



Convergence

Robert J. Barro, Xavier Sala-i-Martin

The Journal of Political Economy, Volume 100, Issue 2 (Apr., 1992), 223-251.

Your use of the JSTOR database indicates your acceptance of JSTOR's Terms and Conditions of Use. A copy of JSTOR's Terms and Conditions of Use is available at <http://www.jstor.org/about/terms.html>, by contacting JSTOR at jstor-info@umich.edu, or by calling JSTOR at (888)388-3574, (734)998-9101 or (FAX) (734)998-9113. No part of a JSTOR transmission may be copied, downloaded, stored, further transmitted, transferred, distributed, altered, or otherwise used, in any form or by any means, except: (1) one stored electronic and one paper copy of any article solely for your personal, non-commercial use, or (2) with prior written permission of JSTOR and the publisher of the article or other text.

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.

The Journal of Political Economy is published by University of Chicago. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ucpress.html>.

The Journal of Political Economy
©1992 University of Chicago

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©1999 JSTOR

Convergence

Robert J. Barro

Harvard University and National Bureau of Economic Research

Xavier Sala-i-Martin

Yale University and National Bureau of Economic Research

A key economic issue is whether poor countries or regions tend to grow faster than rich ones: are there automatic forces that lead to convergence over time in the levels of per capita income and product? We use the neoclassical growth model as a framework to study convergence across the 48 contiguous U.S. states. We exploit data on personal income since 1840 and on gross state product since 1963. The U.S. states provide clear evidence of convergence, but the findings can be reconciled quantitatively with the neoclassical model only if diminishing returns to capital set in very slowly. The results for per capita gross domestic product from a broad sample of countries are similar if we hold constant a set of variables that proxy for differences in steady-state characteristics.

A key economic issue is whether poor countries or regions tend to grow faster than rich ones: are there automatic forces that lead to convergence over time in the levels of per capita income and product? We use the neoclassical growth model as a framework to study convergence across the 48 contiguous U.S. states. We exploit data on personal income since 1840 and on gross state product since 1963. For studying the determinants of economic growth, the experience

We have benefited from research support by the National Science Foundation, the Bradley Foundation, and the Hoover Institution. We appreciate helpful comments from Gary Becker, Paul David, Steve Durlauf, Susan Guthrie, Carol Heim, Anne Krueger, Edward Lazear, Bob Lucas, Greg Mankiw, Kevin M. Murphy, Danny Quah, Sergio Rebelo, and Gavin Wright.

[Journal of Political Economy, 1992, vol. 100, no. 2]

© 1992 by The University of Chicago. All rights reserved. 0022-3808/92/0002-0001\$01.50

of the U.S. states represents a vastly underutilized resource: in effect, we have over a century of data on 48 economies (although surely not 48 closed economies!).

The U.S. states provide clear evidence of convergence in the sense that poor economies tend to grow faster than rich ones in per capita terms. The estimated speed of convergence accords with the neoclassical growth model if we take a broad view of capital so that diminishing returns to capital set in slowly as an economy develops. The findings for the U.S. states can be reconciled with those for a broad cross section of countries if we allow for a notion of conditional convergence in the underlying growth model. Some puzzles arise, however, in reconciling the data with open-economy extensions of the model. In particular, the rates of convergence found for income and product across the U.S. states are similar, whereas theoretical reasoning suggests some important differences.

Convergence in the Neoclassical Growth Model

In neoclassical growth models for closed economies, as presented by Ramsey (1928), Solow (1956), Cass (1965), and Koopmans (1965), the per capita growth rate tends to be inversely related to the starting level of output or income per person. In particular, if economies are similar in respect to preferences and technology, then poor economies grow faster than rich ones. Thus there is a force that promotes convergence in levels of per capita product and income. Since the model is familiar, we provide only a brief sketch.

The production function in intensive form is

$$\hat{y} = f(\hat{k}), \quad (1)$$

where \hat{y} and \hat{k} are output and capital per unit of effective labor, Le^{xt} , L is labor (and population), and x is the rate of exogenous, labor-augmenting technological progress. (We assume the usual curvature properties for the production function.) In a closed economy, \hat{k} evolves as

$$\dot{\hat{k}} = f(\hat{k}) - \hat{c} - (\delta + x + n)\hat{k}, \quad (2)$$

where $\hat{c} = C/Le^{xt}$, δ is the rate of depreciation, and n is the growth rate of L . The representative, infinite-horizon household maximizes utility,

$$U = \int_0^{\infty} u(c) e^{nt} e^{-\rho t} dt, \quad (3)$$

where $c = C/L$, ρ is the rate of time preference, and

$$u(c) = \frac{c^{1-\theta} - 1}{1-\theta}, \quad (4)$$

with $\theta > 0$, so that marginal utility, $u'(c)$, has the constant elasticity $-\theta$ with respect to c . (We assume $\rho > n + [1 - \theta]x$ below to satisfy the transversality condition.)

The first-order condition for maximizing U in equation (3) entails

$$\frac{\dot{c}}{c} = \frac{1}{\theta} \cdot [f'(\hat{k}) - \delta - \rho]. \tag{5}$$

In the steady state, the effective quantities, \hat{y} , \hat{k} , and \hat{c} , do not change and the per capita quantities, y , k , and c , grow at the rate x . The level of \hat{k} in the steady state satisfies

$$f'(\hat{k}^*) = \delta + \rho + \theta x. \tag{6}$$

If the economy starts with \hat{k} below \hat{k}^* , then the usual analysis shows that \hat{k} monotonically approaches \hat{k}^* (see, e.g., Blanchard and Fischer 1989, chap. 2). We have shown (Barro and Sala-i-Martin 1991*b*, chap. 1) that the growth rate of capital per worker, $\dot{\hat{k}}/\hat{k}$, declines monotonically toward the steady-state value, x . This property carries over unambiguously to the growth rate of output per worker, $\dot{\hat{y}}/\hat{y}$, if the production function is Cobb-Douglas, that is, if

$$\hat{y} = f(\hat{k}) = A\hat{k}^\alpha, \tag{7}$$

where $0 < \alpha < 1$. Thus if two economies have the same parameters of preferences and technology, then the key result is that the initially poorer economy—with a lower starting value of \hat{k} —tends to grow faster in per capita terms.

The transitional dynamics can be quantified by using a log linearization of equations (2) and (5) around the steady state. The solution for $\log[\hat{y}(t)]$ in the log-linearized approximation to the model with a Cobb-Douglas technology is

$$\log[\hat{y}(t)] = \log[\hat{y}(0)] \cdot e^{-\beta t} + \log(\hat{y}^*) \cdot (1 - e^{-\beta t}), \tag{8}$$

where the positive parameter β , which governs the speed of adjustment to the steady state, is given by the formula

$$\begin{aligned} 2\beta = & \left\{ \psi^2 + 4 \left(\frac{1 - \alpha}{\theta} \right) (\rho + \delta + \theta x) \right. \\ & \left. \times \left[\frac{\rho + \delta + \theta x}{\alpha} - (n + \delta + x) \right] \right\}^{1/2} - \psi, \end{aligned} \tag{9}$$

where $\psi = \rho - n - (1 - \theta)x > 0$.

The average growth rate of y over the interval between dates 0 and T is

$$\frac{1}{T} \cdot \log \left[\frac{y(T)}{y(0)} \right] = x + \frac{1 - e^{-\beta T}}{T} \cdot \log \left[\frac{\hat{y}^*}{\hat{y}(0)} \right]. \tag{10}$$

The higher β , the greater the responsiveness of the average growth rate to the gap between $\log(\hat{y}^*)$ and $\log[\hat{y}(0)]$, that is, the more rapid the convergence to the steady state. The model implies *conditional convergence* in that, for given x and \hat{y}^* , the growth rate is higher the lower $\hat{y}(0)$. The convergence is conditional in that $\hat{y}(0)$ enters in relation to \hat{y}^* and x , which may differ across economies. In cross-country regressions, it is crucial, but difficult, to hold fixed the variations in \hat{y}^* and x in order to estimate β . One advantage of the U.S. state context is that the differences in \hat{y}^* and x are likely to be minor, so that conditional and absolute convergence need not be distinguished.

Because the crucial element for convergence in the neoclassical model is diminishing returns to capital, the extent of these diminishing returns—that is, the size of the capital-share coefficient α in equation (7)—has a strong effect on β . To assess the relation quantitatively we use a set of baseline values for the other parameters: $\rho = .05$ per year, $\delta = .05$ per year, $n = .02$ per year, $x = .02$ per year, and $\theta = 1$ (log utility). The value $n = .02$ per year is the average of population growth for the United States over the long history. The other baseline parameters come from estimates reported in Jorgenson and Yun (1986, 1990). If we assume $\alpha = .35$ —a capital share appropriate to a narrow concept of physical capital (see, e.g., Maddison 1987)—then equation (9) implies $\beta = .126$ per year, which corresponds to a half-life for the log of output per effective worker of 5.5 years. For $\alpha = .80$, which might apply if capital is interpreted broadly to include human capital, the value $\beta = .026$ per year implies a half-life of 27 years. As α approaches unity, diminishing returns to capital disappear, β tends to zero, and the half-life tends to infinity.

The effects of the other parameters have been explored by Chamley (1981) and King and Rebelo (1989).¹ Quantitatively, the most important effect is that a lower θ (increased willingness to substitute intertemporally) raises β . Another result is that the parameter A in equation (7) does not affect β . Thus the convergence coefficient β can be similar across economies that differ greatly in levels of per capita product because of differences in the available technique (or in government policies or natural resources that amount to differences in the parameter A).

The main result for the subsequent analysis is that the baseline specification—including $\alpha = .35$ —generates a short half-life and a rapid speed of adjustment. The speeds of adjustment that we estimate empirically are much slower: β is in the neighborhood of .02 per year. The theory conforms to the empirical findings only if we assume parameter values that depart substantially from the baseline case.

¹ Sato (1966) presents a related analysis for a model with a constant saving rate.

One possibility is a value of α around .8, that is, in the range in which the broad nature of capital implies that diminishing returns set in slowly. We can reduce the required value of α to around .5 if we assume very high values of θ (in excess of 10) and a value of δ close to zero.

Setup of the Empirical Analysis

Consider a version of equation (10) that applies for discrete periods to economy i and is augmented to include a random disturbance:

$$\log\left(\frac{y_{it}}{y_{i,t-1}}\right) = a_i - (1 - e^{-\beta}) \cdot [\log(y_{i,t-1}) - x_i \cdot (t - 1)] + u_{it}, \quad (11)$$

where $a_i = x_i + (1 - e^{-\beta})\log(\hat{y}_i^*)$, and u_{it} is a disturbance term. Although the coefficient β can vary across economies, we neglect these differences in our analysis. This assumption is tenable for the U.S. states, which are likely to be similar in terms of the underlying parameters of technology and preferences. Also, as mentioned before, the theory implies that pure differences in the level of technology do not affect β . Thus β can be similar for economies that are very different in other respects.

In the application to the U.S. states, we assume that the coefficient a_i in equation (11) is the same for all i ; that is, we assume that the steady-state value, \hat{y}_i^* and the rate of technological progress, x_i , do not differ across states. The time trend, $x_i \cdot (t - 1)$, is then also the same for all i . The conditions $a_i = a$ and $x_i = x$ in equation (11) imply that poor economies tend to grow unconditionally faster than rich ones if $\beta > 0$. Because the coefficient on $\log(y_{i,t-1})$ is $1 - e^{-\beta}$, which is between zero and one, the convergence is not strong enough to eliminate the positive serial correlation in $\log(y_{it})$. Put alternatively, in the absence of random shocks, convergence to the steady state is direct and involves no oscillations. This property reflects the absence of overshooting in the neoclassical growth model.

Convergence in the sense that poor economies tend to grow faster than rich ones, which corresponds to $\beta > 0$ if a_i and x_i are the same for all i in equation (11), does not necessarily imply that the cross-economy dispersion of $\log(y_{it})$ declines over time. The effect from $\beta > 0$, which tends to reduce dispersion, is offset by random shocks, u_{it} , which tend to raise dispersion. If u_{it} has zero mean and variance σ_u^2 and is distributed independently over time and across economies, then the cross-economy variance of $\log(y_{it})$, denoted σ_t^2 , evolves as

$$\sigma_t^2 = (e^{-2\beta})\sigma_{t-1}^2 + \sigma_u^2, \quad (12)$$

which implies

$$\sigma_t^2 = \frac{\sigma_u^2}{1 - e^{-2\beta}} + \left(\sigma_0^2 - \frac{\sigma_u^2}{1 - e^{-2\beta}} \right) e^{-2\beta t}. \quad (13)$$

(We assume here that the cross section is large enough so that the sample variance of $\log[y_{it}]$ corresponds to the population variance, σ_t^2 .) Equation (13) implies that σ_t^2 monotonically approaches the steady-state value, $\sigma^2 = \sigma_u^2/(1 - e^{-2\beta})$, which rises with σ_u^2 but declines with β . The variance σ_t^2 falls (or rises) over time if the initial value σ_0^2 is greater than (or less than) σ^2 . Thus a positive coefficient β does not ensure a falling σ_t^2 .

Shocks that have common influences on subgroups of countries or regions, such as harvest failures and oil shocks, imply that u_{it} in equation (11) would not be independent of u_{jt} for $j \neq i$. An important example of this kind of shock from U.S. history is the Civil War, which had a strong adverse effect on the southern states relative to the northern states. We can handle this type of situation by writing the error term, u_{it} , in equation (11) as the sum of an aggregate influence and an independent disturbance:

$$\log\left(\frac{y_{it}}{y_{i,t-1}}\right) = a - (1 - e^{-\beta}) \cdot [\log(y_{i,t-1}) - x \cdot (t - 1)] + \phi_i s_t + v_{it}, \quad (14)$$

where s_t is an aggregate shock, which has zero mean and variance σ_s^2 , and ϕ_i measures the effect of the aggregate disturbance on the growth rate of economy i . We assume that, with $\phi_i s_t$ held constant, the error term, v_{it} , is cross-sectionally and serially independent with zero mean and constant variance σ_v^2 .²

We assume that the coefficients ϕ_i in equation (14) have mean $\bar{\phi}$ and variance σ_ϕ^2 and are distributed independently of v_{it} . If $\log(y_{i,t-1})$ and ϕ_i are uncorrelated, then estimates of the coefficient β in equation (14) would not be systematically related to the realization of s_t because the composite error term, $u_{it} = \phi_i s_t + v_{it}$, is uncorrelated with the regressor, $\log(y_{i,t-1})$. Suppose, alternatively, that $\text{cov}[\log(y_{i,t-1}), \phi_i] > 0$; for example, if a positive s_t represents an increase in the relative price of oil, then economies that produce a lot of oil ($\phi_i > \bar{\phi}$) tend to have high values of $y_{i,t-1}$. In this case, the least-squares estimate of the coefficient on $\log(y_{i,t-1})$ in equation (14)

² The specification in eq. (14) means that realizations of s_t effectively shift σ_t^2 in eqq. (12) and (13). Thus the approach of σ_t^2 to a steady-state value need no longer be monotonic. We plan in future research to analyze the time series of σ_t^2 for the U.S. states.

is biased for a given realization of s_t . For example, if oil-producing economies have relatively high values of $y_{i,t-1}$, then least-squares procedures tend to underestimate β for a period in which the oil price rises.³

In the empirical analysis, we include variables that we think hold constant the effects of aggregate shocks, s_t , on economy i 's growth rate. One reason to add these variables is to achieve cross-sectional independence of the error terms, v_{it} , in equation (14): the composite error, $u_{it} = \phi_i s_t + v_{it}$, would not exhibit this independence. The second purpose is to obtain consistent estimates of the coefficient β , conditional on the realizations of s_t .

The Data for the U.S. States

We have two measures of per capita income or product across the U.S. states. The first is per capita personal income. The U.S. Commerce Department has published annual data on nominal personal income for the 48 continental states since 1929 (see Bureau of Economic Analysis [1986] and recent issues of *Survey of Current Business*). We use the figures that exclude transfer payments from all levels of government. Easterlin (1960a, 1960b) provides estimates of state personal income for 1840 (29 states or territories), 1880 (47 states or territories), 1900 (48 states or territories), and 1920 (48 states). These data also exclude transfer payments.

We lack useful measures of price levels or price indexes for individual states. Therefore, we deflate the nominal values for each state by the national index for consumer prices. Since we use the same price deflator for each state in a single year, the particular deflator that we use affects only the constant terms in the subsequent regressions. The use of the same deflator for each state introduces two types of potential measurement error. First, if relative purchasing power parity does not hold across the states, then the growth rates of real per capita income are mismeasured. Second, if absolute purchasing power parity does not hold, then the levels of real per capita income are mismeasured.

The second type of data is per capita gross state product (GSP), which is available annually for each state from 1963 to 1986 (see Renshaw, Trott, and Friedenbergl 1988). This variable, which is anal-

³ We assume here that y_{it} represents either real per capita income for residents of economy i (corresponding to the data on state personal income) or the real per capita income derived from production of goods and services in economy i (corresponding to the figures on gross state product). Hence, changes in relative prices show up directly as changes in y_{it} : e.g., if no quantities change, then an increase in the relative price of oil generates a high growth rate of y_{it} for economies that produce a lot of oil.

ogous to gross domestic product (GDP), measures factor incomes derived from production within a state. We deflate the nominal figures by the aggregate GSP deflator for the year. (This deflator is close to that for U.S. GDP.) Since we use a common deflator for each state at a point in time, the particular deflator chosen is again of no consequence. We should stress, however, that the GSP figures that we use are not quantity indexes, but rather represent the incomes accruing to factors from the goods and services produced within a state.

The main differences between state personal income and GSP involve capital income. Personal income includes corporate net income only when individuals receive payment as dividends, whereas GSP includes corporate profits and depreciation. (Neither concept includes capital gains.) Most important, GSP attributes capital income to the state in which the business activity occurs, whereas personal income attributes it to the state of the asset holder.⁴

Evidence on Convergence for the U.S. States

We use the data on real per capita income or product, y_{it} , for a cross section of the U.S. states, $i = 1, \dots, N$. Equations (10) and (11) imply that the average growth rate over the interval between any two points in time, t_0 and $t_0 + T$, is given by

$$\frac{1}{T} \cdot \log\left(\frac{y_{i,t_0+T}}{y_{i,t_0}}\right) = B - \left(\frac{1 - e^{-\beta T}}{T}\right) \cdot \log(y_{i,t_0}) + u_{i,t_0,t_0+T}, \quad (15)$$

where u_{i,t_0,t_0+T} is a distributed lag of the error terms, u_{it} , between dates t_0 and $t_0 + T$.⁵ The constant term is $B = x + [(1 - e^{-\beta T})/T] \cdot [\log(\hat{y}^*) + x t_0]$, which is independent of i because we assumed $\hat{y}_i^* = \hat{y}^*$ and $x_i = x$. The coefficient B shifts because of the trend in technology with a change in the starting date, t_0 .

The coefficient on $\log(y_{i,t_0})$ in equation (15) is $-(1 - e^{-\beta T})/T$, which declines in magnitude with the length of the interval, T , for a given β . As T gets larger, the effect of the initial position on the average growth rate gets smaller; as T tends to infinity, the coefficient tends to zero. We estimate β nonlinearly to take account of the associated value of T in the form of equation (15). Therefore, we should obtain similar estimates of β regardless of the length of the interval.

Table I contains nonlinear least-squares regressions in the form of

⁴ Some of these locational considerations apply also to labor income, although—except for a few cities—the location of a business and the residence of the workers are typically in the same state.

⁵ The error term is $1/T$ times the sum for τ between zero and T of the error terms, $u_{i,t_0+\tau}$, weighted by $e^{-\beta(T-\tau)}$.

TABLE I
CROSS-STATE REGRESSIONS FOR PERSONAL INCOME

Sample	$\hat{\beta}$	Sectoral Composition (s_{ij})	R^2	σ
1. 1880-1988	.0175 (.0046)92	.0014
2. 1880-1900	.0224 (.0040)62	.0054
3. 1900-1920	.0209 (.0063)67	.0062
4. 1920-30	-.0122 (.0074)43	.0111
5. 1930-40	.0127 (.0051)36	.0075
6. 1940-50	.0373 (.0053)86	.0057
7. 1950-60	.0202 (.0052)49	.0048
8. 1960-70	.0135 (.0043)68	.0037
9. 1970-80	.0119 (.0069)36	.0056
10. 1980-88	-.0005 (.0114)51	.0103
11. Nine periods, β restricted*	.0189 (.0019)
12. 1880-1900	.0268 (.0048)	-.0161 (.0079)	.65	.0053
13. 1900-1920	.0269 (.0075)	-.0214 (.0094)	.71	.0060
14. 1920-30	.0218 (.0112)	-.0936 (.0175)	.64	.0089
15. 1930-40	.0141 (.0048)	2.43 (.81)	.46	.0070
16. 1940-50	.0362 (.0055)	-.40 (.57)	.87	.0057
17. 1950-60	.0313 (.0055)	.42 (.09)	.65	.0041
18. 1960-70	.0194 (.0052)	.55 (.25)	.71	.0036
19. 1970-80	.0139 (.0076)	.25 (.37)	.36	.0056
20. 1980-88	.0196 (.0106)	1.35 (.22)	.73	.0077
21. Nine periods, β restricted*	.0249 (.0021)	individual
22. 1840-80	.0254 (.0067)91	.0030

Note.—Standard errors of coefficients are shown in parentheses. Regression 22 has 29 observations, regressions 1 and 2 have 47 observations (excluding Oklahoma), and regression 12 has 46 observations (excluding Oklahoma and Wyoming). All others have 49 observations. The dependent variable is the growth rate of real per capita personal income exclusive of transfers over the indicated sample period. Each regression includes a constant and three regional dummy variables, south, midwest, and west. (Regression 22 includes only south and midwest.) The coefficient β applies to $\log(y_{it}/y_{i0})$, where y_{it} is real per capita personal income at the start of the period. The sectoral composition variable, s_{ij} , is described in the text. The regressions denoted nine periods, β restricted use nonlinear, iterative weighted least squares, with the coefficient β constrained to be equal for all nine subperiods. Individual coefficients are estimated for each subperiod for the constant, regional dummies, and the sectoral composition variable.

* For line 11, the log likelihood ratio is 32.1 (p -value = .000); for line 21, it is 13.9 (p -value = .084). The likelihood ratio statistic refers to the hypothesis of equality for the β coefficients. Under the null hypothesis, this statistic is distributed as χ^2 with eight degrees of freedom.

equation (15) for the U.S. states or territories and for various time periods.⁶ Aside from $\log(y_{i,t_0})$, each regression includes a constant and three regional dummy variables: south, midwest, and west. (To save space, the estimated coefficients for the constant and the regional dummies are not shown in the table.) Because the regional dummies are held constant, the effect of initial per capita income does not reflect purely regional differences, such as the southern states' catching up with the northern states.⁷

For the longest interval, 1880–1988 (for 47 observations), the estimated convergence coefficient shown in line 1 of table 1 is $\hat{\beta} = .0175$ (standard error = .0046). Figure 1 shows the dramatic inverse relation between the average growth rate from 1880 to 1988 and $\log(y_{1880})$: the simple correlation is $-.93$.

The full time series for y_{it} (1880, 1900, 1920, and annually from 1929) potentially provides more information about the coefficient β . For a smaller value of T , however, the error term in equation (15), u_{i,t_0,t_0+T} , represents an average of shocks over a shorter interval. Therefore, the estimates become more sensitive to the specification of the error process. In particular, if there is serial persistence in the error term, u_{it} , then the correlation between u_{i,t_0,t_0+T} and $\log(y_{i,t_0})$ is likely to be negligible for large T but substantial for small T . For this reason, we have not attempted to use the full annual time series that starts in 1929.

Lines 2–10 of table 1 show estimates of β for nine subperiods of the overall sample: 1880–1900, 1900–1920, 10-year intervals from 1920 to 1980, and 1980–88. (There are 47 observations for the first subperiod and 48 for the others.) Each regression includes a constant and the three regional dummies. The results show values of $\hat{\beta}$ that range from $-.0122$ (.0074) for 1920–30 to $.0373$ (.0053) for 1940–50.

If all nine subperiods are restricted to have a single value for β , then the estimate is $\hat{\beta} = .0189$ (.0019) in line 11. This estimation allows each subperiod to have individual coefficients for the constant and the regional dummies.⁸ The joint estimate of β is close to the value $.0175$ estimated for the single interval 1880–1988. But, as would be expected, the standard error from the joint estimation,

⁶ See App. A for a discussion of the effects of measurement error in y_{it} on the estimates of β .

⁷ The estimated β convergence across regions turns out to be similar to that within regions (see Barro and Sala-i-Martin 1991a).

⁸ It would be possible to restrict the constants if it were maintained that each state experienced exogenous technological progress at the constant rate x . We could then use the whole sample to estimate a single constant and the value of x . We have not imposed these restrictions because we have no reason to think that the rate of technological change would be the same over all time periods.

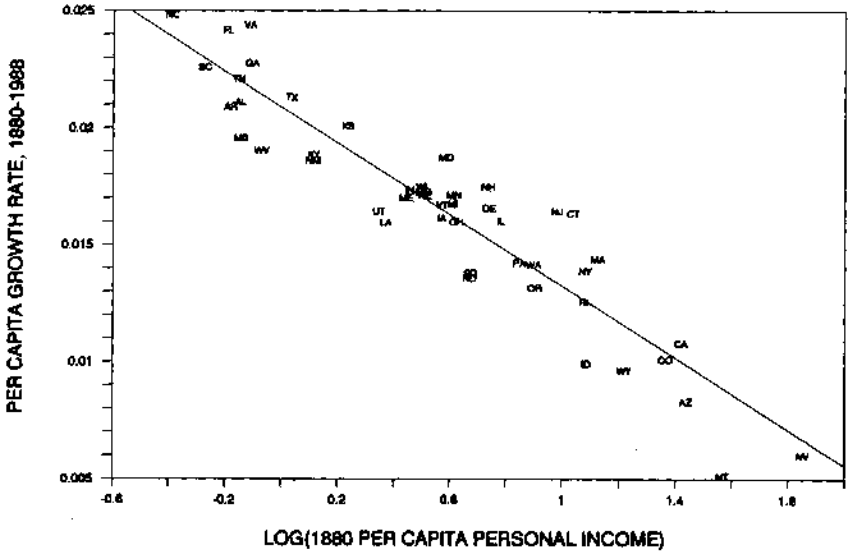


FIG. 1.—Growth rate from 1880 to 1988 vs. 1880 per capita income

.0019, is a good deal smaller than that, .0046, found for the single interval. The problem with the joint estimate is that the data reject the hypothesis that the coefficient β is the same for the nine subperiods. The likelihood ratio statistic for this hypothesis, 32.1, is well above the 5 percent critical value from the χ^2 distribution with eight degrees of freedom of 15.5 (p -value = .000).

The unstable pattern of $\hat{\beta}$ coefficients across subperiods can reflect aggregate disturbances that have differential effects on state incomes, as represented by the term $\phi_i s_i$ in equation (14). For example, during the 1920s, the ratio of the wholesale price index for farm products to the overall consumer price index fell at an average annual rate of 3.5 percent. The agricultural states also had below-average per capita personal income in 1920: the correlation of $\log(y_{1920})$ with the share of national income originating in agriculture in 1920 was $-.67$. Thus the estimated coefficient, $\hat{\beta} = -.0122$, for the 1920–30 period in table 1 likely reflects the tendency of the poorer states to be agricultural and therefore to experience relatively low growth in this decade. This effect reverses for the 1940–50 decade, when the ratio of the wholesale price index for farm products to the overall consumer price index grew at an average annual rate of 9.5 percent.

To hold constant this type of effect, we construct a variable that measures the sectoral composition of income in each state. For the subperiods that begin since 1930, we use a breakdown of the sources of labor income (including income from self-employment) into nine

categories: agriculture; mining; construction; manufacturing; transportation and public utilities; wholesale and retail trade; finance, insurance, and real estate; services; and government and government enterprises. For each subperiod, we construct a sectoral composition variable for state i :

$$s_{it} = \sum_{j=1}^9 w_{ijt} \cdot \log\left(\frac{y_{j,t+T}}{y_{jt}}\right), \quad (16)$$

where w_{ijt} is the weight of sector j in state i 's personal income at time t and y_{jt} is the national average of personal income that originates in sector j at time t , expressed as a ratio to national population at time t . Aside from the effect of changing sectoral weights within a state, the variable s_{it} would equal the growth rate of per capita personal income in state i between years t and $t + T$ if each of the state's sectors grew at the national average rate for that sector. In particular, the variable reflects shocks to agriculture, oil, and so forth in a way that interacts with state i 's concentration in the sectors that do relatively well or badly in terms of income because of the shocks.

We think of the variable s_{it} as a proxy for common effects related to sectoral composition in the error term in equation (15). Note that s_{it} depends on contemporaneous realizations of national variables, but only on lagged values of state variables. Because the impact of an individual state on national aggregates is small, s_{it} can be nearly exogenous with respect to the current individual error term for state i . In any event, we assume that, with s_{it} held constant, the error terms are independent across states and over time.

For the subperiods that begin before 1930, we lack detailed data on the sectoral composition of personal income, but we have data on the fraction of national income originating in agriculture. For these subperiods, we use this fraction as a measure of s_{it} . Note that the different methods of construction and the differing behavior of agricultural relative prices mean that the coefficients of the variable s_{it} will vary from one subperiod to another. Therefore, we estimate a separate coefficient on s_{it} for each subperiod.

Lines 12–20 in table I add the variable s_{it} to the growth rate regressions for each subperiod. (The first subperiod has 46 observations and the others have 48.) As before, these regressions include $\log(y_{i,t_0})$, a constant, and three regional dummies. Not surprisingly, the estimated coefficients on the variable s_{it} for the post-1930 subperiods are typically positive. That is, states in which income originates predominantly in sectors that do well at the national level tend to have higher per capita growth rates. (The estimated coefficient for the 1940–50 subperiod is negative, but not significantly so.) For the

subperiods that begin before 1930, the negative estimated coefficient on s_{it} signifies that, with initial per capita income and region held constant, agricultural states have lower per capita growth rates. This pattern is especially clear for the agricultural price collapse in the 1920–30 decade: the estimated coefficient on s_{it} is $-.0936$ (.0175).

For our purposes, the principal finding from the addition of the sectoral composition variables is that the estimated β coefficients become much more stable across subperiods. The range is now .0139 (.0076) for 1970–80 to .0362 (.0055) for 1940–50. Line 21 shows that the jointly estimated coefficient for the nine subperiods is .0249 (.0021). (This joint estimation allows each subperiod to have individual coefficients for s_{it} as well as for the constant and the regional dummies.) The likelihood ratio statistic for the equality of β coefficients across the nine subperiods is now 13.9, compared to the 5 percent critical value of 15.5. Thus if we hold constant the measures of sectoral composition, we no longer reject the hypothesis of a single β coefficient at the 5 percent level (p -value = .084).

The agriculture share variable, which was included to measure s_{it} for the earlier subperiods in table 1 (lines 12–14 and the joint estimate in line 21), holds constant compositional effects on aggregate state income that reflect shifts of persons out of agriculture and into higher-productivity jobs in industry and services. If we add the agriculture share variable to the later subperiods, then the joint estimate for nine subperiods becomes $\hat{\beta} = .0224$ (.0022), slightly less than the value shown in line 21. This estimate of β is virtually unchanged if we include the change in the agriculture share over each subperiod in the regressions. Thus convergence at a rate of about 2 percent per year is net of effects from changes in agricultural shares.

In general, industry mix effects would matter for the results if changes in income shares among sectors with different average levels of productivity are correlated with initial levels of per capita income. It is unclear that we would want to filter out all these effects to measure convergence, but, in any event, our examination of productivity data from the post–World War II period indicates that shifts between agriculture and nonagriculture would be the main effect of this type. Since we already held constant the compositional effect for agriculture, it is unlikely that industry mix effects are a major element in the estimated convergence for state personal income.

The final result from table 1 is a regression with the 29 available observations from 1840 to 1880.⁹ This regression includes a constant

⁹ Easterlin (1960a, p. 124 ff.) indicates that the data for 1840 do not cover income originating in wholesale and retail trade; finance, insurance, and real estate; government; and most other services. The figures that we use for 1880 in the 1840–80 regressions are comparable in coverage to those for 1840. This more limited coverage

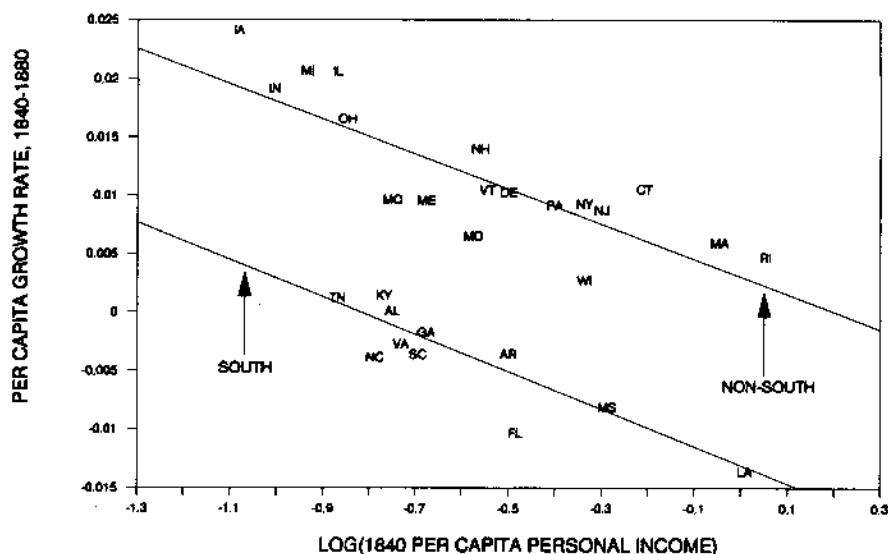


FIG. 2.—Growth rate from 1840 to 1880 vs. 1840 per capita income

and two regional dummies (no western states are in the sample). We exclude the variable s_{it} because the data are unavailable. The estimate in line 22 is $\hat{\beta} = .0254$ (.0067), which accords with the estimate of .0249 (.0021) for the subperiods that begin after 1880 (line 21).

Figure 2 plots the per capita growth rate from 1840 to 1880 against $\log(y_{1840})$. A remarkable aspect of the plot is the separation of the southern and nonsouthern states because of the Civil War. In 1840, the southern and nonsouthern states differed little in terms of average per capita income: the unweighted average of 11 southern states was 94 percent of that for 18 eastern and midwestern states. But in 1880 a wide gap had appeared and the southern average was only 50 percent of the nonsouthern. The figure shows, however, that convergence applies to the southern and nonsouthern states as separate groups. That is, with the regional dummies held constant (which effectively hold constant the impact of the Civil War), there is a strong negative correlation between the per capita growth rate and the initial level of per capita income.

The Civil War affected states differentially, but, in contrast to the shock to agriculture in the 1920s, the effect of the Civil War on state per capita income had little correlation with the initial level of per

for 1880 comprises about half the income included in the measure that we used previously. In any event, the limited figures for 1840 are not comparable to the data for years after 1880.

capita income. For this reason, we do not get a very different point estimate of β for the 1840–80 subperiod if we eliminate the regional dummies: the estimate without these dummies is $\hat{\beta} = .0203$ (.0126). The fall in the R^2 of the regression from .91 in line 22 of table 1 to .19 indicates, however, that the regional dummies have a lot of explanatory power in this period!

Results with Gross State Product

Table 2 and figure 3 deal with the growth of per capita GSP for 48 states over the period 1963–86. Recall that the data are nominal GSP divided by an aggregate, national price deflator. The growth rates

TABLE 2
CROSS-STATE REGRESSIONS FOR GROSS STATE PRODUCT

Sample	$\hat{\beta}$	Sectoral Composition (s_{it})	R^2	σ
1. 1963–86	.0180 (.0059)48	.0038
2. 1963–69	.0154 (.0060)63	.0056
3. 1969–75	.0406 (.0162)41	.0120
4. 1975–81	-.0285 (.0130)17	.0139
5. 1981–86	.1130 (.0244)62	.0168
6. Four periods, β restricted*	.0211 (.0053)
7. 1963–69	.0157 (.0060)	.18 (.25)	.63	.0056
8. 1969–75	.0297 (.0101)	1.56 (.20)	.74	.0081
9. 1975–81	.0258 (.0108)	1.74 (.15)	.78	.0072
10. 1981–86	.0238 (.0091)	1.73 (.13)	.92	.0079
11. Four periods, β restricted*	.0216 (.0042)	individual
12. 1963–86	.0222 (.0065)	.63 (.27)	.54	.0036

NOTE.—All regressions have 48 observations. The dependent variable is the growth rate of real per capita GSP (nominal GSP per capita divided by the national deflator for GSP). The regressions denoted four periods, β restricted use nonlinear, iterative weighted least squares, with the coefficient β constrained to be equal for the four subperiods. See also the notes to table 1.

* For line 6, the log likelihood ratio is 31.2 (p -value = .000); for line 11, it is 1.7 (p -value = .637). Under the null hypothesis of equal coefficients, the likelihood ratio statistic is distributed as χ^2 with three degrees of freedom.

and thereby leads to the negative value for $\hat{\beta}$, $-.0285$, shown for 1975–81 in line 4 of table 2.¹⁰ But when sectoral composition is held constant in line 9, the value of $\hat{\beta}$ for 1975–81 is similar to that found in the other periods.

For the 1981–86 period, the key elements are the sharp decline in oil prices and the high correlation, $.7$, between per capita GSP and the share of GSP originating in oil and natural gas in 1981. The tendency for oil states to do relatively badly in 1981–86 leads to an exaggerated convergence coefficient, $\hat{\beta} = .1130$, in line 5.¹¹ Again, the inclusion of the variable s_{it} in line 9 leads to a normal value for $\hat{\beta}$.

Barro and Sala-i-Martin (1991a) disaggregate the nonagricultural part of GSP into value added per worker for eight sectors. The main finding is that convergence shows up significantly within these sectors of production, especially for manufacturing. For the nonmanufacturing sectors, the overall estimate of β is somewhat less than 2 percent per year, whereas for manufacturing the estimate is over 4 percent per year. The main inference from these results is that poorer states grow faster not only in terms of overall GSP per person, but also in terms of labor productivity within various sectors of production. Thus, as suggested before for personal income, the findings on convergence cannot be explained by changes over time in the composition of production.

Income versus Product

In a closed-economy growth model, the convergence properties of income and product must coincide. Perhaps surprisingly—because the U.S. states do not look like closed economies—the empirical estimates of β for personal income are nearly equal to those for GSP. If the estimation for personal income is limited to a time span similar to that covered by GSP—namely the three subperiods 1960–70, 1970–80, and 1980–88—then the joint estimate is $\hat{\beta} = .0181$ (.0040). Although this point estimate is less than that, $.0216$ (.0042), shown for GSP in table 2, line 11, the principal finding is that the estimates are close.

The assumptions of a closed economy are implausible for the U.S.

¹⁰ This argument does not apply to the subperiod 1969–75 (line 3 of table 2). Although the oil price rose substantially over this period, the oil states did not have especially high values of per capita GSP in 1969.

¹¹ The results for personal income over the period 1980–88 (table 1, line 10) do not show the same pattern. The main difference is that the correlation in 1980 of the logarithm of per capita personal income with the share of income originating in oil and natural gas is close to zero.

states: goods and technologies flow across borders, residents of one state can borrow from residents of other states, and internal migration is possible. In Barro and Sala-i-Martin (1991*b*, chap. 2), we extend the neoclassical growth model to allow for internationally tradable goods and a global capital market. These features create a sharp distinction between domestic product and income or, equivalently, between domestic capital stock and assets. If technologies are the same, then an economy's per capita capital stock and output converge rapidly to those prevailing in other economies. In contrast, even if all economies have the same parameters of preferences and technologies, per capita incomes do not converge because each small economy faces constant returns on the global capital market. Thus our empirical findings—that rates of convergence are similar for income and product across the U.S. states—are puzzling from the perspective of this theory. We offer here some conjectures that may help to resolve this puzzle.

We have modified the analysis along the lines of Cohen and Sachs (1986) to allow for a ceiling on the ratio of an economy's external debt to its capital stock. This restriction on credit markets is reasonable if the capital stock represents the collateral that secures the debt. If we interpret capital broadly to include human capital, then this framework applies to the U.S. states if the residents or government of a state cannot borrow nationally to finance all their desired expenditures on education or other forms of investment in human capital. The key result from the addition of the borrowing constraint is that domestic product behaves eventually like national income. Hence, the convergence properties of product and income can be similar, as in our empirical results.

If technologies (i.e., anything represented by the coefficient A in eq. [7]) differ across economies, then mobility of capital can create divergence of per capita output and capital stocks. Economies with higher \hat{k} tend to have higher values of A , and the higher A offsets the effect of diminishing returns in the determination of capital's marginal product. Therefore, capital (physical or human) may move from poorer to richer economies, and it is no longer clear theoretically that the convergence coefficient for product would exceed that for income. Once we allow for differences in technologies, we also have to consider the diffusion of technology across economies, along the lines of Nelson and Phelps (1966). The potential to imitate is another reason for poor, follower economies to grow at relatively high rates.

We have extended the neoclassical growth model to allow for migration of persons, another force that promotes convergence of per capita product and income across economies. Sala-i-Martin (1990,

table 5.2) and Barro and Sala-i-Martin (1991a) relate net migration for the U.S. states to initial values of per capita personal income over subperiods of the interval from 1900 to 1987. These studies confirm that net in-migration is positively related to initial per capita income. But the results also show that the estimated convergence coefficients, $\hat{\beta}$, are little affected by the inclusion of net migration as an explanatory variable in the growth rate equations. Moreover, we have shown that the minor interplay between migration and convergence is quantitatively consistent with the neoclassical growth model (extended to allow for migration), given the estimated sensitivity of migration to income differentials.

We leave as an unresolved puzzle the similar estimates for the rates of convergence of per capita income and product. We think that a resolution of this puzzle will involve the construction of an open-economy growth model that satisfactorily incorporates credit markets, factor mobility, and technological diffusion.

Comparisons with Findings across Countries

In this section we compare our findings for the U.S. states with analogous results across countries. It is well known that growth rates of real per capita GDP are uncorrelated with the starting level of real per capita GDP across a large group of countries in the post-World War II period. Barro (1991) uses the Summers-Heston (1988) data set along with other data to analyze the growth experiences of 98 countries from 1960 to 1985. (The limitation to 98 countries rather than the 114 market economies with Summers-Heston GDP data from 1960 to 1985 comes from the lack of information on variables other than GDP.) Line 1 of table 3 shows that a regression for the 98 countries in the form of equation (15) leads to the estimate $\hat{\beta} = -.0037$ (.0018). The dependent variable is the growth rate of real per capita GDP from 1960 to 1985. The only independent variables are a constant and the log of 1960 per capita GDP, $\log(y_{1960})$. The main finding, also depicted in figure 4, is the lack of a close relationship between the growth rate and $\log(y_{1960})$. In fact, the convergence coefficient $\hat{\beta}$ has the wrong sign; that is, there is a small tendency for the initially rich countries to grow faster than the poor ones after 1960.

These cross-country results contrast sharply with the findings discussed earlier for the U.S. states. Figures 1 and 3 and tables 1 and 2 showed that, particularly over the longer samples, there is a clear and substantial negative correlation between starting per capita income or product and the subsequent growth rate. Line 5 of table 3 uses a specification for the U.S. states that parallels the one used for

TABLE 3
COMPARISON OF REGRESSIONS ACROSS COUNTRIES AND U.S. STATES

Sample	$\hat{\beta}$	Additional Variables	R^2	σ
1. 98 countries, 1960-85	-.0037 (.0018)	no	.04	.0183
2. 98 countries, 1960-85	.0184 (.0045)	yes	.52	.0133
3. 20 OECD countries, 1960-85	.0095 (.0028)	no	.45	.0051
4. 20 OECD countries, 1960-85	.0203 (.0068)	yes	.69	.0046
5. 48 U.S. states, 1963-86	.0218 (.0053)	no	.38	.0040
6. 48 U.S. states, 1963-86	.0236 (.0013)	yes	.61	.0033

NOTE.—The dependent variable in regressions 1-4 is the growth rate of real per capita GDP from 1960 to 1985; in regressions 5 and 6 it is the growth rate of real per capita GSP (the variable used in table 2) from 1963 to 1986. The coefficient β applies in regressions 1-4 to the logarithm of real per capita GDP in 1960, and in regressions 5 and 6 to the logarithm of real per capita GSP in 1963. Each regression also includes a constant. The additional variables included in regressions 2 and 4 are the primary and secondary school enrollment rates in 1960, the average ratio of government consumption expenditure (standard figures less spending on defense and education) to GDP from 1970 to 1985, the average number of revolutions and coups per year from 1960 to 1985, the average number of political assassinations per capita per year from 1960 to 1985, and the average deviation from unity of the Summers-Heston (1988) purchasing power parity ratio for investment in 1960. See Barro (1991) for details on these variables. The additional explanatory variables included in regression 6 are regional dummies, the sectoral composition variable, δ_{it} , and the fraction of workers in 1960 that had accumulated some amount of college education. The 20 OECD countries (the original membership in 1960) are Austria, Belgium, Canada, Denmark, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom, and United States.

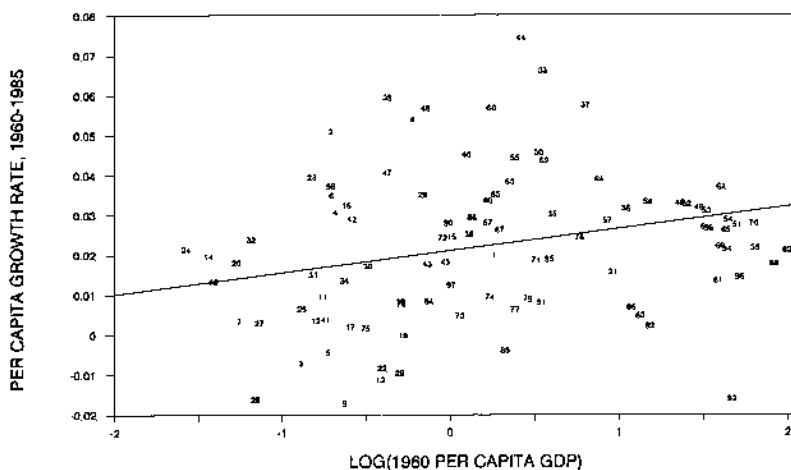


FIG. 4.—Growth rate from 1960 to 1985 vs. 1960 per capita GDP, sample of 98 countries (listed in App. B).

the countries. The variables are based on GSP over the time period 1963–86, and the regression includes only $\log(y_{1963})$ and a constant as regressors. The estimate in this case is $\hat{\beta} = .0218$ (.0053).

Barro (1991, table 1, fig. 2) shows that a significantly negative partial relation between the per capita growth rate from 1960 to 1985 and $\log(y_{1960})$ emerges for the 98 countries if some other variables are held constant. The set of other variables in the main results consists of primary and secondary school enrollment rates in 1960, the average ratio of government consumption expenditure (exclusive of defense and education) to GDP from 1970 to 1985, proxies for political stability, and a measure of market distortions based on purchasing power parity ratios for investment goods. If we include these variables for the 98 countries in the form of equation (15), then line 2 of table 3 shows that the estimated convergence coefficient becomes $\hat{\beta} = .0184$ (.0045). This estimate of β is no longer very much below the cross-state value shown in line 5 of the table.

The theoretical relation in equation (15) predicts conditional convergence, that is, a negative relation between $\log(y_{i,t_0})$ and the subsequent growth rate if we hold constant the steady-state position, $\log(\hat{y}_i^*)$, and the steady-state growth rate, x_i . (The constant B in eq. [15] depends on $\log(\hat{y}_i^*)$ and x_i .) The theory implies that the relation between $\log(y_{i,t_0})$ and the growth rate will be negative unless the correlation between $\log(y_{i,t_0})$ and the two omitted factors, $\log(\hat{y}_i^*)$ and x_i , is substantially positive.

The U.S. states are likely to be reasonably homogeneous with respect to the steady-state values $\log(\hat{y}_i^*)$ and x_i . That is, the differences in initial positions, $\log(y_{i,t_0})$, may be relatively much greater. (This condition is especially compelling if the initial differences reflect exogenous events, such as wars, world agricultural harvests, and oil shocks.) In this case, the negative relation between the growth rate and $\log(y_{i,t_0})$ would show up even if the differences in the steady-state values are not held constant: conditional and absolute convergence would coincide. The result for $\hat{\beta}$ shown in line 5 of table 3 is consistent with this perspective.

In contrast, the sample of 98 countries likely features large differences in the steady-state values, $\log(\hat{y}_i^*)$ and x_i , that is, in the underlying parameters of technology and preferences (and natural resources and government policies) that determine these long-run values. The absence of substantial labor mobility across countries reinforces the possibility of substantial divergences in these steady-state values. The correlation of $\log(y_{i,t_0})$ with $\log(\hat{y}_i^*)$ is likely to be substantially positive; that is, economies with higher steady-state values of output per effective worker would have followed a path that led them today to higher levels of output per person. Similarly, the correlation of $\log(y_{i,t_0})$ with

x_i is likely to be positive. Because of these relations, the simple correlation between the growth rate and $\log(y_{i,t_0})$ could be close to zero, as indicated by the data in figure 4. (This point is made by King and Rebelo [1989, pp. 12–13].) On the other hand, if we include additional variables that hold constant some of the cross-country variations in $\log(\hat{y}_i^*)$ and x_i , then the partial relation between growth rate and $\log(y_{i,t_0})$ should become more negative. We interpret the additional variables that we added to the cross-country regression (table 3, line 2) in this manner. Accordingly, we view the estimate of β in this regression—which is no longer very much below the values from the cross-state regressions—as coming closer to the theoretical convergence coefficient.

We can evaluate these arguments further by considering a group of relatively homogeneous countries, the 20 original members of the OECD.¹² Figure 5 shows that the per capita growth rate is negatively related to the log of initial per capita GDP for this group of countries. The regression in line 3 of table 3 includes only a constant and the log of 1960 per capita GDP. The estimated convergence coefficient is $\hat{\beta} = .0095$ (.0028), which is significant and has the expected sign. The magnitude is, however, about half that applicable to the U.S. states (line 5). Our interpretation is that the OECD countries are intermediate between the 98-country group and the U.S. states in terms of the extent of cross-country variation in steady-state values, $\log(\hat{y}_i^*)$ and x_i , relative to the variation in initial positions, $\log(y_{i,t_0})$. Line 4 of the table shows that the estimate for the OECD countries becomes $\hat{\beta} = .0203$ (.0068) when the additional variables discussed before are added to the regression. This estimate does not differ greatly from the comparable value for 98 countries, .0184 (.0045) in line 2.

We have also explored in a preliminary way the addition of variables as proxies for the steady-state values, $\log(\hat{y}_i^*)$ and x_i , in the cross-state regressions. One variable that has a significantly positive influence on the growth rate is the fraction of the work force in 1960 that had accumulated some amount of college education.¹³ We added this variable along with the regional dummies and sectoral composition variable, s_{it} , that we discussed before. Line 6 of table 3 shows that the estimated convergence coefficient becomes $\hat{\beta} = .0236$ (.0013), compared with .0218 (.0053) in line 5. Thus the inclusion of these

¹² We exclude the four countries added after 1960 (Australia, Finland, Japan, and New Zealand) because of the possibility that the extension of membership was endogenous and related to the growth experience.

¹³ The data on educational attainment come from various issues of *Statistical Abstract*. We have not had much success in finding growth rate effects related to cross-state differences in government expenditures. Also, educational differences aside from college attainment were not important.

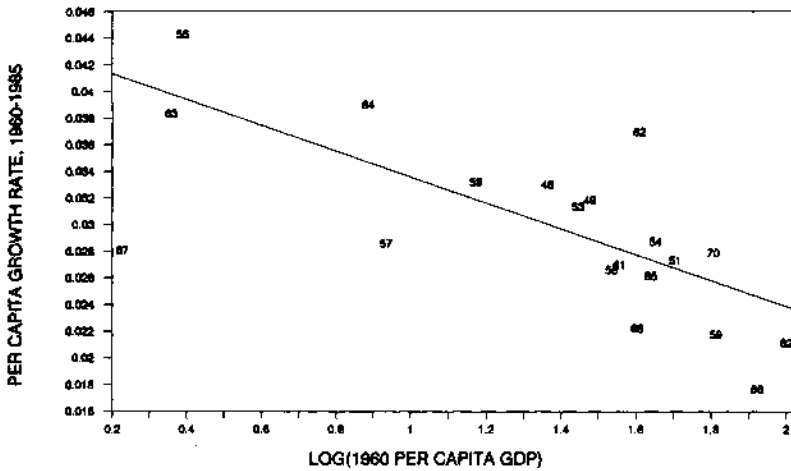


FIG. 5.—Growth rate from 1960 to 1985 vs. 1960 per capita GDP, OECD countries (listed in App. B).

other variables has a positive, but minor, effect on the estimate of β across the U.S. states.

Overall, the impact of the additional variables on $\hat{\beta}$ is greatest for the 98 countries (.0184 in line 2 vs. -.0037 in line 1), next most important for the 20 OECD countries (.0203 in line 4 vs. .0095 in line 3), and least important for the 48 U.S. states. These findings are consistent with the idea that, first, the other variables help to hold constant cross-sectional differences in the long-run values, $\log(\hat{y}_i^*)$ and x_i , and, second, that the ranking of the extent of these differences (relative to the differences in $\log[y_{i,t_0}]$) goes from the 98 countries to the 20 OECD countries to the 48 U.S. states.

Conclusions

Our empirical results document the existence of convergence in the sense that economies tend to grow faster in per capita terms when they are further below the steady-state position. This phenomenon shows up clearly for the U.S. states over various periods from 1840 to 1988. Over long samples, poor states tend to grow faster in per capita terms than rich states even if we do not hold constant any variables other than initial per capita income or product. If we hold constant the region and measures of sectoral composition, then the speed of convergence appears to be roughly the same—around 2 percent per year—regardless of the time period or whether we consider personal income or GSP.

We find evidence of convergence for a sample of 98 countries from 1960 to 1985 only in a conditional sense, that is, only if we hold constant variables such as initial school enrollment rates and the ratio of government consumption to GDP. We interpret these variables as proxies for the steady-state value of output per effective worker and the rate of technological progress. If we hold constant these additional variables, then the estimated rates of convergence are only slightly smaller than those found for the U.S. states.

The standard neoclassical growth model with exogenous technological progress and a closed economy predicts convergence. To match our quantitative estimates, however, we have to assume underlying parameters for preferences and technology that depart substantially from usual benchmark cases. In particular, for reasonable values of the other parameters, the model requires a capital share coefficient, α , in the neighborhood of .8. Lower values of α , which imply that diminishing returns to capital set in more quickly, imply a more rapid rate of convergence than that revealed by the data.

If technologies are the same, then the introduction of a global capital market tends to speed up the convergence for output but to slow down the convergence for income. The empirical results for the U.S. states indicate that the speed of convergence for output is only slightly faster than that for income. At this point, we can reconcile this finding with the theory only if we include elements of capital market imperfections, such as a limited ability to borrow to finance accumulations of human capital. Other elements of an open economy—the mobility of labor and technology—tend to speed up the predicted rate of convergence. Therefore, we require even higher values of the capital share parameter, α , to match the empirical results.

Some recent models of endogenous economic growth, such as Rebelo (1991), assume constant returns to a broad concept of capital that includes human capital. This specification corresponds to $\alpha = 1.0$ in the neoclassical model. As mentioned, our empirical results indicate that the neoclassical model requires a value of α of about .8 to fit the observed speeds of convergence. The difference between $\alpha = .8$, where diminishing returns to capital set in slowly, and $\alpha = 1.0$, where diminishing returns do not set in at all, may seem to be minor. But the difference amounts to a half-life of 27 years in the former case versus infinity in the latter. To put it another way, the convergence coefficient $\beta = 2$ percent per year, corresponding to $\alpha \approx .8$, implies that the poor countries of sub-Saharan Africa should have experienced growth of real per capita GDP from 1960 to 1985 at an average rate above 6 percent per year, compared to 2 percent per year for the United States, *if the African countries were approaching the*

same steady-state path as that for the United States. (The actual average growth rate of 0.8 percent per year for the sub-Saharan African countries is "explained" in the regression in line 2 of table 3 by the additional variables that proxy for steady-state positions.) The main point is that a value for α of .8 is very far from 1.0 in an economic sense.

In open-economy versions of the neoclassical growth model, it is possible to find convergence effects associated with technological diffusion even if the returns to capital are constant ($\alpha = 1$). Also, in closed-economy models with constant returns to a broad concept of capital, convergence effects can reflect the working out of initial imbalances among the various kinds of capital. For example, Mulligan and Sala-i-Martin (1991) show that the per capita growth rate is inversely related to initial physical capital per worker for a given initial quantity of human capital per worker. Thus we would like to break down the observed convergence into various components: first, effects related to diminishing returns to capital and to imbalances among types of capital in the context of a closed economy; second, effects involving the mobility of capital and labor across economies; and third, effects that involve the gradual spread of technology. The present empirical results, which exploit only cross-sectional differences in growth rates, do not allow us to separate the observed convergence patterns into these components. We hope to make these distinctions in future research, which will also exploit the time-series variations of growth rates.

Appendix A

Some Effects of Measurement Error

The regressions shown in tables 1 and 2 can exaggerate the estimated convergence coefficient, $\hat{\beta}$, if real income or product is measured with error. Aside from the usual measurement problems, one reason to expect errors is that we divide all nominal values in each year by a common price index.

Equation (15) can be rewritten as

$$\frac{1}{T} \cdot \log(y_{i,t_0+T}) = B + \frac{e^{-\beta T}}{T} \cdot \log(y_{i,t_0}) + u_{i,t_0,t_0+T} \quad (A1)$$

Assume that the observed value at date t , $\log(\hat{y}_i)$, differs from the true value, $\log(y_i)$, by a random measurement error:

$$\log(\hat{y}_i) = \log(y_i) + \eta_{it} \quad (A2)$$

For purely temporary measurement error, η_{it} would be white noise. Then, as is well known, the measurement error in $\log(\hat{y}_{i,t_0})$ implied by equation (A2) leads to a bias toward zero in least-squares estimation of the coefficient, $e^{-\beta T}/T$, in equation (A1). Because the term $e^{-\beta T}/T$ in equation (A1) is decreasing in β , the nonlinear estimate $\hat{\beta}$ provides a corresponding overestimate of β in large samples.

We can obtain a bound for the inconsistency induced by temporary measurement error. Equation (11) implies that the growth rate of income between any two future dates, $t_0 + \tau$ and $t_0 + T$, is given by

$$\frac{1}{T - \tau} \cdot \log \left(\frac{y_{i,t_0+T}}{y_{i,t_0+\tau}} \right) = B - \frac{e^{-\beta\tau} - e^{-\beta T}}{T - \tau} \cdot \log(y_{i,t_0}) + u_{i,t_0+\tau,t_0+T} \quad (A3)$$

where $T > \tau > 0$ and $u_{i,t_0+\tau,t_0+T}$ depends on the error terms, u_{it} , between dates $t_0 + \tau$ and $t_0 + T$. Equation (A3) relates the growth rate from $t_0 + \tau$ to $t_0 + T$ to the level of per capita income or product at an earlier time, t_0 . Note that equation (15) is the special case in which $\tau = 0$.

We assume that the measurement error, η_{i,t_0} , is independent of η_{i,t_0+t} for $t \geq \tau$. This condition holds for all $\tau > 0$ if η_{it} is white noise but also applies for large enough τ to measurement error with some persistence over time. We assume that η_{i,t_0} is independent of $u_{i,t_0+\tau,t_0+T}$. In this case, least-squares estimation of equation (A3) leads to an underestimate of the magnitude of the coefficient, $(e^{-\beta\tau} - e^{-\beta T})/(T - \tau)$. We can show that this term is increasing in β if $\beta < [\log(T/\tau)]/(T - \tau)$. In practice, we use the values $\tau = 10$ years and $T = 20$ years or $\tau = 5$ years and $T = 10$ years. For the first pair of values, the term $(e^{-\beta\tau} - e^{-\beta T})/(T - \tau)$ is increasing in β if $\beta < .07$ per year; for the second pair, the term is increasing in β if $\beta < .14$ per year. Therefore, for these ranges of β and in large samples, the underestimate of the coefficient on $\log(y_{i,t_0})$ in equation (A3) corresponds to a large-sample underestimate of β . Because this bias is opposite in direction to that found for equation (15), we can use regressions in the form of equation (A3) to bound the size of the bias.

Consider the regressions for personal income in which each subperiod has individual coefficients for the constant, three regional dummies, and the sectoral composition variable, s_{it} . If we use only the five equal-length subperiods from 1930–40 to 1970–80, then the joint estimate β in the form of equation (15) is .0244 (.0025), which is close to the value for nine subperiods from 1880 to 1988 shown in line 20 of table 1. The comparable result in the form of equation (A3) with $\tau = 10$ years and $T = 20$ years is $\beta = .0278$ (.0049). Although we expected the asymptotic bias induced by temporary measurement error to be positive in the first case and negative in the second, the result for β turns out to be higher in the second case. (The theoretical result can be affected by the inclusion of additional explanatory variables in the regressions.) In any event, we infer from the similarity of the two estimates of β that temporary measurement error is unlikely to have a major influence on the results.

For GSP, we use the three equal-length subperiods 1970–75, 1975–80, and 1980–85. The joint estimate β in the form of equation (15) is .0280 (.0058), somewhat higher than that, .0216 (.0042), shown for four subperiods from 1963 to 1986 in line 11 of table 2. With $\tau = 5$ years and $T = 10$ years, joint estimation of equation (A3) over the three subperiods from 1970 to 1985 leads to the estimate $\beta = .0366$ (.0091). Again, in contrast to expectations, the estimated value in the second case exceeds that in the first case. But the main inference is that the results are similar and, hence, that temporary measurement error is unlikely to be important.

Appendix B**Key for Countries in Figures 4 and 5**

- | | |
|-----------------------------|-------------------------|
| 1. Algeria | 51. Denmark |
| 2. Botswana | 52. Finland |
| 3. Burundi | 53. France |
| 4. Cameroon | 54. Germany |
| 5. Central African Republic | 55. Greece |
| 6. Egypt | 56. Iceland |
| 7. Ethiopia | 57. Ireland |
| 8. Gabon | 58. Italy |
| 9. Ghana | 59. Luxembourg |
| 10. Ivory Coast | 60. Malta |
| 11. Kenya | 61. Netherlands |
| 12. Liberia | 62. Norway |
| 13. Madagascar | 63. Portugal |
| 14. Malawi | 64. Spain |
| 15. Mauritius | 65. Sweden |
| 16. Morocco | 66. Switzerland |
| 17. Nigeria | 67. Turkey |
| 18. Rwanda | 68. United Kingdom |
| 19. Senegal | 69. Barbados |
| 20. Sierra Leone | 70. Canada |
| 21. South Africa | 71. Costa Rica |
| 22. Sudan | 72. Dominican Republic |
| 23. Swaziland | 73. El Salvador |
| 24. Tanzania | 74. Guatemala |
| 25. Togo | 75. Haiti |
| 26. Tunisia | 76. Honduras |
| 27. Uganda | 77. Jamaica |
| 28. Zaire | 78. Mexico |
| 29. Zambia | 79. Nicaragua |
| 30. Zimbabwe | 80. Panama |
| 31. Bangladesh | 81. Trinidad and Tobago |
| 32. Burma | 82. United States |
| 33. Hong Kong | 83. Argentina |
| 34. India | 84. Bolivia |
| 35. Iran | 85. Brazil |
| 36. Israel | 86. Chile |
| 37. Japan | 87. Colombia |
| 38. Jordan | 88. Ecuador |
| 39. Korea | 89. Guyana |
| 40. Malaysia | 90. Paraguay |
| 41. Nepal | 91. Peru |
| 42. Pakistan | 92. Uruguay |
| 43. Philippines | 93. Venezuela |
| 44. Singapore | 94. Australia |
| 45. Sri Lanka | 95. Fiji |
| 46. Taiwan | 96. New Zealand |
| 47. Thailand | 97. Papua New Guinea |
| 48. Austria | 98. Indonesia |
| 49. Belgium | |
| 50. Cyprus | |

References

- Barro, Robert J. "Economic Growth in a Cross Section of Countries." *Q.J.E.* 106 (May 1991): 407-43.
- Barro, Robert J., and Sala-i-Martin, Xavier. "Convergence across States and Regions." *Brookings Papers Econ. Activity*, no. 1 (1991), pp. 107-82. (a)
- . "Economic Growth." Manuscript. Cambridge, Mass.: Harvard Univ., 1991. (b)
- Blanchard, Olivier J., and Fischer, Stanley. *Lectures on Macroeconomics*. Cambridge, Mass.: MIT Press, 1989.
- Bureau of Economic Analysis. *State Personal Income by State: 1929-1982*. Washington: Government Printing Office, 1986.
- Cass, David. "Optimum Growth in an Aggregative Model of Capital Accumulation." *Rev. Econ. Studies* 32 (July 1965): 233-40.
- Chamley, Christophe. "The Welfare Cost of Capital Income Taxation in a Growing Economy." *J.P.E.* 89 (June 1981): 468-96.
- Cohen, Daniel, and Sachs, Jeffrey. "Growth and External Debt under Risk of Debt Repudiation." *European Econ. Rev.* 30 (June 1986): 529-60.
- Easterlin, Richard A. "Interregional Differences in Per Capita Income, Population, and Total Income, 1840-1950." In *Trends in the American Economy in the Nineteenth Century*. Princeton, N.J.: Princeton Univ. Press (for NBER), 1960. (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 Soc., 1960. (b)
- Jorgenson, Dale W., and Yun, Kun-Young. "Tax Policy and Capital Allocation." *Scandinavian J. Econ.* 88, no. 2 (1986): 355-77.
- . "Tax Reform and U.S. Economic Growth." *J.P.E.* 98, no. 5, pt. 2 (October 1990): S151-S193.
- King, Robert G., and Rebelo, Sergio T. "Transitional Dynamics and Economic Growth in the Neoclassical Model." Working Paper no. 3185. Cambridge, Mass.: NBER, November 1989.
- Koopmans, Tjalling C. "On the Concept of Optimal Economic Growth." In *The Econometric Approach to Development Planning*. Amsterdam: North-Holland, 1965.
- Maddison, Angus. "Growth and Slowdown in Advanced Capitalist Economies: Techniques of Quantitative Assessment." *J. Econ. Literature* 25 (June 1987): 649-98.
- Mulligan, Casey B., and Sala-i-Martin, Xavier. "Two-Capital-Goods Models of Endogenous Growth: Steady State, Transitional Dynamics, and Empirical Implications." Manuscript. New Haven, Conn.: Yale Univ., June 1991.
- Nelson, Richard R., and Phelps, Edmund S. "Investment in Humans, Technological Diffusion, and Economic Growth." *A.E.R. Papers and Proc.* 56 (May 1966): 69-75.
- Ramsey, Frank P. "A Mathematical Theory of Saving." *Econ. J.* 38 (December 1928): 543-59.
- Rebelo, Sergio. "Long-Run Policy Analysis and Long-Run Growth." *J.P.E.* 99 (June 1991): 500-521.
- Renshaw, Vernon; Trott, E. A., Jr.; and Friedenberg, Howard L. "Gross State Product by Industry, 1963-1986." *Survey Current Bus.* 68 (May 1988): 30-46.

- Sala-i-Martin, Xavier. "On Growth and States." Ph.D. dissertation, Harvard Univ., 1990.
- Sato, Kazuo. "On the Adjustment Time in Neo-classical Growth Models." *Rev. Econ. Studies* 33 (July 1966): 263-68.
- Solow, Robert M. "A Contribution to the Theory of Economic Growth." *Q.J.E.* 70 (February 1956): 65-94.
- Summers, Robert, and Heston, Alan. "A New Set of International Comparisons of Real Product and Price Levels: Estimates for 130 Countries, 1950-1985." *Rev. Income and Wealth* 34 (March 1988): 1-25.