It's Still 2%: Evidence on Convergence from 116 Years of the US States Panel Data

Etsuro Shioji*
Department of Economics and Business
Universitat Pompeu Fabra†

September 21, 1997

Abstract

A new debate over the speed of convergence in per capita income across economies is going on. Cross sectional estimates support the idea of slow convergence of about two percent per year. Panel data estimates support the idea of fast convergence of five, ten or even twenty percent per year. This paper shows that, if you "do it right", even the panel data estimation method yields the result of slow convergence of about two percent per year.

*This research was partly financed by the DGICYT grant PB94-1095. I would like to thank Fabio Canova and Xavier Sala-i-Martin for stimulating discussions.

†From October 1997, Department of Economics, Yokohama National University, 79-3 Tokiwadai, Hodogaya-ku, Yokohama 240 Japan. E-mail: shioji@econ.ynu.ac.jp.

1. Introduction

There has been a debate on what the right speed of convergence in per capita incomes across economies is. On one hand, economists who use cross sectional approach tend to find a low estimated speed of convergence of about two to three percent per year. On the other hand, recent studies that take advantage of the panel dimension of the data using fixed effect approach tend to find much faster speed of convergence, ranging from 5% to as large as 20%. This difference is important not only because the two sharply different conclusions have very different welfare implications but also because they have very different implications for model building in growth theory.

A problem with the existing panel data studies is that the sample periods are too short. As some of the panel data techniques, most notably OLS with fixed effects, are plagued by small sample biases, it would be ideal to use a longer time series.. The US states personal income data offers an ideal ground for this kind of empirical studies in this respect. Not only has the BEA collected the data since 1929 till 1996 (so far), earlier estimates for the years 1880, 1900 and 1920 are available from Easterlin (1960). All together, we currently have the richness of 116 years of data on growth, although data is missing for some earlier years. OLS with fixed effects cannot easily deal with this kind of data with missing points, but I show that, with maximum likelihood approach, this situation can be dealt with in a straightforward manner.

What I find is the following. If I use the data from after 1929, estimated speed of convergence is much faster than two percent. However, once I take into account the period prior to 1929, the estimated speed becomes much closer to two percent. One may argue that this is because there was an important structural break around 1929, and that we could not put the periods before and after the break together to estimate the speed of convergence. The core part of the paper argues that this is not the case. Indeed, this difference across the periods is exactly what one would expect from econometric theory, even if there was no structural break, given the nature of the data. In fact it will be argued that the true speed of convergence must be much closer to the estimate that includes periods both before and after 1929. Hence, it is concluded that, even when one uses panel data approach, if the data is "right", the method uncovers the true speed of convergence — which is far from 10% or 20%, as some authors who took this approach argue for, but is close to 2%, as researchers who use the cross sectional approach claim.

The rest of the paper is organized as follows. The next section offers a quick overview of the literature. Section 3 explains the econometric approach. Section

4 demonstrates results from estimation. It is shown that estimated speed of convergence is fairly different, depending on whether the period prior to 1929 is included or not. Section 5 explains why this does not have to mean that there was an important structural break and that we could not put the two periods together into a single sample. This section further argues that the estimates from the sample that includes years prior to 1929 are much more likely to be close to the true speed of convergence. Section 6 concludes by saying that the true speed of convergence seems to be close to 2%, rather than 10% or 20%.

2. Overview

Estimating the speed of convergence has been one of the central issues of recent empirical literature of economic growth. At some point, the general consensus among economists on this issue was that economies converge to their respective steady states at a very slow rate of two percent per year. This speed means a half life of deviation from one's steady state of about 34.3 years. However, many of the recent studies cast doubt on this view using panel data estimation with fixed effects, which almost always produces much higher estimates for the speed of convergence that ranges from 4 to 20% For example, De la Fuente (1996) applies OLS with Fixed Effects to the Spanish regional data and comes up with the estimated speed of about 12%. Canova and Marcet (1995) apply their Bayesian estimation method to the regional data from Western Europe and countries in Europe, and come up with the estimated speed of about 20% for the former and 10% for the latter (though the values vary depending on the precise specification). Islam (1995) applies OLS with Fixed Effects as well as the minimum distance estimation method to the Summers Heston data set and comes up with estimates of between 4% and 9.3 % depending on the method and countries included in the regression. Also, Casselli, Esquivel and Lefort (1996), Evans (1995), and Knight, Laoyza and Villanueva (1993) apply different methods of panel data estimation to the Summers Heston data set and come up with estimated speed of convergence much higher than 2%. A speed of 10% (which seems to be more or less the mean of these estimates) would mean a half life of deviation from one's own steady state of about 6.6 years. Different estimates have very different implications for growth theories as well. The neoclassical growth model of Solow (1956) and Swan (1956), when one takes a narrow view on capital stock and identifies it solely as physical capital, implies a speed of 5-6 % (Barro and Sala-i-Martin (1995)). Hence, if the results from cross section regressions are to be trusted, the definition of capital stock in the model must be extended to include human capital or some other types of capital. On the other hand, if the mean of the panel data estimates of about

10% is to be trusted, there is no way the neoclassical growth model alone can explain such a high speed of convergence, as the speed of 10% implies a negative share of capital if the neoclassical growth model is taken literally (De la Fuente (1996))[1].

A problem with the past studies on convergence that used panel data is that the sample period tends to be very short. In an extreme case, the European regions sample in Canova and Marcet (1995) spans a period of about ten years only. The Summers and Heston data set runs for about 40 years. As our focus is to uncover long run relationship among per capita regional (national) per capita income, there is a worry that these are not sufficiently long. In fact, it is known that OLS with Fixed Effect estimation of convergence is subject to important small sample biases (Nerlove (1971), Nickell (1981), Hsiao (1986), Shioji (1997)), when the sample is small in the time dimension. In a sample of 40 years, assuming that the true speed of convergence is under 10%, the estimated speed of convergence is asymptotically (in the sense of the number of regions or countries approaching infinity) biased upwards by about 5%. In that sense, the US state personal income series seems to be the ideal data set to examine the robustness of the results of the recent studies, as the available data spans the period of 1880 to 1996, or 116 years.

3. Maximum Likelihood Estimation

It is known that, when the data exists for every year between the starting year and the end year, OLS with Fixed Effects estimation coincides with maximum likelihood estimation (Hsiao (1986)). However, in my sample, data from earlier years estimated by Easterlin is available for years 1880, 1900 and 1920 only, and the BEA data set starts from 1929. This situation is tricky to handle by OLS etc., but it can be handled in a straightforward manner by maximum likelihood. Assume that the model is

$$y_{it} - y_i^* = \rho \cdot (y_{it-1} - y_i^*) + u_{it} \tag{3.1}$$

where y_{it} is output per capita of region i in year t relative to, say, the mean of the whole country, y_i^* is its steady state value which is assumed to be time-invariant but could vary across the states, $\beta \equiv 1 - \rho$ is the **speed of convergence**, that is common across regions and periods ($|\rho| < 1$), α_i is the region specific fixed effect term, u_{it} is shock to the true output per capita which follows $N(0, \sigma_u^2)$ and

[1] On the other hand, if one takes the lower bound estimate of 4%, this value is consistent with the neoclassical growth model with the narrow definition of capital stock.

is assumed to be serially and spatially independent. Suppose that the data is available for N regions ($i = 1, 2, \cdots N$) for a set of periods denoted by $\bar{T}$ which is

$$\bar{T} \equiv \{t_j\}_{j=0}^{z}$$

where $t_0 = 0$, $t_z = T$ and $t_{j+1} > t_j$ for all j. Equation (3.1) implies, for each j between 1 and z,

$$y_{it_j} - y_i^* = \rho^{t_j - t_{j-1}} \cdot (y_{it_{j-1}} - y_i^*) + e_{it_j}$$

where

$$e_{it_j} = u_{it_j} + \rho \cdot u_{it_j - 1} + \rho^2 \cdot u_{it_j - 2} + \cdots + \rho^{t_j - t_{j-1}} \cdot u_{it_{j-1}+1}.$$

Note that

$$VAR(e_{it_j}) = \frac{1 - \rho^{2(t_j - t_{j-1})}}{1 - \rho^2} \cdot \sigma_u^2.$$

The maximum likelihood estimators for ρ, σ_u, and y_i^* are given by maximizing the following log likelihood function

$$\begin{aligned} L \equiv & -\frac{Nz}{2} \cdot \ln 2\pi - \frac{Nz}{2} \cdot \ln \sigma_u^2 + \frac{Nz}{2} \cdot \ln(1 - \rho^2) \\ & -\frac{N}{2} \cdot \sum_{j=1}^{z} \ln \left(1 - \rho^{2(t_j - t_{j-1})}\right) \\ & -\frac{1 - \rho^2}{2\sigma_u^2} \sum_{i=1}^{N} \sum_{j=1}^{z} \frac{1}{1 - \rho^{2(t_j - t_{j-1})}} \cdot \left[\left(y_{it_j} - y_i^*\right) - \rho^{t_j - t_{j-1}} \cdot \left(y_{it_{j-1}} - y_i^*\right)\right]^2. \end{aligned}$$

The estimated variance-covariance matrix of the unknown parameter is given by the negative of the inverse of the Hessian of the above function evaluated at the maximum.

4. Empirical Evidence

4.1. Evidence from Non-Skipping Data

First I present the result for the case when the whole available data is used. As was stated previously, the data is from 1880, 1900, 1920 and every year from 1929 to 1996. Out of 50 US states, earlier data is not available for Alaska and Hawaii, and the 1880 data is not available for Oklahoma. That leaves 47 states in my sample ($N = 47$). Table 1a (subtitled **No-SKIP**) demonstrates the result. To save space, only estimated speed of convergence $\equiv 1 - \rho$ together with its standard error and estimated σ_u are reported. In the first column I report the results when the whole sample is used. In the second column I report the result when only the data from 1880, 1900, 1920 and 1929 are used and in the third column I report the result for the case when only the data from 1929-1996 is used (the method for this case is OLS with fixed effects which coincides with maximum likelihood).

Table 1a (**No-SKIP**)
Estimation result when the whole data is used

period	all	1880,1900,20,29	1929-96
$1-\rho$	3.91%	3.64%	8.6%
(std.)	(0.30%)	(0.74%)	(0.67%)
σ_u	0.042	0.032	0.042

First, note that the estimated speed of convergence for the case when all the data is used is much lower than what is normally found in the literature of panel data estimation of convergence. Second, the estimated speed is much lower in the earlier period than in the later period. Third, the estimated speed from the overall period is much closer to that of the earlier period than that of the later period. Hence, several questions arise. (1) Why is the estimated speed so much lower than what is normally found using similar methods? (2) Why is the estimated speed so much lower than the earlier period? Isn't that an indication that there was an important structural break between the two periods? That is, isn't it unjustifiable to put the two periods together into one sample and estimate a single equation? I investigate these questions in the rest of the paper.

4.2. Measurement Error?

One difference between the earlier data and the later data is that the former is an estimate done retrospectively after many years from whatever data was available, whereas the latter is a series of estimates made contemporaneously using the data collected with the specific purpose of estimating personal incomes. Hence, it is conceivable that the earlier data is subject to measurement errors of greater extent. However, this does not explain the difference in the estimated speed of convergence between the two periods. As is shown in Shioji (1997), measurement errors tend to bias the estimated speed of convergence upwards. That is, if anything, the true speed of convergence in the earlier period is likely to be even lower than the estimated one. Hence, it does not explain why the earlier estimate is so much lower than the later one.

4.3. Evidence from the Skipping Estimation

Another difference between the earlier data and the later one is that the former is the "skipped data", that is, it does not come from every single year, whereas the latter does come from every single year. Shioji (1997) shows that, for the case of OLS with Fixed Effects estimation, when the data is subject to large measurement errors, skipping (taking data from every m years ($m \geq 2$) as opposed to every single year) tends to cut down the upward bias that exists due to these errors, at least for m not too large. That is, skipping tends to lower the estimated speed of convergence (and brings it closer to the true speed). Hence, the presence of skipping in the earlier sample and its absence in the later sample may explain why the estimated speed is lower in the former. And Shioji (1997) also shows that, using the data from 1929 to 1996, skipping indeed lowers the estimated speed of convergence for the US states, which suggests that the data is in fact subject to large measurement errors. And he argues that, when we talk about the convergence regression, the word "measurement errors" must be defined in the broad sense: whatever is a deviation from the growth model, such as business cycle components and temporary shocks to income, has to be considered as measurement errors from the viewpoint of growth theory.

Now I investigate this possibility. Obviously we cannot solve this problem by finding data for the earlier period, which does not exist. Instead I will see what happens when I do "skipping" for the later sample. In Table 1b (labelled "**SKIP-2**"), I use data from every two years for the period between 1930 and 1996. In Table 1c (labeled "**SKIP-5**"), I use data from every five years for the period between 1930 and 1995. In Table 1d (labeled "**SKIP-10**"), I use the data from every 10 years for the period between 1930 and 1990. In each of these tables, in the first column I report the estimated speed of convergence for the case where the data from 1880, 1900, 1920 are combined with the data from the later period. In the second column I use data from only after 1930. In Table 1e (labelled "**SKIP-20**"), I use the data from every 20 years. In the first column, I use the data from 1880, 1900, 1920, 1940, 1960, 1980, and 1996. In the second column, the data from the first half, namely 1880, 1900, 1920 and 1940 is used. In the third column the data from 1940, 1960, 1980 and 1996 is used.

Table 1b (**SKIP-2**)
For 1930 onwards, every two years

period	1880, 1900, 1920, 1930-96	1930-96
$1-\rho$	3.07%	7.6%
(std.)	(0.40%)	(0.60%)
σ_u	0.036	0.034

Table 1c (**SKIP-5**)

For 1930 onwards, every five years

period	1880, 1900, 1920, 1930-95	1930-95
$1-\rho$	2.46%	6.2%
(std.)	(0.28%)	(0.52%)
σ_u	0.030	0.042

Table 1d (**SKIP-10**)

For 1930 onwards, every ten years

period	1880, 1900, 1920, 1930-90	1930-90
$1-\rho$	2.44%	4.73%
(std.)	(0.29%)	(0.50%)
σ_u	0.030	0.045

Table 1e (**SKIP-20**)

Every twenty years (except between 1980 and 1996)

period	1880-1996	1880-1940	1940-1996
$1-\rho$	2.21%	3.48%	5.68%
(std)	(0.23%)	(0.51%)	(0.66%)
σ_u	0.028	0.029	0.017

A quick comparison from Table 1a through Table 1e reveals that the estimated speed of convergence goes down monotonically as the skipping becomes larger and larger, when the whole period is used as a sample. When only the latter period is used, the estimate goes down monotonically except between "SKIP-10" and "SKIP-20". These are evidence that the "non-skipping" estimates in Table 1a are subject to large upward bias due to measurement errors. Thus the true speed of convergence must be close to about 2% in Table 1d or 1e than about 4% in Table 1a. Also, note that the difference between the whole period estimate and the later period estimate shrinks somewhat as the skipping becomes larger. This suggests that the huge difference between these two in Table 1a was at least partially due to the fact that there was a fair amount of skipping for the earlier period but there was no skipping in the later period. However, the most important message from Table 1d and 1e is that, even with a large skipping, a large difference between the whole period estimate and the later period estimate remains. Hence, it looks like the doubt that there was an important structural change between the earlier period and the later period is justified.

In the next section it is shown that this needs not be the case. This difference between the whole period (or the earlier period) estimate and the later period counterpart is exactly what is expected from econometric theory. Estimates from the whole sample in tables 1d and 1e are fairly reliable estimates of the true speed of convergence.

5. Initial Value Distribution and Biases in Estimated Speed of Convergence

5.1. Case of OLS (Analytical results)

In this section I argue that the difference between the whole period estimate (or the earlier period estimate) and the later period estimate is exactly what is expected from theory. I will first discuss the case of OLS with Fixed Effects, as analytical results are easy to derive in this case. In the next subsection I will discuss the case of Maximum Likelihood using Monte Carlo experiments. Intuition behind the two results is the same.

I will consider the general case where measurement errors are present, as this seems to be the case in the data I am dealing with. Suppose that the true model is equation (3.1), but that the observed relative income per capita, denoted x_{it}, is subject to measurement errors:

$$x_{it} = y_{it} + v_{it}$$

where v_{it} is the measurement error with mean 0 and variance σ_v^2 which is assumed to be serially and spatially independent. Distributions of both u_{it} and v_{it} are assumed to be space- and time-invariant. Let's say an econometrician use this observed incubi per capita to estimate an equation

$$x_{it} = \rho \cdot x_{it-1} + \alpha_i + u_{it}$$

using OLS with Fixed Effects. Nickell (1981) and Hsiao(1986) derive the asymptotic bias (as N goes to infinity) in the estimate for the case without measurement errors. Shioji (1997) generalizes their result to the case with measurement errors:

$$\hat{\rho} - \rho = -(B_1 + B_3)/(B_2 + B_4)$$

where

$$B_1 = \frac{1}{T^2} \cdot \frac{(T-1) - T \cdot \rho + \rho^T}{(1-\rho)^2}$$

$$B_2 = \frac{1}{1-\rho^2} \cdot \left\{ 1 - \frac{1}{T} - \frac{2 \cdot \rho}{(1-\rho)^2} \cdot \frac{(T-1) - T \cdot \rho + \rho^T}{T^2} \right\}.$$

where

$$B_3 = \left(\rho - \rho \cdot \frac{1}{T} + \frac{T-1}{T^2} \right) \cdot \left(\frac{\sigma_v}{\sigma_u} \right)^2$$

and

$$B_4 = \left(1 + \frac{1}{T} \right) \cdot \left(\frac{\sigma_v}{\sigma_u} \right)^2.$$

The bias is negative unless T is very small.

A potential problem with the derivation of the above expression is that it assumes that the initial values, y_{i0}, are drawn from their stationary distribution. That is, it is implicitly assumed that the economies start from the stationary states. This assumption is not necessarily appropriate in many cases in practice, and is definitely unrealistic in the case of the US states. Shortly before the sample starts there was the Civil War, which most likely sent states far out of their stationary distributions. In fact, cross sectional dispersion of state per capita incomes declined almost monotonically between 1880 and the late 1970s (Barro and Sala-i-Martin (1992)) which seems to indicate that the states were far away from the stationary distribution in 1880 and spent nearly a century to get back to their stationary distribution. When the initial value distribution is allowed to differ from the stationary distribution, the above expression undergoes a slight but potentially important change. The stationary distribution of y_{i0} has mean y_i^* and standard deviation $\sigma_u/\sqrt{1-\rho^2}$. Now assume that the initial value distribution has the same mean but the standard deviation is η (>0) times larger. Then the above expression becomes:

$$\hat{\rho} - \rho = -(B_1 + B_3)/(B_2 + B_4 + B_5)$$

where

$$B_5 = \left[\frac{1}{T} \cdot \frac{1-\rho^{2(T-1)}}{1-\rho^2} - \frac{1}{T^2} \cdot \frac{(1-\rho^T)^2}{(1-\rho)^2}\right] \cdot \left(\eta^2 - 1\right) \cdot \frac{1}{1-\rho^2}.$$

The only change is the presence of B_5 in the new expression. When η is greater than 1, that is, when the initial value distribution is more disperse than the stationary distribution (as seems to be the case), B_5 is positive for a large T. And if that is the case, the absolute value of the bias decreases with η. That is, **a large dispersion in the initial value distribution cuts down the bias** in the estimated speed of convergence. The intuition is the usual signal-noise argument. When the economies start out far away from their respective non-stochastic steady states, the data contains more information on how quickly (or slowly) they move back toward their steady states. This fact could potentially explain why the whole period (or the earlier period) estimated speed of convergence was much lower than the later period estimate. As I argued, it is quite conceivable that the states started far away from their stationary distribution in 1880. Hence, when the sample that starts from this year is used, the bias in the estimated speed of convergence is likely to be small. However, between 1880 and 1930, the economies moved closer to their respective stationary distribution. Hence, when the sample that starts from 1930 (or 1940) is used, the upward bias in the estimated speed of convergence is likely to be large. This could explain why the estimated speed turned out to be much larger for the later period in the previous section. And if this is true, there is no need to presume that there was an important structural change. In fact, it could be concluded that the whole period estimate of about 2% is a fairly bias-free estimate.

5.2. Case of ML (Numerical results)

As analytical results are hard to derive in the case of Maximum Likelihood, I rely on the Monte Carlo experiment. The experiment is in line with the empirical exercise in Section 3. Number of regions, N, is set to 47. Total number of years, T, is 116. For each trial, I estimate the speed of convergence for the earlier period: $t = 0$ (for the year 1880), 20 (for 1900), 40 (for 1920) and 49 (for 1929) and for the later period: $t = 50$ to 116 (for the years 1930 to 1996), and also for the combined period, using maximum likelihood. I set the underlying parameter values as follows. I set ρ to be 0.98, which means the true speed of convergence of 2%. I set σ_v, the standard deviation of the measurement error term, to be half of σ_u. This follows the result of Shioji (1997). If we define measurement errors as literally errors in measuring incomes, this is probably too large. However, as has already been argued, when talking about growth theory, any short run movements in incomes that are not captured by growth theory should be considered as measurement errors. From this viewpoint the above choice of value is by no means extreme. The question is how to choose η, the ratio between the standard deviation of the initial value distribution and that of the stationary distribution. The latter can in principle be obtained from the estimates in section 3. However, I discarded this approach as the standard error around its estimate was huge. Instead, in Table 2, I compute standard deviation in the difference between y_{it} and the estimated y^* across the states. The estimates are based on the case of Table 1e ("Skip-20"), and I only report the average across years for the later period. Results are similar when other estimates from section 3 are used.

Table 2
std. of the deviation from steady state

1880	0.494
1900	0.387
1920	0.223
1929-39	0.234
1940-49	0.166
1950-59	0.116
1960-69	0.080
1970-79	0.092
1980-89	0.084
1990-96	0.079

One gets an impression that, starting from a dispersion far larger than that of the stationary state, around 0.5, the states more or less converged to the stationary distribution since around 1950, with the standard deviation floating around somewhat below 0.1. Hence, the standard deviation of the initial distribution (of $y_{it} - y_i^*$) seems to be 4-5 times larger than that of the stationary distribution.. From this exercise, I set $\eta = 4$. I also set $\sigma_u = 0.019$, so that the standard deviation in the steady state distribution is 0.1. Finally, for each trial, I draw non-stochastic steady state values y_i^* from a uniform distribution with mean 0 and standard deviation = 0.5.

I run the experiment 100 times and simply take the average. The result is shown in Table 3.

Table 3: Monte Carlo Experiments

	whole	0,20,40,49	50-116
average	3.56%	4.11%	7.40%
std.	0.44%	0.94%	0.95%

The first column refers to the case when the whole period, namely $t = 0, 20, 40, 49$, and $50 - 116$ is all used for estimation. The second column reports the case when only $t = 0, 20, 40$ and 49 are used for estimation. In the third column I use $t = 50 - 116$ only and use OLS for the estimation. The first row reports the average of the point estimate over the 100 trials, and the second row reports the standard deviation of the point estimate over these trials. Note that, despite that the true parameter is the same across periods, the average estimates for the whole sample and for the earlier period are much lower than that for the later period. And the former two are much closer to the true value than the latter one. Hence, it can be concluded that the differences in the estimated speed of convergence across periods in section 4 do not necessarily imply the presence of a structural break between the periods. In fact, the presence of these differences is exactly what you would expect from econometric theory, once you take into account that the initial distribution of the US states incomes was probably far more disperse than the stationary distribution. Note how close the experiment results in Table 3 are to the empirical estimates in Table 1a. And the true speed of convergence has to be much close to the whole period estimate (or the earlier period estimate) than the latter period estimate, that is, it is likely to be much closer to 2% than 8.6% in the third column of Table 1a. Definitely not 10% or 20%.

6. Conclusion

It is always a truth in econometrics that a good data set is a data set that contains large variations. The US states data set is a good data set to study convergence not only because a long time series is available but also because it shows a huge initial distribution due to the Civil War and it records how the states came back toward their respective steady states from this distribution. That is why it eloquently tells how the process of convergence looks like. Using this data set, I have shown that even the panel data estimation method yields the result of slow convergence of about 2% per year.

Appendix

A. Data Sources

The original data for 1929-1996 is from Bureau of Economic Analysis (BEA). For 1929-1992, I downloaded the data from Xavier Sala-i-Martin's home page. I combined the two series that appear in his file using the growth rate. For 1992-1996, I downloaded the data from the home page of BEA. I linked this series with the previous series using the growth rate.

References

[1] Anderson, T.W. and C. Hsiao, "Estimation of Dynamic Models with Error Components", *Journal of the American Statistical Association*, 76, 598-606, 1981.

[2] Arellano, Manuel, and Stephen Bond, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, 58, 277-297, 1991.

[3] Barro, Robert J. and Xavier Sala-i-Martin, "Convergence", *Journal of Political Economy*, vol. 100, 223-251, April 1992.

[4] Barro, Robert J. and Xavier Sala-i-Martin, *Economic Growth*, McGraw Hill, 1995.

[5] Canova, Fabio and Albert Marcet, "The Poor Stay Poor: Non-Convergence across Countries and Regions", Universitat Pompeu Fabra working Paper 137, October 1995.

[6] Caselli, Francesco, Gerardo Esquivel and Fernando Lefort, "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics", *Journal of Economic Growth*, 1, 363-389, September 1996.

[7] De la Fuente, Angel, "On the Sources of Convergence: A Close Look at the Spanish Regions", mimeo, Universitat Autònoma de Barcelona, October 1996.

[8] Easterlin, Richard A., "Regional Growth of Income: Long Term Tendencies", in *Population Redistribution and Economic Growth, United States, 1870-1950*, vol. 2, *Analysis of Economic Change*, edited by Simon Kuznets, Ann Ratner Miller, and Richard A. Easterlin, Philadelphia, American Philosophical Society, 1960.

[9] Evans, Paul, "How to Estimate Growth Regressions Consistently", mimeo, Ohio State University, January 1995.

[10] Hsiao, Cheng, *Analysis of Panel Data*, Cambridge University Press, Cambridge, UK, 1986.

[11] Islam, Nazrul, "Growth Empirics: A Panel Data Approach", *Quarterly Journal of Economics*, 110, 1127-1170, 1995.

[12] Knight, Malcolm, Norman Loayza and Delano Villanueva, "Testing the Neoclassical Theory of Economic Growth: A Panel Data Approach", *IMF Staff Papers*, 40, 512-541, September 1993.

[13] Nerlove, Marc, "Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections", *Econometrica*, 39, no. 2, 359-382, March 1971.

[14] Nickell, Stephen, "Biases in Dynamic Models with Fixed Effects", *Econometrica*, 49, no. 6, 1417-1426, November, 1981.

[15] Shioji, Etsuro, "Convergence in Panel Data: Evidence from the Skipping Estimation", mimeo, Universitat Pompeu Fabra, 1997.

[16] Solow, Robert M., "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, 70, 65-94, February 1956.

[17] Swan, Trevor W., "Economic Growth and Capital Accumulation", *Economic Record*, 32, 334-361, November 1956.