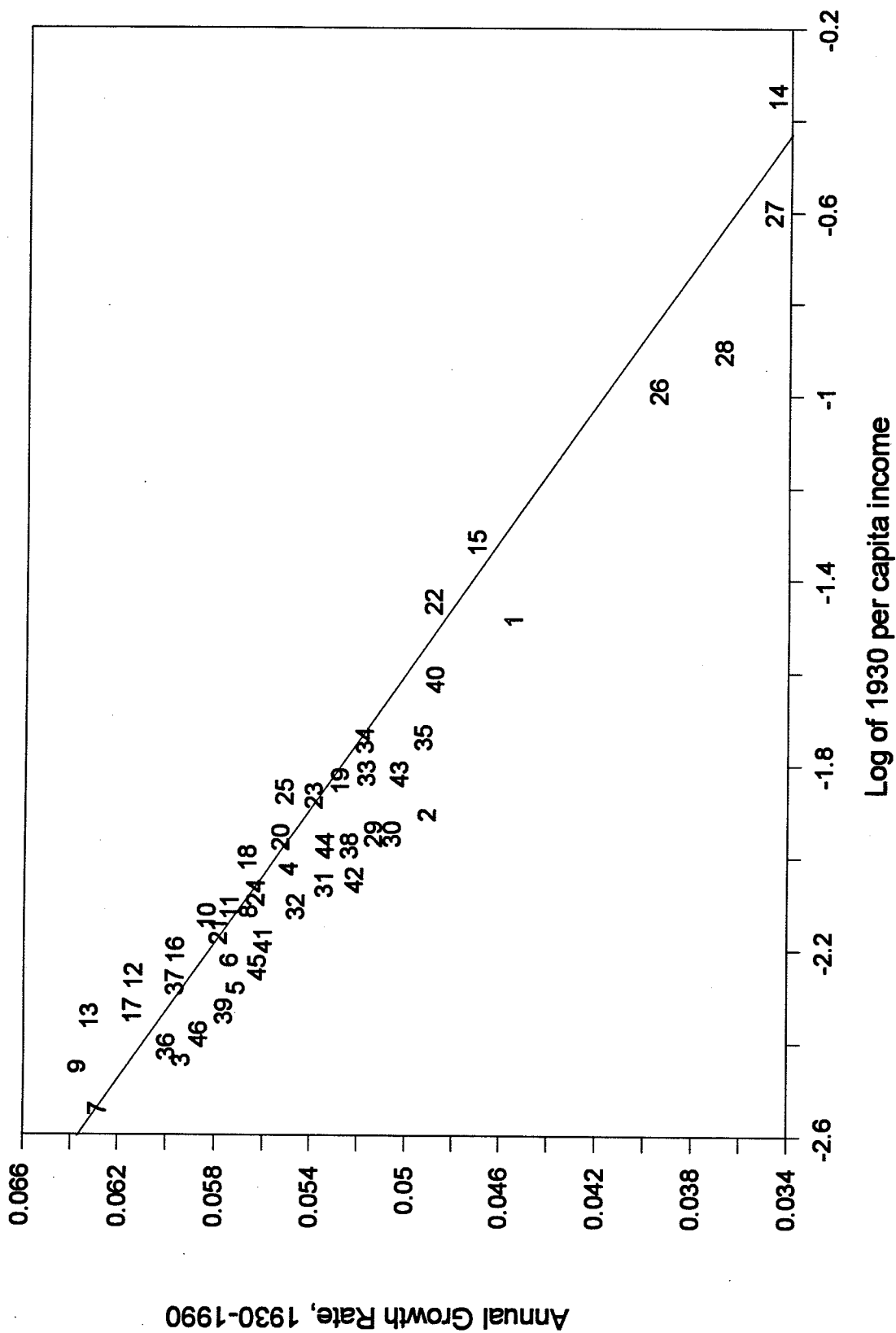


Figure 7:
Dispersion of Personal Income Across Japanese Prefectures
1930-1990

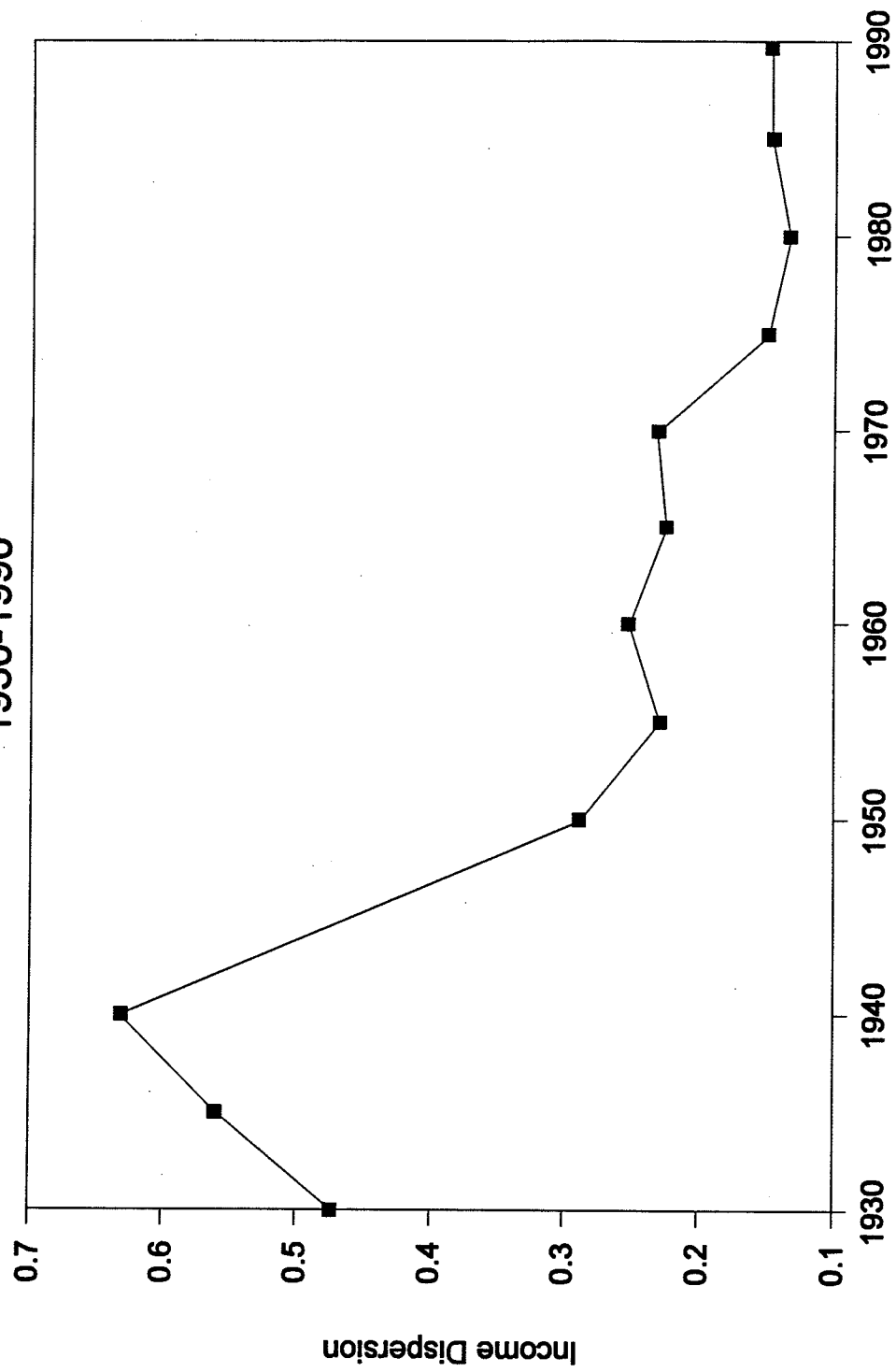


Figure 8: Growth Rate from 1950 to 1990
versus 1950 per capita GDP for 90 Regions in Europe

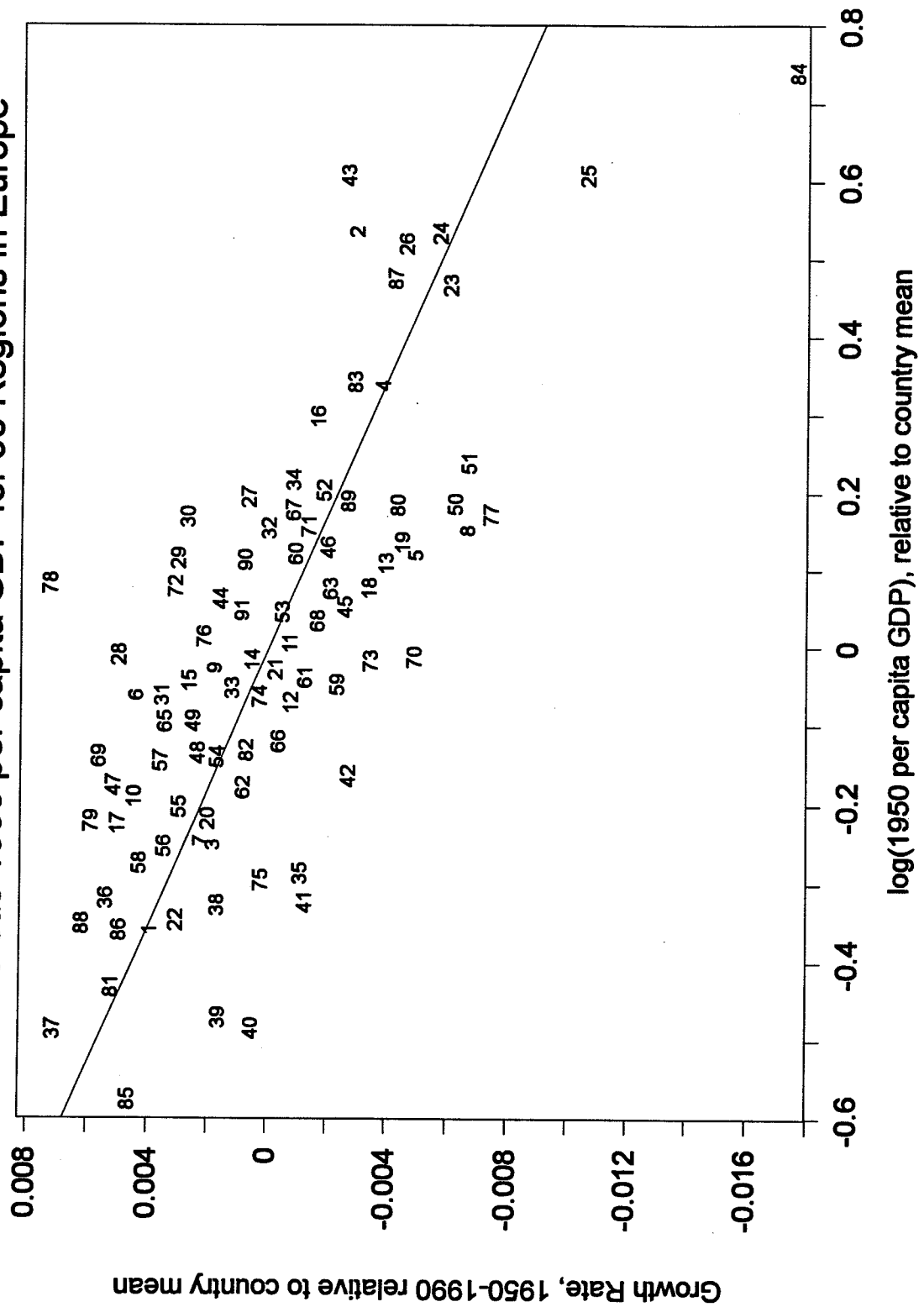


Figure 9: Dispersion of GDP per Capita within Five European Countries

