Reviewing the Evidence against Absolute Convergence

Rómulo A. Chumacero
University of Chile and Central Bank of Chile

Few subjects in applied economic research have been studied as extensively as the convergence hypothesis advanced by Solow (1956) and documented by Baumol (1986).[1] In simple terms, the hypothesis states that poor countries or regions tend to grow faster than rich ones. In its strongest version (known as absolute convergence), the hypothesis implies that in the long run, countries or regions should not only grow at the same rate, but also reach the same per capita income.[2] This hypothesis has been tested using different methodologies and datasets, and it appears to be strongly rejected by the data. In view of these results, several modifications of the absolute convergence hypothesis have been advanced and tested, although they usually lack both theoretical foundations and econometric rigor and discipline.

This paper analyzes whether the econometric methods usually applied to test for absolute convergence have given this hypothesis a fair chance. The paper is organized as follows. Section 1 presents a brief review of some of the tests for convergence advanced in the empirical literature and documents their shortcomings. Section 2 develops a simple theoretical model that implies absolute convergence. Section 3 discusses the likelihood that time series generated from the model can accommodate the results of the tests described on section 1. Finally, section 4 draws some conclusions.

I would like to thank William Easterly, Rodrigo Fuentes, Eduardo Morón, Klaus Schmidt-Hebbel, and Raimundo Soto for helpful comments and suggestions. Francisco Gallego provided able research assistance.

1. Representative studies in this line of research include Aghion and Howitt (1997); Barro (1991); Barro and Sala-i-Martin (1992); Mankiw, Romer, and Weil (1992); Durlauf and Johnson (1995); Jones (1995); Kocherlakota and Yi (1996,1997).
2. This interpretation has been challenged by Bernard and Durlauf (1996).

Economic Growth: Sources, Trends, and Cycles, edited by Norman Loayza and Raimundo Soto, Santiago, Chile.

1. The Empirical Literature

The first stylized fact that appears to be uncontroversial is that whatever the type of dataset used (a cross-section of countries or panel data), the data strongly reject absolute convergence (Barro and Sala-i-Martin, 1995). The simplest test that can be devised to verify this claim using cross-sectional observations takes the form

$$g_i = \varsigma + \vartheta \ln y_{i0} + \varepsilon_i \,, \qquad (1)$$

where y_{it} is per capita GDP in period t for country i, and g_i is the average growth rate of per capita GDP in country i. That is,

$$g_i = \frac{1}{T}\sum_{t=1}^{T} \Delta \ln y_{it} = \frac{1}{T}\left(\ln y_{iT} - \ln y_{i0}\right) \;.$$

When pooled data are used, tests for absolute convergence usually take the form

$$\Delta \ln y_{it} = \varsigma + \vartheta \ln y_{it-1} + \varepsilon_{it} \,. \qquad (2)$$

In both cases, the data are said to favor absolute convergence if the estimate of ϑ is negative and statistically different from zero. If the null hypothesis ($\vartheta = 0$) is rejected, it would support the conclusion not only that poor countries grow faster than rich countries, but also that they all converge to the same level of per capita GDP.

As table 1 and figure 1 show, the convergence hypothesis is strongly rejected by the data.[3] In fact, if these results are taken seriously, the evidence appears to favor divergence instead of convergence. That is, the countries that grew faster were those that had a higher initial per capita GDP.

Because the null hypothesis being tested in both cases is that ϑ is equal to zero versus the alternative that it is negative, equation 2 makes explicit that a test for absolute convergence is essentially a test for a unit root on y. As abundantly documented elsewhere, these tests not only have nonstandard asymptotic properties, but they also lack power.

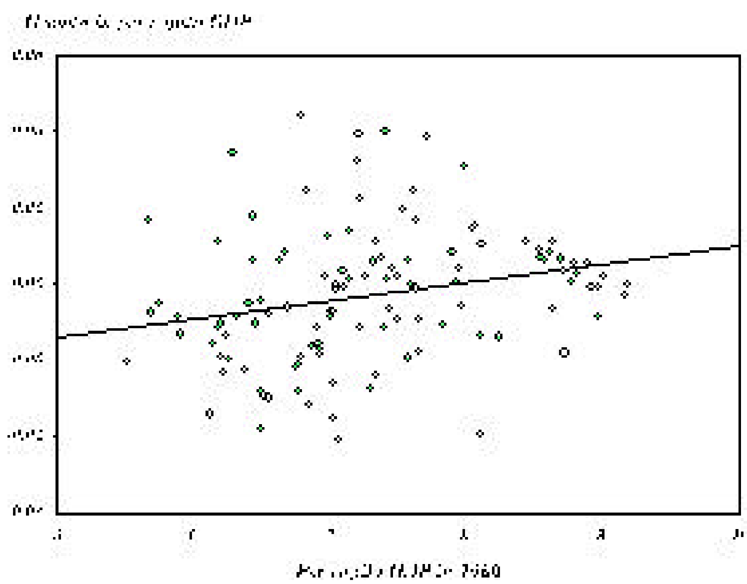
3. All tests using panel data were conducted using the latest version of the Penn World Tables dataset described in Summers and Heston (1991), with data for most variables ranging from 1960 to 1998. Cross-section regressions were conducted using the dataset described in Doppelhofer and others (2000).

Table 1. Tests for Absolute Convergence[a]

Variable or summary statistic	*Cross-section*	*Pooled data*
$\hat{\vartheta}$	0.0047 (0.0014)	0.0048 (0.0010)
Adjusted R^2	0.051	0.007
Number of countries	116	85
Number of observations	116	3,219

a. Standard errors consistent with heteroskedasticity are in parentheses.

Figure 1. Growth Rate from 1960 to 1998 versus 1960 per Capita GDP



In fact, if a traditional (augmented Dickey-Fuller) unit root test on ln*y* were performed for each country, none would reject the null at standard significance levels. Moreover, the first-order autocorrelation coefficient of ln*y* for each country ranges from 0.610 to 0.999, with an average value of 0.947. These results suggest that ln*y* is extremely persistent even in the absence of a unit root, and initial conditions take a long time to dissipate.

Barro (1991) therefore considers a modification of equation 1. Convergence is still understood as the situation in which poor countries grow faster than rich countries (unconditionally), but other factors may influence their growth rates and thereby prevent convergence in per capita GDP levels. Tests for conditional convergence using cross-sectional observations usually take the form

$$g_i = \varsigma + \vartheta \ln y_{i0} + \varphi' \mathbf{x}_i + \varepsilon_i \,, \tag{3}$$

where $\mathbf{x}$ is a vector of k variables that may influence growth. Given that the $\mathbf{x}$ variables are different for each country, incomes might never converge even if ϑ is negative.

Table 2 presents the results of cross-sectional and panel regressions that include some of the usual candidates for specifications such as equation 3.[4] Serious problems plague this strategy, as noted by Durlauf (2001). First, as economic theory is usually silent with respect to the set of $\mathbf{x}$ variables to be included, empirical studies often abuse the resulting flexibility for selecting among the potential candidates. Durlauf and Quah (1999) report that as of 1998, over ninety different variables had appeared in the literature, despite the fact that no more than 120 countries are available for analysis in the standard datasets. Second, important biases in the results may stem from the endogeneity of most of the control variables used (Cho, 1996). Third, the estimated coefficients of the convergence parameter (ϑ) are rather small, suggesting that ln*y* continues to be extremely persistent even after controlling for the $\mathbf{x}$ variables. Fourth, as a corollary of the previous observation, initial conditions may play a crucial role in the results. Fifth, the robustness of results in terms of the potential determinants of long-run growth is subject to debate (see, for example, Levine and Renelt, 1992; Sala-i-Martin, 1997; and Doppelhofer, Miller, and Sala-i-Martin 2000). Finally, several of the variables included in the $\mathbf{x}$ vector are fixed effects that cannot be modified; if these variables were actually long-run determinants of growth, convergence

4. The model that uses cross-sectional observations includes the following $\mathbf{x}$ variables (signs on the coefficients associated with the variables are in parentheses): life expectancy in 1960 (+); equipment investment (+); years of open economy (+); a rule of law index (+); a dummy variable for sub-Saharan African countries (–); and the fraction of people that profess the Muslim (+), Confucian (+), and Protestant (–) religions. The model that uses panel data was estimated using fixed effects and the following $\mathbf{x}$ variables: investment-to-GDP ratio (+); growth rate of the population (–): ratio of exports plus imports to GDP (+); ratio of liquid liabilities to GDP (–): inflation rate (–); and ratio of government consumption to GDP (–).

Table 2. Tests for Conditional Convergence[a]

Variable or summary statistic	*Cross-section*	*Pooled data*
$\hat{\vartheta}$	–0.0154 (0.0028)	–0.0456 (0.0062)
Adjusted R^2	0.811	0.181
Number of countries	79	85
Number of observations	79	2,552

a. Standard errors consistent with heteroskedasticity are in parentheses.

would never be achieved (even with $\vartheta < 0$).[5]

Durlauf and Johnson (1995) suggest that cross-sectional growth behavior may be determined by initial conditions. They explore this hypothesis using a regression tree methodology, which turns out to be a special case of a threshold regression (Hansen, 2000). The basic idea is that the level of per capita GDP on which each country converges depends on some initial condition (such as initial per capita GDP), such that some countries converge on one level and others on another, depending on this characteristic. A common specification used to test this hypothesis considers a modification of equation 1 that takes the form

$$g_i = \begin{cases} \zeta_1 + \vartheta_1 y_{i0} + \varepsilon_i & \text{if } y_{i0} < \aleph \\ \zeta_2 + \vartheta_2 y_{i0} + \varepsilon_i & \text{if } y_{i0} \geq \aleph \end{cases}, \quad (4)$$

where $\aleph$ is a threshold that determines whether country *i* belongs to the first or the second "club." In this case, convergence would not be achieved by the sample as a whole, but it would be by members of each group.

If equation 4 were the actual data-generating process, results such as those in table 1 could be easily motivated: if two regimes are present, with each regime converging to a different state and at a different rate, then estimations based on a single regime might produce a nonsignificant estimate for the convergence parameter. On the other hand, equation 4 states that if the threshold variable (in this case, initial per capita GDP) is correlated with some of the **x** variables included in equation 3, results such as those reported in table 2 are likely to be

5. A curious example of such a variable is absolute latitude, which measures how far a country is from the Equator. When statistically significant, its coefficient is usually positive, implying that one way to enhance growth would be for a country to move its population toward the North or the South Pole.

encountered, even if the **x** variables are not (necessarily) determinants of long-run growth. Equation 4 has an unequivocal implication, however, in terms of the distribution of per capita GDP across countries: if the parameters that characterize each regime are different, a threshold process should be consistent with a bimodal distribution for ln*y*.

Quah (1993, 1997) notices that relative per capita GDP (defined as the ratio of the per capita GDP of country *i* to average world per capita GDP, represented here by $\tilde{Y}_{it}$) displays such bimodality. He conjectures that if convergence clubs were present, even if the unconditional distribution of initial per capita GDP were unimodal, the existence of such clubs would imply that countries would not converge to a degenerate distribution in the long run (as absolute convergence would seem to imply). Rather, one group may converge to one level of per capita GDP and another group to another, in which case twin peaks would arise.

Figure 2 presents kernel estimators of the unconditional density of relative per capita GDP in 1960 and 1995. Consistent with Quah's claim, twin peaks are present in 1995; a bimodal distribution also appears to be present in 1960, however. If Quah is right, rich countries will converge to one distribution, while initially poor countries will never be able to catch up, converging instead to a distribution with a permanently lower per capita GDP. On the other hand, figure 3 presents surface and contour plots of the log of relative per capita GDP, which shows that a bimodal joint density does indeed appear to be consistent with the data. Given that the initial distribution is also bimodal, it is difficult to assess whether the bimodal distribution of 1995 is due to the presence of twin peaks or to the persistence of the per capita GDP level.

2. A Model with Absolute Convergence

This section presents a simple exogenous growth model in which absolute convergence holds and then asks whether the tests for convergence presented in the previous section would be robust. That is, if time-series realizations were generated using a model in which convergence holds, would tests for convergence find convergence? Simply put, the models that I discuss imply that countries should converge to a stationary distribution, that countries with initially lower GDP should grow faster, and that twin peaks should not be present in the long run. To clarify concepts, I now present the type of model to be used, describe its properties and the data-generating process that ln*y* would obey, and ask whether the tests discussed in the previous section are really tests

Figure 2. Densities of Relative per Capita GDP

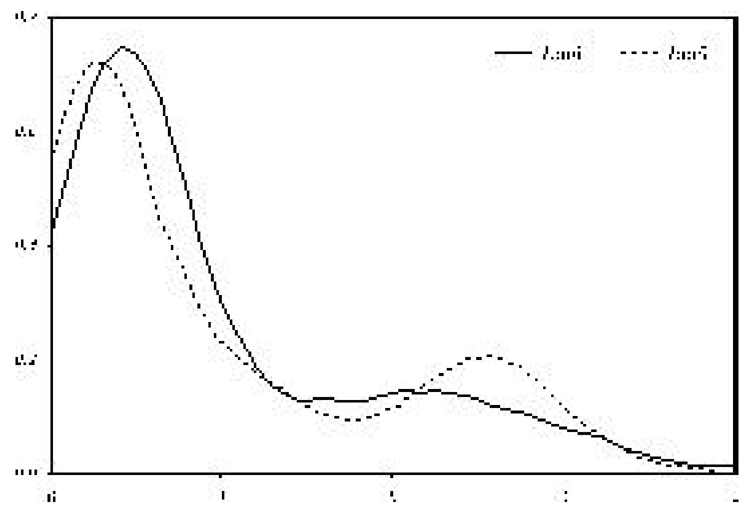
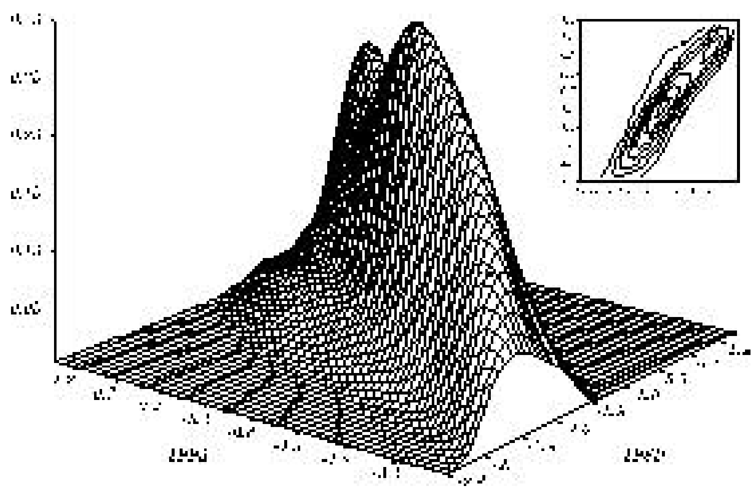


Figure 3. Surface and Contour Plots of Log of Relative per Capita GDP



for convergence. The model considers a representative, infinitely lived household that maximizes

$$u_0 = E_0 \sum_{t=0}^{\infty} \beta^t L_t \ln c_t \ ,$$

where $0 < \beta < 1$ is the subjective discount factor, c_t ($= C_t / L_t$) is per capita consumption, and E_t is the expectations operator conditional on information available for period t.[6] There is no utility from leisure, and the labor force is equal to L_t. Utility is maximized with respect to per capita consumption and the per capita capital stock, k_{t+1}, subject to the budget constraint:

$$K_{t+1} + C_t = e^{z_t} K_t^{\alpha} \left[(1+\lambda)^t L_t \right]^{1-\alpha} + (1-\delta) K_t \ ,$$

where α is the compensation of capital as a share of GDP. In this economy, technological progress is labor augmenting and occurs at the constant rate, λ. Note that production is affected by a stationary productivity shock, z_t. It is straightforward to show that capital and consumption per unit of effective labor, $\hat{k}_t$ and $\hat{c}_t$, are stationary.[7] In fact, one can transform the above economy to a stationary economy and obtain exactly the same solutions for $\hat{k}_t$ and $\hat{c}_t$. Such an economy can be characterized by the following maximization problem:

$$\max_{\{\hat{k}_{t+1}, \hat{c}_t\}} E_0 \sum_{t=0}^{\infty} \left[\beta (1+\lambda)^{1-\gamma} \right]^t L_t \ln \hat{c}_t \ , \quad (5)$$

subject to

$$(1+\eta_{t+1})(1+\lambda) \hat{k}_{t+1} + \hat{c}_t = e^{z_t} \hat{k}_t^{\alpha} + (1-\delta) \hat{k}_t \ , \quad (6)$$

where η_t is the rate of population growth for period t.

Given that this model is used to compare the dynamics of different economies, I include a simple channel to induce correlation between

6. Lower case letters denote per capita values, upper case totals, and a hat above a variable denotes that the value is per unit of effective labor.

7. $\hat{k}_t = k_t / (1+\lambda)^t$ and $\hat{c}_t = c_t / (1+\lambda)^t$.

each economy's income (following den Haan, 1995). Specifically, I obtain correlated incomes by assuming that the law of motion of technology shocks in country *i* can be written as

$$z_{it} = \rho z_{it-1} + \varepsilon_{it}, \qquad \varepsilon_{it} = (1-\phi)v_t + \phi w_{it}, \tag{7}$$

where v_t and w_{it} are independent $N(0, \sigma_i^2)$ random variables (for $i = v, w$). If ϕ is equal to zero, all countries face the same aggregate shock; if f is equal to one, each country faces only an idiosyncratic shock.

Fully characterizing the model requires taking a stance on the rate of population growth. Here I consider the case in which fertility is exogenous and has the following law of motion:

$$\ln(1+\eta_{it}) = \overline{\eta}(1-\tau) + \tau \ln(1+\eta_{it-1}) + n_{it}, \tag{8}$$

where n_{it} is an independent $N(0, \sigma_n^2)$ random variable.

Once values for the preference and technology parameters are chosen, this dynamic programming problem can be solved using numerical methods to generate artificial realizations of the variables of interest. Here, I am interested in generating realizations of per capita GDP for several samples of countries and applying the convergence tests discussed in section 1. As shown below, this model implies convergence (in a specific sense defined below). The goal is to evaluate how likely it is for the tests to conclude otherwise, even though the main feature of this model is that countries converge.

3. Convergence Tests and the Model

To understand whether the tests discussed in section 1 are useful in testing for convergence, I tailor the model to instances in which a closed-form expression for the data-generating process of the log of per capita GDP is available. This simplification imposes a very rigid structure on the theoretical model and makes it harder for its realizations to present the features that are considered signs of rejection of the absolute convergence hypothesis.

If $\delta = 1$, an analytical expression for the capital stock policy function is available and is expressed as

$$\ln \hat{k}_{t+1} = \ln(\alpha\beta) - \ln(1+\lambda) + \ln \hat{y}_t, \tag{9}$$

where $\hat{y}_t = e^{z_t}\hat{k}_t^{\alpha}$ is GDP per unit of effective labor.

Because $\ln \hat{y}_t$ can be expressed as

$$\hat{y}_t = z_t + \alpha \ln \hat{k}_t , \tag{10}$$

we can substitute equations 7 and 9 into equation 10 to obtain a simple expression for $\hat{y}_t$:

$$\ln \hat{y}_{it} = A + (\alpha + \rho) \ln \hat{y}_{it-1} - \alpha\rho \ln \hat{y}_{it-2} + \varepsilon_{it} , \tag{11}$$

where $A = \alpha(1 - \rho) [\ln(\alpha\beta) - \ln(1 + \lambda)]$. Because $\hat{y}_t (1+\lambda)^t = y_{it}$ (from above), one can use equation 11 to obtain a compact representation of the data-generating process of per capita GDP, as follows:

$$\ln y_{it} = B + Dt + (\alpha + \rho) \ln y_{it-1} - \alpha\rho \ln y_{it-2} + \varepsilon_{it} , \tag{12}$$

where B and D are constants.[8]

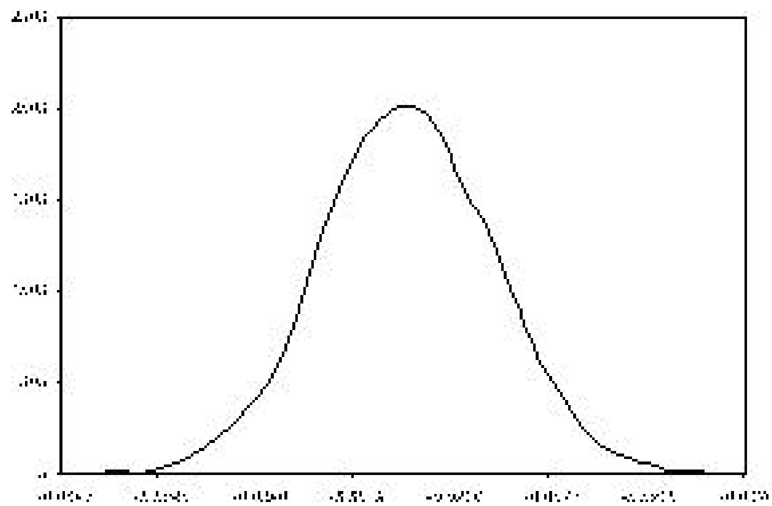
Four features of equation 12 are worth mentioning. First, as is typical of exogenous growth models, per capita GDP is trend stationary.[9] Second, given that the technology shock follows an AR(1) process, ln$y$ follows an AR(2) process.[10] Third, even without exogenous growth ($\lambda = 0$), an AR(1) process for ln$y$ such as equation 2 is consistent with equation 12 only if white-noise technology shocks ($\rho = 0$) are present. Finally, this model suggests that convergence on growth rates and GDP levels should eventually be achieved. The type of convergence on GDP levels depends on the characteristics of the aggregate and idiosyncratic shocks that are present in equation 7. In particular, if the only source of variation in technology shocks is the aggregate shock ($\phi = 0$), all countries should eventually converge on the same per capita GDP, independent of their initial conditions and independent of the persistence of z. On the other hand, if at least part of the variation in technology shocks is due to the idiosyncratic component ($\phi > 0$), then per capita GDP will converge to a nondegenerate distribution that does not display a mass point. In other words, *ln y* will converge to a normal distribution with positive variance, in which

8. More precisely, $B = \alpha(1-\rho)\ln(\alpha\beta) + \rho(1-\alpha)\ln(1+\lambda)$ and $D = (1-\alpha)(1-\rho)\ln(1+\lambda)$.

9. In fact, a case for divergence can only be made when ln$y$ has a unit root, which requires either r = 1 (a unit root in the technology shock) or $\alpha = 1$ (a model of endogenous growth of the AK type).

10. In general, if the productive shocks follow an AR(j) process, *ln y* follows an AR(j + 1) process.

Figure 4. Distribution of $\hat{\vartheta}$ from Absolute Convergence Tests with I.I.D. Shocks[a]



a. Estimates obtained from 2,000 artificial samples for one hundred countries.

case the probability of observing identical levels of *y* is zero.

Next, I focus on the implications of different parameterizations of equation 12 for the convergence tests discussed in section 1.

3.1 Independently and Identically Distributed Shocks

An absolute convergence test such as equation 2 is correctly specified only when the technology shocks are independently and identically distributed (i.i.d.). In that case equation 12 reduces to

$$\ln y_{it} = \alpha \ln(\alpha\beta) + (1-\alpha)\ln(1+\lambda)t + \alpha \ln y_{it-1} + \varepsilon_{it} . \quad (13)$$

Independent of the initial distribution of per capita GDP and population growth rates, $\hat{\vartheta}$ in equation 2 consistently estimates the coefficient $\alpha - 1$, and convergence should occur.[11]

11. That is, $\hat{\vartheta}$ should be negative and statistically different from zero, provided that $0 < \alpha < 1$. Of course, equation 2 should also include a deterministic trend.

Figure 4 presents the empirical distribution of $\hat{\vartheta}$, computed from artificial samples of countries. Each sample consists of a hundred countries, and the initial per capita GDP is obtained from bootstrapping realizations of per capita GDP in 1960. Based on these initial conditions, values of $\ln y_{it}$ are simulated from equation 13 for a thirty-six-year period. Finally, for each sample an estimate for ϑ was obtained by running a regression like equation 1.[12] The probability of obtaining estimates of $\hat{\vartheta}$ consistent with the results from section 1 is, of course, zero, because even if the distribution of per capita GDP in 1960 is considered as the initial condition, i.i.d. shocks with realistic figures for α are unable to produce enough persistence in $\ln y$.

Furthermore, the precise nature of absolute convergence is dictated by ϕ. If $\phi = 0$, then in the long run countries converge (in probability) to the same per capita GDP, whereas if some shocks are idiosyncratic, per capita GDP converges to a nondegenerate distribution in the long run.

Figures 5 and 6 reveal another characteristic of i.i.d. productivity shocks: even when they begin with a bimodal distribution for initial per capita GDP, the bimodality quickly disappears because *y* is not persistent enough. In fact, after thirty-six years, per capita GDP would not feature twin peaks. Thus, i.i.d. shocks are inconsistent with the evidence reported on section 1.

3.2 Persistent Shocks

If the unrealistic setup of i.i.d. technology shocks is abandoned, $\ln y$ can be made significantly persistent by choosing a value of r close to one. Persistence of technology shocks is routinely invoked in the literature on real business cycles, and it is broadly consistent with key stylized facts of modern economies. Once persistence in $\ln y$ is obtained without having to resort to unrealistic values of α, the conclusions regarding i.i.d. shocks change radically.

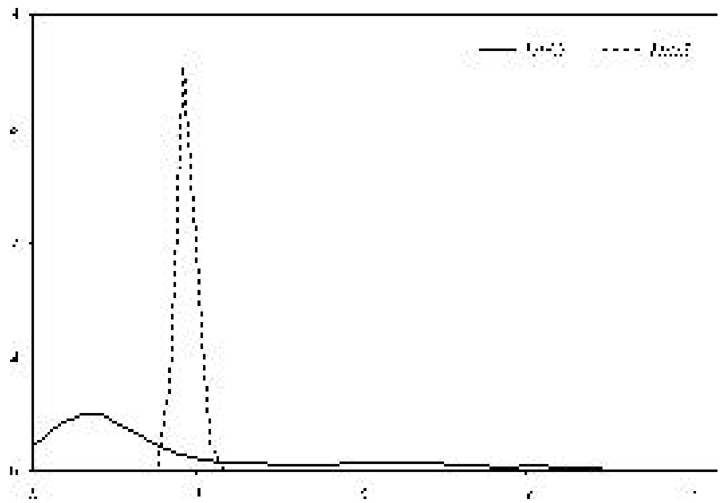
Remember that the law of motion of the univariate representation for $\ln y_{it}$ is expressed by equation 12; that is,

$$\ln y_{it} = B + Dt + (\alpha + \rho)\ln y_{it-1} - \alpha\rho \ln y_{it-2} + \varepsilon_{it} .$$

Convergence tests such as equation 2 are clearly misspecified. Furthermore, as demonstrated by den Haan (1995), the estimated value of

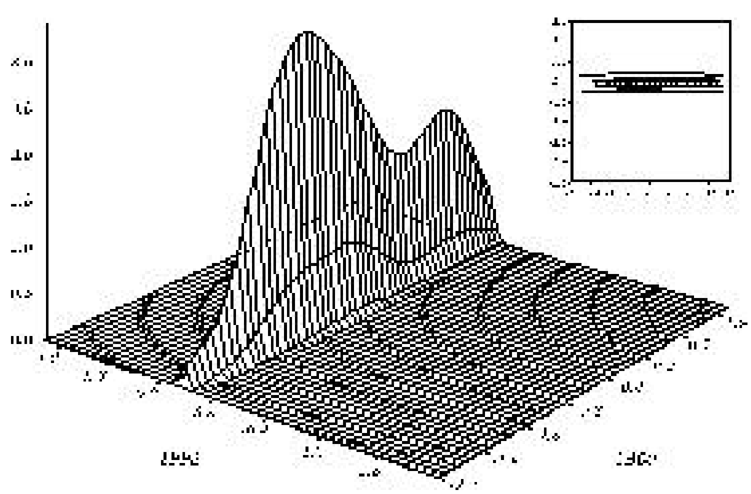
12. The parameter values for this model were set as follows: $\alpha = 0.35$, $\beta = 0.96$, $\lambda = 0$, $\phi = 1$, and $\sigma_w^2 = 0.05^2$.

Figure 5. Densities of Relative per Capita GDP with I.I.D. Shocks[a]



a. Empirical densities for an artificial realization of one hundred countries.

Figure 6. Surface and Contour Plots of Log of Relative per Capita GDP for I.I.D. Shocks[a]



a. Results for an artificial realization of one hundred countries.

ϑ in equation 1 is inconsistent and biased toward zero. In other words, even if the model implied convergence, the estimated value of ϑ is biased toward the rejection of this hypothesis. Furthermore, using pooled observations in equation 2 shows that

$$\hat{\vartheta} \xrightarrow{p} \psi - 1 = \frac{(1-\alpha)(1-\rho)}{1+\alpha\rho},$$

where $\psi = (\alpha + \rho) / (1 + \alpha\rho)$ is the first-order autocorrelation of ln$y$. This implies that the more persistent the technology shocks, the closer the probability limit of $\hat{\vartheta}$ will be to zero.

Figure 7 presents an exercise similar to that reported in figure 4 for the i.i.d. case. Here I consider exactly the same parameterization, but I now set $\rho = 0.97$. The difference is that even when the model implies convergence, the results of estimating equation 1 by bootstrapping the initial distribution of ln$y$ that was observed in 1960 presents a nonnegligible probability (11 percent) that the estimated coefficient would indeed be positive (implying divergence).

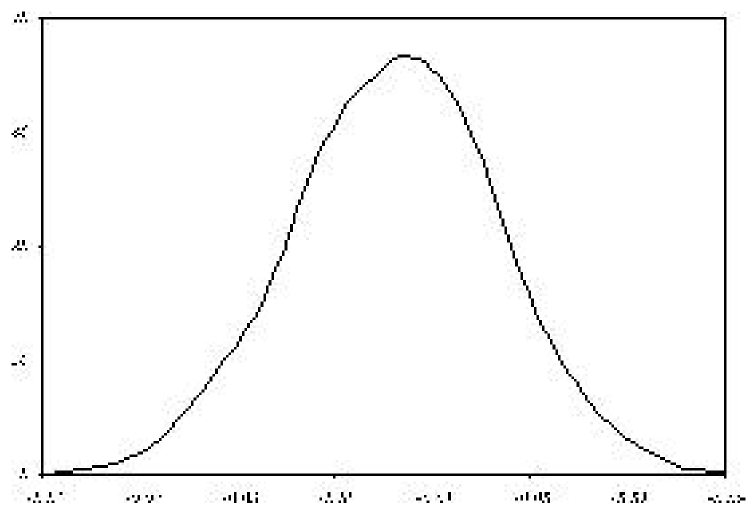
Furthermore, as figure 8 reveals, persistent technology shocks can replicate a bimodal joint distribution of the initial log of per capita GDP (consistent with that observed in 1960) and the figures that would be obtained thirty-five years later. Because initial conditions do not dissipate as fast as in the i.i.d. case, an initially bimodal distribution would persist even over long periods. Thus bimodality in the short run is not inconsistent with a model that displays convergence in the long run.

In summary, persistent technology shocks can be broadly consistent with the evidence reported in section 1, in the sense that whatever the initial conditions of the distribution of per capita GDP, they fade slowly. In particular, this simple model, which displays convergence to a unimodal distribution in the long run, is consistent with twin peaks in the distribution of per capita GDP, even over relatively prolonged horizons.

3.3 The Model and Conditional Convergence

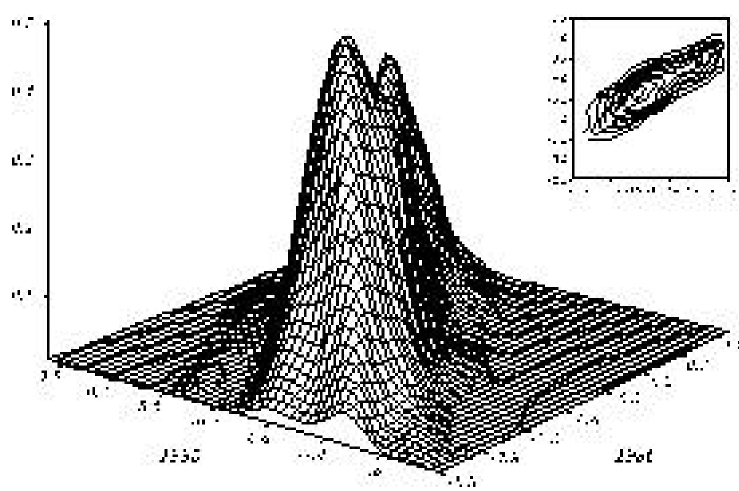
Once persistent shocks are allowed, even the simplest of the exogenous growth models can display several of the features that are considered evidence of divergence or club convergence. Given an initially bimodal distribution of the log of per capita GDP, therefore, persistence by itself could generate an illusion of bimodality for prolonged periods.

Figure 7. Empirical Density of Absolute Convergence Tests with AR(1) Shocks[a]



a. Empirical distribution of the $\hat{\vartheta}$ coefficients obtained with 2,000 artificial samples for one hundred countries.

Figure 8. Surface and Contour Plots of Log of Relative per Capita GDP for AR(1) Shocks[a]



a. Results for an artificial realization of one hundred countries.

However, the models just described are not consistent with evidence of conditional convergence. A few lags added to an equation like equation 2 would become sufficient statistics for ln*y*, and no other variable in the econometrician's information set should be informative. The results of conditional convergence (statistically significant **x** variables) can still be found when a misspecified law of motion for ln*y* is considered. In particular, if some **x** variables are correlated with the initial distribution of *y*, models that do not include as many lags of the variable as necessary can easily be found to be significant.

Furthermore, the models discussed above are among the simplest that can be generated from the theoretical model. In models in which population growth rates can be determinants of ln*y* (such as those described in Chumacero, 2002), the exclusion of ln*y* from growth regressions could generate results consistent with conditional convergence, provided that technology shocks and population growth are persistent and that the chosen **x** variables correlate with initial conditions. In fact, as stressed in section 1, most of the supposedly robust **x** variables that are included in growth regressions are both persistent and strongly correlated with initial conditions.

If the economy is better characterized using parameters that do not allow for an analytical solution for the law of motion of ln*y*, equations 1 and 2 can, at best, be viewed as linear approximations. The more nonlinear the model, the more inaccurate this approximation will be, and any omitted nonlinear terms may be approximated by any **x** variable that is correlated with the initial conditions.

4. Concluding Remarks

This paper takes issue with interpretations of cross-country growth models that contend that the convergence hypothesis is strongly rejected by the data. It shows that even the simplest exogenous growth model that in the long run displays absolute convergence can present several features that such studies argue to be evidence against convergence. In particular, if persistent and moderately volatile productivity shocks are allowed, exogenous growth models can display features such as bimodality and asymmetries in the unconditional distribution of relative per capita GDP. Furthermore, there is a nonnegligible probability that misspecified econometric models will reject absolute convergence even when it is present.

Persistence of technology shocks is not enough to generate these results, however. In this case persistence implies that initial conditions will eventually dissipate, and if bimodality is present in a given period, it will not dissipate for long periods.

The paper also presented simple (and realistic) variations of the models, which ultimately imply convergence, and showed how they can be made consistent with conditional convergence results. This occurs when the chosen determinants of growth are correlated with initial conditions and when the models being tested are misspecified (with an incorrect law of motion of per capita GDP or omission of nonlinearities).

Finally, the paper does not explain the initial bimodality that seems to be present in the data. Apparently relevant policy variables in conditional convergence regressions may have something to do with this. McGrattan and Schmitz (1999) argue that distortionary policies are at fault. If so, the model presented here implies that convergence to an ergodic distribution of per capita GDP should be achieved if these policies also converge.

References

Aghion, P., and P. Howitt. 1997. *Endogenous Growth Theory*. MIT Press.

Barro, R. 1991. "Economic Growth in a Cross-Section of Countries." *Quarterly Journal of Economics* 106(2): 407–43.

Barro, R. J., and X. Sala-i-Martin. 1992. "Convergence." *Journal of Political Economy* 100 (April): 223–51.

———. 1995. *Economic Growth*. New York: W. W. Norton.

Baumol, W. J. 1986. "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show." *American Economic Review* 76(5): 1072–85.

Bernard, A., and S. Durlauf. 1996. "Interpreting Tests of the Convergence Hypothesis." *Journal of Econometrics* 71(1–2): 161–73.

Cho, D. 1996. "An Alternative Interpretation of Conditional Convergence Results." *Journal of Money, Credit and Banking* 28: 161–73.

Chumacero, R. 2002. "On the Power of Absolute Convergence Tests." Santiago: Banco Central de Chile and Universidad de Chile. Unpublished paper.

Den Haan, W. 1995. "Convergence in Stochastic Growth Models: The Importance of Understanding Why Income Levels Differ." *Journal of Monetary Economics* 35(1): 65–82.

Doppelhofer, G., R. Miller, and X. Sala-i-Martin. 2000. "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach." Working paper 7750. Cambridge, Mass.: National Bureau of Economic Research.

Durlauf, S. 2001. "Manifesto for a Growth Economics." *Journal of Econometrics* 100(1): 65–9.

Durlauf, S., and P. Johnson. 1995. "Multiple Regimes and Cross-Country Growth Behavior." *Journal of Applied Econometrics* 10(4): 363–84.

Durlauf, S., and D. Quah. 1999. "The New Empirics of Economic Growth." In *Handbook of Macroeconomics*, edited by J. Taylor and M. Woodford. Amsterdam: North-Holland.

Hansen, B. 2000. "Sample Splitting and Threshold Estimation." *Econometrica* 68(3): 575–603.

Jones, C. 1995. "Time Series Tests of Endogenous Growth Models." *Quarterly Journal of Economics* 110(2): 495–525.

Kocherlakota, N., and K. Yi. 1996. "A Simple Time Series Test of Endogenous versus Exogenous Growth Models." *Review of Economics and Statistics* 78(1): 126–34.

———. 1997. “Is There Endogenous Growth in the Long Run? Evidence from the United States and the United Kingdom.” *Journal of Money, Credit and Banking* 29(2): 235–62.

Levine, R., and D. Renelt. 1992. “A Sensitivity Analysis of Cross-Country Growth Regressions.” *American Economic Review* 82(4): 942–63.

Mankiw, G., D. Romer, and D. Weil. 1992. “A Contribution to the Empirics of Economic Growth.” *Quarterly Journal of Economics* 107 (May, *Papers and Proceedings, 1991*): 407–37.

McGrattan, E., and J. Schmitz. 1999. “Explaining Cross-Country Income Differences.” In *Handbook of Macroeconomics,* edited by J. Taylor and M. Woodford. Amsterdam: North-Holland.

Quah, D. 1993. “Empirical Cross-Section Dynamics in Economic Growth.” *European Economic Review* 37: 426–34.

———. 1997. “Empirics for Growth and Distribution: Polarization, Stratification, and Convergence Clubs.” *Journal of Economic Growth* 2(1): 27–59.

Sala-i-Martin, X. 1997. “I Just Ran Four Million Regressions.” Working paper 6252. Cambridge, Mass.: National Bureau of Economic Research.

Solow, R. (1956). “A Contribution to the Theory of Economic Growth.” *Quarterly Journal of Economics* 70: 65–94.

Summers, R., and A. Heston. 1991. “The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1998.” *Quarterly Journal of Economics* 106(2): 327–68.