Convergence and Modernization*

Robert J. Barro, Harvard University

August 2014

Abstract

In a panel of countries since 1960, the estimated convergence rate for per capita GDP is around 1.7% per year, conditional on an array of explanatory variables that hold constant countries' long-run characteristics. The introduction of country fixed effects generates a much higher—and, I argue, misleading—convergence rate. In a longer time frame—28 countries with GDP data starting between 1870 and 1896—estimation with country fixed effects is more appropriate, and the estimated convergence rate is around 2.6% per year. Combining the point estimates from the post-1960 and post-1870 panels suggests that the conditional convergence rate is between 1.7% and 2.6% per year, an interval that contains the "iron-law" rate of 2%. In the post-1960 panel, estimation without country fixed effects supports the modernization hypothesis, in the form of positive effects of per capita GDP and schooling on maintenance of law and order and democracy. In the post-1870 panel, estimation without or with country fixed effects supports the modernization idea, in the sense of positive effects of per capita GDP and schooling on the Polity indicator for democracy.

***An earlier version of this paper was presented at the Nobel Symposium on Growth and Development, Stockholm, September 3-5, 2012. I appreciate excellent research assistance from David Hsu and Ricardo Perez-Truglia and comments from Philippe Aghion, Abhijit Banerjee, Xavier Gabaix, Elhanan Helpman, Chad Jones, Peter Klenow, Jong-Wha Lee, Andrei Shleifer, Ugo Troiano, Jose Ursúa, and Romain Wacziarg.

According to the "iron law of convergence," countries eliminate gaps in levels of real per capita GDP at a rate around 2% per year.[1] Convergence at a 2% rate implies that it takes 35 years for half of an initial gap to vanish and 115 years for 90% to disappear. Convergence-rate parameters are important to pin down because they provide guidance on how fast countries like China and India are likely to catch up to richer countries. The convergence rate may also reveal how fast a poor African country could develop or how rapidly North Korea could catch up to South Korea, and so on.

Empirically, the iron law takes the form of unconditional or absolute convergence in some samples of economies; those that are reasonably homogeneous in terms of long-run or steady-state characteristics. For example, a roughly 2% convergence rate emerged for per capita personal income in a long-term panel of U.S. states in Barro and Sala-i-Martin (1992).[2] This convergence was absolute in the sense of not having to be conditioned on a set of variables that capture differences in long-run positions. The results implied—in accordance with the data—that it would take the U.S. South about a century after the Civil War to get close in per capita income to the rest of the country. Applying these results to East versus West Germany suggested that a short time frame for convergence was not a realistic expectation.[3] And, looking forward to

[1]I first heard this term applied to my empirical findings on economic growth by Rudi Dornbusch. However, Larry Summers said that Rudi got the term from him. In any event, the term is reminiscent of the "iron law of wages." According to Wikipedia, this phrase came from Lassalle, but Marx and Engels argued that Lassalle got the idea from Malthus's theory of population and the specific terminology from Goethe.

[2]Baumol (1986, Figure 2) reported unconditional convergence from 1870 to 1979 for 16 countries (all subsequently OECD members), using data from Maddison (1982). However, De Long (1988) showed that Baumol's results depended on a sample-selection issue, whereby only countries that were rich toward the end of the sample (1979) were considered. Unconditional convergence did not hold for an expanded sample of 22 countries that were selected based on per capita income in 1870 (De Long [1988, Figure 2]). This sample-selection criticism of Baumol's (1986) findings was presented earlier by Romer (1986, pp. 1012-1013). Rodrik (2013) finds unconditional convergence in labor productivity across manufacturing industries for recent decades in 118 countries.

[3]Barro (2002) found that the predicted slow convergence between East and West Germany accorded with regional data on GDP per worker through the late 1990s. However, wage rates converged faster because of the German government's transfer and subsidy policies.

the potential reunification of North and South Korea, the iron law presents a pessimistic outlook on how rapidly the large gap in per capita product could be eliminated.

In a collection of heterogeneous economies that differ substantially in terms of long-run properties, the 2% convergence rate holds in a conditional sense. That is, convergence applies only with an allowance for differences in constant or slowly varying cross-economy characteristics, such as saving rates or fertility rates or quantity of human capital or institutional quality or colonial history or geographical features. For example, a convergence rate around 2% per year appeared in a cross section of 98 countries in Barro and Sala-i-Martin (1992, Table 3), after conditioning on an array of variables that differed by country.[4] Because of the conditioning variables, these results were more pessimistic than the iron-law convergence rate would suggest. Poor places—for example, many sub-Saharan African countries or North Korea or Burma or Bolivia or Venezuela—might not converge at all if key underlying variables, such as the quality of human capital and institutions, were not improved.

The present study assesses convergence behavior within two empirical contexts. One data set comprises the large number of countries with observations for many variables since 1960. Another data set exploits recent advances in long-term national-accounts information. These data go back to 1870 but apply to fewer countries and variables. In both contexts, the distinction between absolute and conditional convergence is important. And, within the context of conditional convergence, a key technical issue is whether the cross-country growth regressions include country fixed effects.

[4] In earlier work, Barro (1991) reported conditional convergence for the cross section of 98 countries but did not express the results in terms of a convergence rate.

Many analyses of economic growth stress effects from the quality of institutions, gauged particularly by maintenance of the rule of law and democracy.[5] A prominent feature of this analysis is two-way causation between economic development and institutional quality. Specifically, according to the "modernization hypothesis," economic development spurs the introduction and maintenance of higher quality institutions, including well-functioning representative democracy.[6] The validity of the modernization thesis is important for its own sake—particularly for understanding how democracy and rule of law evolve—as well as for assessing institutional determinants of economic growth.

I use the post-1960 and post-1870 data sets to assess the modernization hypothesis. From an econometric standpoint, the analysis of modernization parallels the study of convergence in per capita GDP. Both analyses involve convergence rates and an array of explanatory variables that determine long-run positions. And empirical inferences in both contexts are sensitive to the treatment of country fixed effects.

I. Country Fixed Effects

Cross-country empirical findings on convergence and modernization depend on whether the panel regressions include country fixed effects. Although the incorporation of these fixed effects into cross-country panel regressions has become almost routine,[7] the merits are not straightforward, because they involve a tradeoff between two forces, highlighted by Nerlove

[5] Knack and Keefer (1995) and Mauro (1995) studied growth effects from rule of law and corruption. Przeworski and Limongi (1993) and Barro (1997, Ch. 2) assessed growth effects from democracy. King and Levine (1993) examined effects of financial institutions on economic growth. Glaeser, La Porta, Lopez-de-Silanes, and Shleifer (2004) argued that institutions should be measured by basic legal constraints on the government, rather than political outcomes, which include official corruption and risk of expropriation.

[6] Contributions to the modernization literature include Aristotle's *Politics*, Lipset (1959), Dahl (1991), and Huntington (1991). Marx (1904) extended the modernization idea to a predicted collapse of organized religion under capitalism.

[7] This approach applied to economic growth seems to have begun with Knight, Loayza, and Villanueva (1993); Islam (1995); and Caselli, Esquivel, and Lefort (1996). Acemoglu, Johnson, Robinson, and Yared (2005, 2008) advocate the use of country fixed effects in studies of the modernization hypothesis.

(2000). The appendix brings out details, using Monte Carlo methods. This analysis expands on Kiviet (1995, Tables 1-3), Judson and Owen (1999, Tables 1-2), and Hauk and Wacziarg (2009, Table 4) to allow for sample lengths (such as 50 or 140 years) and convergence rates (around 0.02 per year) that match up with data on economic growth.

To fix ideas, consider cross-country panel regressions for the growth rate of per capita GDP. Country fixed effects are attractive as a way to allow for unobserved, persistent country characteristics that influence long-run per capita GDP and are also correlated with observed per capita GDP. That is, rich countries tend to have prospered because they possess persistently favorable characteristics, such as high quality institutions, that lead to high steady-state per capita GDP. From this omitted-variables perspective, the exclusion of country fixed effects tends to bias upward the estimated effect of lagged GDP on current GDP and, thereby, bias downward the estimated convergence rate. One familiar example of this effect is the tendency to estimate an absolute convergence rate near zero in a panel of heterogeneous countries. However, the bias may be small if the framework without country fixed effects includes a rich set of explanatory variables so that little remains of omitted variables that are conditionally correlated with per capita GDP.

The second force involves the Hurwicz (1950)-type bias in the estimated coefficient of a lagged dependent variable. In panels that are small in the time dimension, Nickell (1981), Arellano and Bond (1991), Kiviet (1995), and Nerlove (2000), among others, show that this force biases down the fixed-effects estimator (based on least squares with dummy variables) for the coefficient of the lagged dependent variable.[8] On this ground, the estimated convergence rate

[8]Let y_{it} be the dependent variable and ε_{it} the associated error term for country i at time t. (Hurwicz [1950] dealt with only one country.) The ε_{it} are serially uncorrelated, i.i.d. random variables. The Hurwicz (1950)-type bias arises because the realized error terms appear in the sample mean of $y_{i,t-1}$. That is, a higher ε_{it} implies a higher sample mean of $y_{i,t-1}$ and, therefore, a lower $y_{it,t-1}$ when expressed relative to its sample mean. The implied negative

tends to be overestimated (because the persistence in the level of the dependent variable is underestimated).

Nickell (1981, p. 1422) provides a formula for the Hurwicz-type bias in the least-squares-with-dummy-variables estimate for the coefficient of a lagged dependent variable. The Nickell specification includes cross-sectional fixed effects, no other explanatory variables (X variables), and—unlike Hurwicz—a large (effectively infinite) number of cross sections. The Nickell analysis also treats the initial value of the dependent variable as given; specifically, independent of the fixed effects.

To get a simplified version of the Nickell formula for the Hurwicz (1950)-type bias, let β denote the magnitude of the convergence rate per year. The coefficient of the lagged dependent variable, labeled ρ in Nickell's analysis, is then $\rho = e^{-\beta\tau}$, where τ is the period length in years. Let T be the sample length in years (differing from Nickell's notation). Nickell's formula for the proportionate bias in the estimated β can then be expressed, if $\beta > 0$, as

$$(1) \qquad [(\widehat{\beta} - \beta)/\beta] \approx \frac{2\cdot(e^{-\beta T} - 1 + \beta T)}{\beta^2 T^2 - 2\cdot(e^{-\beta T} - 1 + \beta T)} > 0,$$

where the approximation uses $\beta\tau << 1$ (so that, in effect, the convergence that occurs within an observation period is small).

An important implication of the Nickell formula in equation (1) is that the proportionate bias depends only on the product βT. Therefore, a change in the period length, τ, with T held fixed, does not affect the bias. (The Monte Carlo analyses in the appendix and the empirical findings discussed later accord with this result.) In particular, the bias depends on the overall length of the sample, T, not the number of periods, T/τ, over which the data are observed (given

covariance between ε_{it} and the deviation of $y_{i,t-1}$ from its sample mean biases downward the OLS estimate of the coefficient of $y_{i,t-1}$.

the condition $\beta\tau<<1$). This finding means that the bias cannot be reduced by raising the frequency of observation; for example, from 10 years to 5 years or 1 year. It also does not help to have a large number of cross sections (since this number is already taken to be infinity in Nickell's analysis).

Quantitatively, if β=0.02 per year, Nickell's formula in equation (1) generates an upward bias of 0.056 when T=50 years and 0.018 when T=140 years (corresponding to the long-term sample used later). Hence, although the bias approaches zero as T approaches infinity, the bias tends to be large in samples of realistic length. However, these results are only suggestive, because Nickell's formula depends on a number of unrealistic assumptions, particularly about the exclusion of X variables and the treatment of the initial value of the dependent variable as given, independent of the fixed effects. Monte Carlo results related to these issues are in the appendix.

When country fixed effects are included and the time dimension of the sample is moderate—say 20-50 years—the panel effectively comprises multiple cases of moderate length. Since the regression for each country features a Hurwicz (1950)-type bias, this bias applies also to the overall panel. Therefore, although the Nickell (1981) extension to include a large number of cross sections matters quantitatively, that extension does not alter the essential source of the bias. In any event, I refer subsequently to the Hurwicz-Nickell bias when considering the estimated coefficient of a lagged dependent variable.

If country fixed effects are excluded, the observations are effectively stacked in the time and country dimensions. That is, with 151 countries and 10 time periods (as in the subsequent analysis of data since 1960), the roughly 1500 observations amount to a large sample in the relevant time dimension, and the Hurwicz-Nickell bias is negligible. In other words, the bias arises because of the inclusion of country fixed effects. These conjectures are confirmed by the Monte Carlo studies described in the appendix.

Inclusion of country fixed effects also affects the estimated coefficients and, especially, standard errors of explanatory variables—X variables—other than lagged dependent variables. Coefficients on country variables that are constant (such as geographical features and colonial history) cannot be estimated at all, and variables that have little within-country time variation cannot be estimated with precision. In effect, the inclusion of country fixed effects throws out much of the information in isolating the effects of X variables on growth rates or other variables.

Problems in estimating coefficients of X variables in a fixed-effects context apply to the recent debate about the modernization hypothesis in Glaeser, La Porta, Lopez-de-Silanes, and Shleifer (2004); Glaeser, Ponzetto, and Shleifer (2007); and Acemoglu, Johnson, Robinson, and Yared (2005, 2008). The failure in the Acemoglu, et al. studies to find statistically significant effects on democracy from per capita GDP and education depends, in their main analysis, on the inclusion of country fixed effects. This result is not surprising because, with country fixed effects, it is challenging to estimate statistically significant coefficients on X variables that do not have a lot of independent variation over time within countries. In contrast, without country fixed effects, as in the Glaeser, et al. studies and Barro (1999), the typically substantial cross-sectional variations in the X variables make it easier to isolate statistically significant effects.

The perspective changes in the context of panel data observed for over a century. In this setting, the econometric problems posed by the inclusion of country fixed effects are less serious. In particular, the Hurwicz-Nickell bias is smaller in this context, as confirmed by equation (1) and the Monte Carlo studies in the appendix.

II. The Framework of Conditional Convergence

This section summarizes well-known implications of the neoclassical growth model and its extensions for empirical analyses of conditional convergence. The model features a production function,

$$Y = A \cdot F(K, L), \tag{2}$$

where $F(\cdot)$ satisfies the usual neoclassical properties, including constant returns to scale in capital, K, and labor, L. Output per worker, $y \equiv Y/L$, depends on capital per worker, $k \equiv K/L$:

$$y = f(k). \tag{3}$$

The economy is closed, so that saving and investment coincide. In the baseline model, there is no government sector, but extensions allow for government purchases and taxes.

The labor force, measured as worker-hours per year, is related in a fixed way to population, which grows at rate n. The labor force is fully employed and, therefore, corresponds to labor input, L, in equation (2). Given these assumptions, the quantities y and k, measured per worker in equation (3), can be interpreted as quantities per capita. In the baseline model, n is constant over time within an economy but may differ across economies.

In the baseline model, the gross saving rate (which equals the ratio of gross investment to GDP) is a constant, s (as in Solow [1956] and Swan [1956]), which may differ across economies. The first major extension of the Solow-Swan framework was to endogenize the saving rate. In a setting based on Ramsey (1928), Cass (1965), and Koopmans (1965), the representative household determines the optimal saving rate at each point in time in the context of an iso-elastic, time-additive utility function. The equilibrium value of s then varies over time within economies along a transition path to the steady state. Barro and Sala-i-Martin (2004, Ch. 2) show that, if F(·) is Cobb-Douglas, the transitional behavior of the saving rate is monotonic—always rising toward its steady-state value, always falling toward its steady-state value, or always constant at its steady-state value. With rising (falling) s, the convergence rate is slower (faster) than in the model with fixed s. With alternative utility functions, the transition may not be monotonic; for example, the transition path for s may be hump-shaped. Across economies, differences in preference and other parameters can shift the path of s when considered in relation to GDP per capita, y. An economy with higher s at a given y tends to grow faster.

Just as the Ramsey-Cass-Koopmans analysis endogenized the fixed and exogenous saving rate of Solow-Swan, the model can be extended along the lines of Malthus (1798) to endogenize the population growth rate, n.[9] Barro and Sala-i-Martin (2004, section 9.2) allow for choices over time of the fertility rate for a given mortality rate. In the case where child-rearing costs rise linearly with k (because, as in Becker [1991], child-rearing is intensive in parental time, which is valued in accordance with the wage rate), n falls montonically during the transition to the steady state. This pattern reduces the convergence rate compared to the setting in which n is fixed. However, the allowance for goods costs of child rearing can generate a non-monotonic pattern in which fertility first rises and later falls as per capita GDP, y, increases. Across economies, differences in preference and other parameters (including costs related to the rearing of children) can shift the path of n when considered in relation to y. An economy with lower n at a given y tends to grow faster.

In the baseline model, the productivity factor, A, can differ across economies. This factor may rise over time due to exogenous technical progress. The usual steady-state properties

[9]Malthus had a model of endogenous population growth, though he sometimes got the signs wrong. Specifically, he predicted a positive effect of real per capita GDP on fertility, whereas, in the modern data, the relation goes strongly in the opposite direction. See Manuelli and Seshadri (2009), who emphasize the distinction between income effects (stressed by Malthus) and substitution effects associated with increases in wage rates.

depend on technical progress taking the labor-augmenting form. In this setting, effective labor input, $\hat{L}$, replaces L in equation (1), $\hat{L}$ is the multiple e^{xt} of L, and $x \geq 0$ is the rate of technological progress. In this context, the results for the growth-rate transition go through in terms of quantities per effective labor; that is, as $\hat{y} \equiv Y/\hat{L}$ and $\hat{k} \equiv K/\hat{L}$. In the Cobb-Douglas case, labor-augmenting technical progress is equivalent to exogenous growth in A in equation (1). In extensions, differences in A (or in the path of A) within or across economies can reflect the quality of institutions, including maintenance of property rights and the efficiency of the tax system. That is, the economy's response to weak property rights or high marginal income-tax rates is essentially equivalent to its response to a reduction in productivity, A.

The research on endogenous growth due to endogenous technical change began with Romer (1987, 1990), Aghion and Howitt (1992), and Grossman and Helpman (1991, Chs. 3 and 4).[10] In these models, technological advances derive from purposeful and successful R&D activity. In "varieties" models (Romer [1987, 1990] and Grossman and Helpman [1991, Ch. 3]), discoveries of new types of intermediate inputs raise productivity by (metaphorically) expanding the number of inputs employed in production. In "quality-ladders" models (Aghion and Howitt [1992] and Grossman and Helpman [1991, Ch. 4]), innovations raise the quality of intermediate inputs within sectors (or, equivalently, improve the efficiency of production processes). From the perspective of the Solow-Swan model, the theories of technological progress effectively endogenize the growth rate of the productivity factor, A. Endogenous growth theory can be

[10]Other research on endogenous growth—exemplified by Romer (1986), Lucas (1988), and Rebelo (1991)—built on earlier work by Arrow (1962) and Uzawa (1965) and did not provide a theory of technical change. In these models, growth may go on indefinitely because the returns to investment in a broad class of capital goods, which include human capital, may not diminish as economies develop. (This idea goes back to Knight [1944]). Spillovers of knowledge across producers and external benefits from human capital may be parts of the process because they help to avoid the tendency for diminishing returns to capital. Models of this type are sometimes described as "AK models," because the growth dynamics looks like that in a simple framework with constant returns to capital.

viewed, accordingly, as another extension of the Solow-Swan model to endogenize a key parameter; in this case, A (that is, the time path of A), rather than s or n.

Endogenous growth theory has not played a major role in empirical studies of the determinants of economic growth across countries. The main empirical applications of the theory have involved effects of R&D outlays. This research builds on the (pre-endogenous-growth) approach described by Griliches (1973), which starts by computing TFP growth rates as Solow residuals in a growth-accounting framework.[11] The TFP growth rates can then be related econometrically to measures of R&D expenditures. This methodology was applied to U.S. firms and industries by Griliches and Lichtenberg (1984) and Griliches (1988). Coe and Helpman (1995) applied this framework to aggregate data for OECD countries and reported large positive effects of R&D outlays on economic growth.[12] However, a problem with this approach is that—in the absence of good instruments—a positive relation between R&D spending and economic growth can reflect reverse causation from growth opportunities to R&D, rather than effects of R&D and technological progress on growth.

One reason that endogenous growth theory has not played a major role in cross-country growth analysis is that the theory may apply mostly to the worldwide average growth rate. This perspective would apply if international trade, foreign investment, and the flow of ideas lead to rapid diffusion of technology across countries. The speed of diffusion of technology was assessed theoretically in Nelson and Phelps (1966), who stressed the role of human capital, and in subsequent models summarized in Barro and Sala-i-Martin (2004, Ch. 8). Caselli and Coleman (2001) found empirically for developing countries that imports of computers and other

[11]This idea began with Solow (1957) and was extended to allow for changing quality of an array of inputs by Jorgenson and Griliches (1967). See also Barro and Sala-i-Martin (2004, section 10.4).
[12]Aghion, Bloom, Blundell, Griffith, and Howitt (2005) have used a model related to endogenous-growth theory as the framework for an empirical study of the relation between competition and innovation across firms.

high-tech equipment—viewed as a proxy for technology absorption—were spurred by increased imports from technologically advanced countries. Technological diffusion was also higher when the home country had higher levels of school attainment at secondary and higher levels and better institutional quality.

A further extension of the neoclassical growth model, introduced by Mankiw, Romer, and Weil (1992), distinguishes human from physical capital. One way that this extension affects the dynamics of economic growth involves the greater difficulty in adjusting human capital, H, compared to physical capital, K. In this case, the growth rate of y tends to be higher at a given y if the ratio H/K is higher. For an economy below its steady-state position, a higher H/K (perhaps generated from a war that destroyed much more physical than human capital) means that subsequent growth focuses on K, which is easier than H to expand rapidly.

Suppose that L is the number of workers and h is human capital per worker, so that quality-adjusted labor input is H=hL. The input H then replaces L in equation (1). Assume, further, that h relates to years of schooling along Mincerian lines, so that

$$h = e^{\lambda S},$$

where S is (average) years of schooling and λ is the rate of return on schooling (if the cost of schooling is the income foregone by not employing human capital in production). If the production function in equation (1) is Cobb-Douglas with exponents a on K and b on H, we can derive:

$$\text{(3)} \qquad \log(H/K) = (1/a)\cdot\log(A) - (1/a)\cdot\log(y) + \lambda\cdot(1 + b/a)\cdot S.$$

Therefore, for given A, a higher S signals a higher H/K at a given y, which predicts higher economic growth. However, for a given H/K, a higher S would signal a lower A at a given y, which predicts lower economic growth. In the empirical analysis, I assume that, although S and y are observable, A (corresponding to total factor productivity) cannot be measured (because K cannot be measured accurately). In this case, the overall effect of S on economic growth, for given y, is ambiguous. In the empirical analysis, the estimated effect on growth from the level of S turns out to be small and typically statistically insignificantly different from zero.

III. Cross-Country Growth Regressions

The preceding section implies that the growth rate of real per capita GDP, Dy_{it}, for country i at date t can be written as

$$Dy_{it} = \Phi(\underset{(-)}{y_{it}}, \underset{(+)}{s_{it}}, \underset{(-)}{n_{it}}, \underset{(+)}{A_{it}}, \ldots), \tag{5}$$

where the negative sign under y_{it} reflects convergence, conditional on the other variables. These other influences include variables related to the saving rate, s_{it}, the population growth rate, n_{it}, and the level of productivity, A_{it}. In the main analysis, I estimate equation (5) using per capita growth rates of GDP averaged over 5-year periods. Sometimes the analysis includes constant terms that are specific to countries (country fixed effects). The estimation always includes constants for each time period (time effects)—therefore, the analysis does not attempt to explain variations over time of world average growth rates.

One well-known problem with cross-country growth regressions is endogeneity of some of the X variables. For example, it is unclear whether good institutions cause economic growth or are a reaction to rising living standards—or, perhaps, that GDP and institutional quality are responses to common influences. Previous studies have proposed instruments to deal with this problem. Examples are gravity variables such as country size and trade restrictions that influence international trade (Lee [1993]); ethnolinguistic fractionalization (Mauro [1995]); population density and settler mortality at the time of colonial settlement (Engerman and

Sokoloff [1994],[13] Acemoglu, Johnson, and Robinson [2001, 2002]); the form of legal origins (La Porta, Lopez-de-Silanes, Shleifer, and Vishny [1998]); absolute degrees latitude and primary language (Hall and Jones [1999]); the presence of state religion (Barro and McCleary [2003]); and physical characteristics of islands (Feyrer and Sacerdote [2009]). One problem is that the proposed instruments typically do not vary over time in the sample within countries and, therefore, do not help when country fixed effects are included. A more basic issue, if one allows for the multi-dimensional set of X variables that matters for economic growth, is that there are never enough convincing instruments to allow for full identification.

The present empirical analysis emphasizes panel least-squares but compares these results with two-state least-squares estimates based on lagged values of the X variables as instruments. For example, in considering the growth rate from 2005 to 2010, the average of the investment ratio from 2005 to 2009 enters into the regression equation but the investment ratio for 2005 is on the instrument list. This use of lagged X values as instruments helps to deal with endogeneity in some contexts (and also to alleviate problems of temporary measurement error). However, this approach is not fully satisfactory because of the strong serial correlation in some of the X variables.

Another well-known issue is the robustness of the results with respect to which X variables are included and in what functional form. Barro and Sala-i-Martin (2004, Ch. 12) discuss the Bayesian model-averaging approach to this problem. This technique effectively weights each possible specification by the fits to the dependent variable, taken to be the growth rate of per capita GDP from 1960 to 1996. This method was applied to 67 X variables that have

[13]Engerman and Sokoloff (1994) emphasize the density of the indigenous population and the nature of land and climate as influences on institutions and, thereby, on long-term economic growth. Thus, the exogenous elements in their historical analysis, such as population density at the time of colonial settlement, can be viewed as informal instruments for descriptive regressions.

been proposed in the empirical growth literature, using data for 88 countries. The conclusion is that only 5 of the variables have posterior inclusion probabilities above 0.5 and 18 have probabilities above 0.1. However, with 67 variables considered, many are conceptually similar, so that low inclusion probabilities are not surprising.

My view is that pinpointing precisely which X variables matter for growth is impossible. However, what is feasible is interpreting the results in terms of broad influences that matter for growth; for example, quality of institutions, openness to markets, and so on. In addition, one can interpret results on conditional convergence—gauged empirically by the estimated coefficient on the log of lagged per capita GDP—holding fixed an array of X variables. These results on conditional convergence tend not to be highly sensitive to exactly which X variables are included.

IV. Empirical Results

Table 1 contains empirical results for the cross-country panel. The dependent variable is each country's growth rate of real per capita GDP over 10 5-year intervals from 1960-65 to 2005-10. Note that the growth rates are expressed per year, not per five-year period. The sample of countries was chosen based on the availability of data on GDP at least by 1970, so that each country has at least 40 years of data. This sample, corresponding to columns 1 and 2, comprises 151 countries. Subsequent estimation in columns 3 and 4 includes an array of X variables. In this case, the criterion of data availability at least by the 1970-75 period led to the selection of 89 countries, listed in Table 2. Estimation in Table 1, columns 1-4, is by panel least-squares[14] and includes time effects.

[14]Standard errors of coefficient estimates allow each country's error term to be correlated over time.

Column 1 of Table 1 includes as a right-hand-side variable only the 5-year lag of the log of per capita GDP. The estimated coefficient is positive (indicating divergence rather than convergence) and statistically significantly different from zero. However, the coefficient is small in magnitude, in the sense of indicating divergence at a rate of only 0.2% per year. This result reproduces the typical pattern whereby absolute convergence of per capita GDP does not appear in a heterogeneous collection of countries. From the standpoint of equation (5), the interpretation is that the 5-year lag of the log of per capita GDP is positively correlated with determinants of the steady-state position that raise the long-run level of per capita GDP. This omitted-variables effect offsets the convergence force and leads to an estimated coefficient on the lagged log of per capita GDP that is positive but close to zero. The Monte Carlo analysis in the appendix (Table A1, line 11) reproduces this kind of result.

Column 2 adds country fixed effects, which are jointly highly statistically significant. The estimated coefficient of the log of lagged per capita GDP is now significantly negative, -0.0335 (s.e.=0.0039).[15] This result suggests conditional convergence (conditional on individual constants for each country) at a rate around 3.4% per year. An interpretation is that each country's fixed effect proxies for the influences of the various determinants of long-run per capita GDP that appear in equation (5)—at least to the extent that these determinants do not vary much over time within countries. Therefore, the estimated coefficient on the log of lagged per capita GDP now picks up the predicted conditional convergence, indicated by the negative sign in equation (5). However, as noted before, there is a tendency to overestimate the convergence speed in the presence of country fixed effects because of the Hurwicz-Nickell bias in the

[15]In this and the subsequent regressions, one can think equivalently of the dependent variable as the log of the level of per capita GDP (observed for 1965, ..., 2010), with the five-year lag of this variable included on the right-hand side. In this version of the regression from Table 1, column 2, the coefficient on the lagged dependent variable is 0.83 (which equals 1 minus 5 times the estimated convergence coefficient of 0.0335 per year).

estimated coefficient of a lagged dependent variable. The Monte Carlo analysis (Table A1, lines 1 and 2) supports this interpretation.[16]

Column 3 includes, instead of country fixed effects, an array of time-varying X variables for each country.[17] These variables proxy for the various growth determinants, aside from the log of lagged per capita GDP, that enter into equation (5). Unlike in column 1, the estimated coefficient of the log of lagged per capita GDP is significantly negative, -0.0170 (s.e.=0.0021), and indicates convergence at about 1.7% per year. This convergence is conditional in the sense of holding for given values of the X variables. The details on the X variables used in the estimation are in the notes to Table 1.

The estimated coefficients of the X variables in column 3 can be viewed mostly as effects on long-run or steady-state positions for each country. For example, with respect to institutional quality, the results suggest that a country's long-run economic position is enhanced by better maintenance of law and order (and the rule of law).[18] Greater democracy (gauged by the level

---

[16]Several dynamic panel estimators have been developed to alleviate this problem, but none of these approaches seem to be reliable. For example, Blundell and Bond (1998, pp. 115, 116, 120) note that the method of Arellano and Bond (1991)—which uses lagged levels of the dependent variable as instruments in a first-difference form of the model—suffers from a weak-instruments problem and has large finite-sample bias. The problems are especially serious when the coefficient of the lagged dependent variable in a level form of the annual equation is close to one—as in the growth context—and when the relative variance of the fixed effect is large. I find in the Monte Carlo setting described in Appendix A that the Arellano and Bond (1991) estimator sharply over-estimates the convergence rate. In contrast, the Blundell and Bond (1998) estimator (which uses lagged levels and differences of the dependent variable as instruments in a system of equations) turns out to sharply under-estimate the convergence rate. These Monte Carlo findings accord with results obtained with the data and functional form from Table 1, column 2. In this setting, the estimated coefficient of the log of lagged per capita GDP (using annual data) is -0.048 (s.e.=0.010) with the Arellano and Bond (1991) estimator and 0.005 (s.e.=0.001) with the Blundell and Bond (1998) estimator. Another alternative is the bias-adjustment procedure of Kiviet (1995), made operational by Bruno (2005). One problem is that this approach requires, in the first stage, a consistent estimate of the coefficient of the lagged dependent variable. In practice, Bruno (2005) relies on Arellano and Bond (1991), which is problematic because of the large bias in this estimator. Using the data and functional form from Table 1, column 2, the Bruno (2005) procedure yields an estimated coefficient of the log of lagged per capita GDP of -0.007 (s.e.=0.003).

[17]I considered some variables that were constant over time within countries—the absolute value of degrees latitude, land-locked status, and aspects of colonial history and legal origins—but these variables turned out to be unimportant.

[18]The law-and-order variable (previously called rule-of-law) comes from *International Country Risk Guide*, which is produced by Political Risk Services. These data were first used in academic research by Knack and Keefer (1995). The law-and-order indicator starts in 1982 or later. The panel regressions in Table 1, column 3, use each country's

and square of the Freedom House indicator of political rights[19]) tends initially to be positively associated with growth, but the sign switches for higher values of democracy. The break point between marginal effects being positive or negative occurs roughly at the half-way mark between full dictatorship (value zero) and full representative democracy (value one). Analogous relations between democracy and economic growth were reported in Barro (1997, Ch. 2).

Other findings in Table 1, column 3, are that countries' long-run positions are positively associated with a lower mortality rate (gauged by the reciprocal of life expectancy at birth), a lower fertility rate, and higher female relative to male school attainment for persons aged 15 and over.[20] The schooling effect does not relate to the overall level of human capital in the sense of total years of attainment. In fact, the estimated effect from a general increase in attainment—where average years of female and male schooling rise by equal amounts—differs insignificantly from zero, a result that is consistent with the conceptual analysis of human capital presented earlier. A reasonable interpretation is that an expansion in female relative to male attainment signals an improvement more generally in political and social arrangements that support economic growth.

The results in Table 1, column 3, show a significantly positive effect from greater international openness (exports plus imports relative to GDP) and from a higher growth rate of a country's terms of trade (entered as an interaction with the international-openness variable). The estimates also reveal a significantly positive coefficient for the ratio of investment to GDP and a significantly negative coefficient for the inflation rate. However, instrumental estimates

---

first value available for this variable for the periods that start in 1980-85 or earlier. If the sample covers only the five periods that begin with 1985-90, the results are similar to those in column 3. An alternative to the ICRG data is the information on perceived quality of governance assembled by the World Bank (available at www.govindicators.org). However, these data are available only since the late 1990s.

[19]Analogous data from Bollen (1980) were used for 1960 and 1965.

[20]The school-attainment data are updated values described in Barro and Lee (2013) and available at www.barrolee.com

discussed later suggest that these last results may reflect reverse causation from growth to investment (positive) or inflation (negative), rather than the reverse. Finally, the estimated coefficient on the ratio of government consumption to GDP is negative but not statistically significantly different from zero.

As discussed before, the Hurwicz-Nickell bias in the estimated coefficient of a lagged dependent variable is likely to be small in the context of column 3, where country fixed effects are absent. The magnitude of the estimated coefficient of the log of lagged per capita GDP would tend to be underestimated if there are still important omitted determinants of long-run per capita GDP that are conditionally correlated with per capita GDP. However, given the substantial list of growth determinants included, this omitted-variables bias may be small (unlike in column 1). Therefore, it is possible that OLS without country fixed effects can deliver accurate estimates of the convergence rate, as confirmed in the appendix by Monte Carlo analysis (Table A2, lines 9-12).

Table 1, column 4, allows for country fixed effects along with the X variables. The fixed effects are still jointly highly statistically significant. The estimated coefficient of the log of lagged per capita GDP, -0.0458 (s.e.=0.0045), is negative, statistically significant, and larger in magnitude than in columns 2 or 3. The indicated conditional convergence rate is around 4.6% per year. However, as discussed before, the convergence rate tends to be overestimated because of the Hurwicz-Nickell bias (as confirmed by the Monte Carlo studies considered in the appendix; see Table A2, lines 1 and 3).[21]

[21]Hauk and Wacziarg (2009, Table 4, column 1) used a Monte Carlo analysis to evaluate the Arellano and Bond (1991) and Blundell and Bond (1998) estimators in this context. They find, consistent with the results for a model without X variables described in n. 18, that Arellano and Bond (1991) substantially over-estimates the convergence rate, whereas Blundell and Bond (1998) substantially under-estimates it.

Another effect from the introduction of country fixed effects is that the standard errors of the estimated coefficients of the X variables are mostly higher in column 4 than in column 3. This pattern arises because, with country fixed effects included, only within-country variation over time is used to identify the coefficients. Consequently, some of the estimated coefficients that were statistically significantly different from zero in column 3 are no longer statistically significant in column 4. As an example, although the estimated coefficient on the indicator for the maintenance of law and order was significantly positive in column 3, this variable no longer has significant explanatory power in column 4. The likely explanation is that there is insufficient within-country variation in the measured institutional quality to isolate a statistically significant effect on economic growth.

Table 3 contains additional panel regressions for growth rates. Column 1 parallels Table 1, column 3, but uses two-stage least squares instead of OLS. The instrument list replaces the 5-year lag of the log of per capita GDP with the 6-year lag and replaces averages of several variables lagged 1-to-5 years with 5-year-lagged values. Thus, the idea is to use longer lags of some of the X variables as instruments. The results in Table 3, column 1, are similar in most respects to the OLS results (Table 1, column 3), except that the estimated coefficients of the investment ratio and the inflation rate are no longer statistically significantly different from zero. These results likely reflect the joint short-run determination of GDP with investment and inflation, whereby the typical pattern is for the investment ratio to be procyclical and the inflation rate to be countercyclical. Instrumenting with five-year lags eliminates most of this high-frequency interaction, and the remaining partial association of economic growth with investment and inflation turns out to be weak. The bottom line is that the OLS results on these estimated coefficients (in Table 1, column 3) cannot be reliably interpreted as isolating causation from the investment ratio or the inflation rate to GDP growth.[22]

An important finding in Table 3, column 1, is that the estimated coefficient on lagged per capita GDP, -0.0173 (s.e.=0.0022), is virtually the same as that in Table 1, column 3. Thus, the estimated convergence rate—around 1.7% per year—is robust to using longer lags of some of the X variables as instruments. These instruments would be imperfect for making causal inferences about the effects of X variables when these variables are serially correlated. However, the robustness in the estimated coefficient of lagged GDP suggests that endogeneity of the X variables may not be a central issue for the purpose of estimating convergence rates.

Table 3, column 2, uses 10-year averages for growth rates, with corresponding adjustments in the definitions of the X variables. The results are similar to those from the 5-year case in Table 1, column 3. In particular, the estimated convergence rates per year are close: -0.0166 (s.e.=0.0020) for the 10-year case, versus -0.0170 (s.e.=0.0021) for the 5-year. The same general conclusions hold with annual data on GDP growth rates. For the convergence rate, in a system without country fixed effects, the OLS estimate of the convergence coefficient is -0.0167 (s.e.=0.0023), similar to those in Table 1, column 3, and Table 3, column 2.[23]

The results show that changing the time dimension of the sample by observing the data more or less frequently—shifting, say, from 10-year to 5-year to 1-year periods—has minor implications, particularly for the estimated convergence rate. As discussed before, this finding is

[22]This conclusion accords with the one reached with respect to the investment ratio in Blomströ̋m, Lipsey, and Zejan (1996).
[23]The annual system is problematic because some of the data (school attainment, changes in the terms of trade, and the inflation rate) were constructed only over 5-year intervals. Moreover, other data—notably on life expectancy and the fertility rate—are provided at an annual frequency but are not really annual variables. In any event, the annual results (using interpolated values of the variables available at 5-year intervals) show much poorer fits than the systems at 5-year and 10-year intervals. The R-squared value for the annual case is 0.132, and the standard error of the regression is 0.0505. (This system for 89 countries has 3760 total observations.)

consistent with the formula for the Hurwicz-Nickell bias derived by Nickell (1981, p. 1422) and shown in equation (1).

Table 3, column 3, goes further in the exploration of the implications of country fixed effects by allowing for two separate sets of country dummy variables—one applying to the first five periods (1960-65, …, 1980-85) and the other to the second five (1985-90, …, 2005-10). The break in the set of country fixed effects from the first half to the second half of the sample is jointly highly statistically significant, as are the fixed effects overall. Not surprisingly, the standard errors of the estimated coefficients of all the X variables are higher than those in Table 1, column 4 (with one set of country fixed effect), which were in turn mostly higher than those in Table 1, column 3 (with no fixed effects). That is, the richer structure of country fixed effects makes it even more difficult to use the within-country variation to estimate the effects of the X variables. For present purposes, the most interesting result is that the estimated coefficient on the log of lagged per capita GDP, -0.0745 (s.e.=0.0066), becomes even larger in magnitude than that in Table 1, column 4, now indicating convergence at about 7.4% per year. A reasonable interpretation is that the extra set of country fixed effects effectively cuts the time dimension of the sample by half (from 50 years to 25 years) and, thereby, intensifies the Hurwicz-Nickell bias. Hence, these findings provide a further warning that country fixed effects can cause serious upward bias in estimated convergence rates.

Finally, Table 3, column 4, uses as a democracy indicator the Polity measure of democracy/autocracy, rather than the Freedom House measure of political rights. The overall results are similar to those based on the Freedom House indicator (Table 1, column 3), but the estimated coefficients on the Polity variable and its square are neither individually nor jointly statistically significantly different from zero.

My inference from the cross-country data starting in 1960 is that the most reliable estimates of convergence rates come from systems that exclude country fixed effects but include an array of X variables to mitigate the consequences of omitted variables. That is, I would emphasize the results in Table 1, column 3 (based on OLS) or Table 3, column 1 (which uses longer lags of some of the X variables as instruments). However, as I discuss later, the fixed-effects results seem more reliable in systems for the more limited set of countries that have data over a century or more.

V. Modernization

The modernization thesis applied to democracy is that economic development—gauged particularly by per capita GDP and education—promotes democratic institutions. This idea was emphasized by Lipset (1959), who credits the concept to Aristotle.[24] Glaeser, Ponzetto, and Shleifer (2007) provide a theoretical rationale for the effect of education on democracy through the channel of higher education motivating greater participation in political and other social activities. The Aristotle-Lipset hypothesis can be extended beyond democracy to apply to institutional quality, including the indicator used earlier for maintenance of law and order.

Barro (1999) provided empirical confirmation of the Aristotle-Lipset hypothesis in a cross-country panel, with stress on the Freedom House measure of political rights. Additional supporting evidence along these lines appears in Glaeser, La Porta, Lopez-de-Silanes, and Shleifer (2004) and Glaeser, Ponzetto, and Shleifer (2007). However, these results have been challenged by Acemoglu, Johnson, Robinson, and Yared (2005, 2008), who argue that education and per capita GDP do not have statistically significant influences on democracy. As in

[24]According to Lipset (1959, p. 75): "From Aristotle down to the present, men have argued that only in a wealthy society in which relatively few citizens lived in real poverty could a situation exist in which the mass of the population could intelligently participate in politics and could develop the self-restraint necessary to avoid succumbing to the appeals of irresponsible demagogues."

estimating convergence rates for economic growth, the essence of much of this recent debate about modernization turns on whether the empirical framework includes country fixed effects.

Table 4 contains panel regressions in which the dependent variables are indicators of institutional quality. The estimation method, parallel to Table 1, is panel least-squares. Columns 1 and 2 use *International Country Risk Guide*'s indicator for maintenance of law-and-order (previously described by ICRG as maintenance of the rule of law). Columns 3 and 4 use the Freedom House measure of democracy (political rights), and columns 5 and 6 use the Polity measure of democracy (calculated as differences between the Polity measures of democracy and autocracy). To make the analysis comparable to that for economic growth, the estimation applies to the sample of countries used in Table 1, columns 3 and 4, and listed in Table 2.[25] Because of data availability, the time frame for the law-and-order indicator is from 1985 to 2010 and that for the Freedom House measure of democracy is from 1970 to 2010. The Polity analysis is limited to 1970-2010 to make the results comparable to those for Freedom House.

The framework for the modernization equations parallels that for economic growth. Thus, for five-year periods, if Y_t is the level of an indicator of institutional quality in year t, then the dependent variable is $(Y_t-Y_{t-5})/5$, and the lag associated with this dependent variable is Y_{t-5}. The system includes on the right-hand side a set of X variables, analogous to those used to explain economic growth. In this setting, the coefficient on Y_{t-5} gives the convergence rate per year over a five-year period. This convergence is conditional on the X variables, just as in the case of economic growth.

In the present study of institutional indicators, the X variables are limited to those stressed in the literature on modernization—the log of per capita GDP and measures of school attainment (distinguished by gender). The variables used in the regressions—for the log of per

[25]Iceland, Luxembourg, and Malta are excluded in the case of Polity because of missing data.

capita GDP and the schooling variables—are 5-year lags, as in the equations for economic growth.

Table 4, column 1, applies to the law-and-order indicator. The estimated convergence coefficient, -0.0577 (s.e.=0.0056), is statistically significant and indicates conditional convergence at around 5.8% per year. Therefore, the convergence rate for this measure of institutional quality is notably higher than that for the log of per capita GDP, as computed from the analogous growth-rate equation in Table 1, column 3.

With respect to X variables, the estimated coefficient on the log of per capita GDP (lagged five years) in Table 4, column 1, is significantly positive, 0.0051 (s.e.=0.0018). The sum of the estimated coefficients for female and male school attainment (lagged five years) is also significantly positive (p-value=0.020). However, the estimated coefficient for male schooling is positive and statistically significant (0.0056, s.e.=0.0015), whereas that for female schooling is negative and statistically significant (-0.0037, s.e.=0.0014). The results reject with a p-value of 0.000 the joint hypothesis that the coefficient on the GDP variable is zero and the sum of the coefficients for female and male schooling is zero against the alternative that each of these is positive. Hence, the system that excludes country fixed effects provides evidence of modernization with regard to the overall association between the law-and-order indicator and the GDP and schooling variables.

As usual, this least-squares estimation does not provide clear evidence of causation; in this case, from GDP and school attainment to law-and-order, rather than the reverse. However, using two-stage least-squares, with longer lags of the GDP and school-attainment variables used as instruments, produces only minor changes in the conclusions. This finding applies also to the other equations in Table 4.

Table 4, column 2, shows the impact of allowing for country fixed effects, which are jointly highly statistically significant. The results change compared with column 1 in ways familiar from the analysis of economic growth: the estimated convergence rate rises sharply, to around 13% per year, and the standard errors of all the estimated coefficients of the explanatory variables increase substantially. Consequently, none of the estimated coefficients on the three X variables (for the log of per capita GDP, female schooling, and male schooling) are individually statistically significantly different from zero. The sum of the two coefficients for the schooling variables is also insignificant (p-value=0.82). Hence, the inclusion of country fixed effects eliminates the evidence in favor of modernization with regard to the law-and-order indicator.

Results for democracy are in Table 4, columns 3-6. For the equations without country fixed effects, the estimated convergence coefficients are negative and statistically significant for the Freedom House measure (column 3) and the Polity measure (column 5). The estimated rates of convergence are 5-6% per year, similar to that found for the law-and-order indicator (column 1). The estimated coefficient on the lagged GDP variable is significantly positive for the Freedom House data (column 3) but not for the Polity data (column 5). The sums of the estimated coefficients on the schooling variables are significantly positive, with p-values of 0.001 for Freedom House (column 3) and 0.003 for Polity (column 5). The results reject with p-values of 0.000 the joint hypothesis that the coefficient on the GDP variable is zero and the sum of the coefficients for female and male schooling is zero against the alternative that each of these is positive (columns 3 and 5). Thus, the results without country fixed effects support the modernization hypothesis with respect to the two measures of democracy.

The introduction of country fixed effects again raises the estimated rates of convergence—to 12% per year in column 4 and 10% in column 6. As before, the standard errors of all the estimated coefficients rise substantially. This change makes the estimated coefficients on the GDP and schooling variables individually and jointly insignificant (columns 4 and 6). Hence, results with country fixed effects do not support the modernization hypothesis with regard to the measures of democracy.

Acemoglu, et al. (2005, 2008) rely on results with country fixed effects to argue that the modernization hypothesis is not supported empirically in cross-country panel data. My inference is that their results signal the dangers from the inclusion of country fixed effects: this procedure reduces the information contained in the panel data (by blowing up the standard errors of the estimated coefficients of X variables) and biases upward the estimated convergence rates.[26] I show in a later section that the results with country fixed effects look different—and supportive of the modernization hypothesis—in systems estimated over time frames of over a century.

For systems estimated with data spanning 25 to 40 years—starting between 1970 and 1985 and ending in 2010—the most reliable (though imperfect) information on the determinants of institutional quality likely comes from systems that exclude country fixed effects. Thus, in Table 4, I would emphasize the findings in columns 1, 3, and 5. These results support the modernization hypothesis with regard to the relationships of GDP and schooling to the quality of institutions, gauged by law and order and democracy.

[26]Acemoglu, et al. (2008, p. 820) argue that this bias in the fixed-effects estimated coefficient of a lagged dependent variable can be mitigated by observing the data at a higher frequency, such as annually: "Column 6 [of Table 3] estimates … with OLS using annual observations. This is useful since the fixed effect OLS estimator becomes consistent as the number of observations becomes large. With annual observations, we have a reasonably large time dimension." The Nickell formula for the bias in equation (1) implies that this argument is incorrect. The proportionate bias in the estimated coefficient of a lagged dependent variable depends on the overall length of the sample in years (or, actually, on the product βT, where β is the convergence rate per year and T is the sample length in years), not on the frequency of observation of the data.

VI. Long-Term Panels

Until recently, the best long-term macroeconomic panel data were the per-capita GDP series assembled by Maddison (2003). These series constitute a monumental contribution that has been widely used. However, the data have serious shortcomings, discussed and largely rectified in a new data set on per capita GDP and consumer expenditure assembled by Ursúa and me and described in Ursúa (2011).[27] The construction of these data was challenging, described by Ursúa as macroeconomic archeology, and various methods were implemented to include periods and countries originally missing or inadequately treated in standard sources. The new data set covers 42 countries (40 with nearly continuous annual data) and goes back at least to 1913 and in many cases to 1870 or earlier.

For the present analysis, I use a sample of 28 countries selected by the availability of annual data on per capita GDP by 1896 and by the availability of the Polity indicator for democracy. The list of countries is in the notes to Table 5. There is a limited supply of X variables available over this long time frame. My analysis uses the Polity indicator of democracy and recently constructed data since 1870 on average years of school attainment by females and males.[28]

A. Growth rates of per capita GDP

Columns 1-4 of Table 5 use as the dependent variable the growth rate of real per capita GDP over 5-year periods from 1870-75 to 2005-10. These systems include no X variables and,

[27]The data are available as "Barro-Ursúa Macroeconomic Data" at www.rbarro.com/data-sets. Full annual time series back to 1913 or earlier (except for a few missing observations around WWII) are provided for 40 countries for real per capita GDP and 28 countries for real per capita personal consumer expenditure. The data on the website are expressed as indexes for each country, with the value in 2006 normalized to 100. For the present study, the values were converted to levels comparable across countries based on data on 2005 PPP-adjusted real per capita GDP from the World Bank's *World Development Indicators*.
[28]The data from 1870 to 1945 at five-year intervals are unpublished estimates to be described in Barro and Lee (2015).

therefore, parallel the results in Table 1, columns 1 and 2. Column 1 excludes country fixed effects and includes as a regressor only the log of the five-year lag of per capita GDP (along with time effects). Unlike in Table 1, column 1, the estimated coefficient on the lagged log of per capita GDP is significantly negative at the 5% level, -0.0044 (s.e.=0.0013). However, the indicated convergence rate is only around 0.4% per year. Again, this low estimate likely reflects omitted X variables.

Table 5, column 2, adds country fixed effects. The estimated coefficient of the log of lagged per capita GDP, -0.0258 (s.e.=0.0041), shows convergence at around 2.6% per year.

I redid the analysis from Table 5, columns 1 and 2, by adding as X variables the Polity indicator of democracy and its square (as in Table 3, column 4) and the measures of average years of schooling for females and males aged 15 and over (as in Table 1, column 3, and Table 3, column 4).[29] Without country fixed effects (Table 5, column 3), the estimated convergence rate becomes -0.9% per year, compared to -0.4% in Table 5, column 1. The interpretation is that the inclusion of the X variables lessens the problem of omitted variables and, thereby, tends to raise the estimated convergence rate. In the equations with country fixed effects (Table 5, columns 2 and 4), the estimated convergence rate depends little on whether the X variables are included. That is, the estimated convergence rate in column 4 is around 2.6% per year, similar to that in column 2.

In the long samples, the lack of information on the rich array of X variables considered in Table 1, columns 3 and 4, and Table 3, column 4, suggests that the omitted-variables bias would be substantial in systems that omit country fixed effects. Thus, the estimated convergence rate of

[29]One concern is that the comparison between the long-term and short-term results depends on sample selection with respect to which countries have long-term data. However, the post-1960 results in Table 1, column 3, do not differ greatly if the sample comprises only the 28 countries included in the long-term sample in Table 5. The estimated convergence rate for this 28-country sample observed since 1960 (266 total observations) is -0.0198 (s.e.=0.0040).

less than 1% per year in Table 5, columns 1 and 3, is likely to be a serious underestimate of the true convergence rate. On the other hand, because of the long time series, equation (1) implies that the Hurwicz-Nickell bias on the coefficient of the log of lagged per capita GDP—in Table 5, columns 2 and 4—is likely to be much less serious than that in Table 1, columns 2 and 4. Therefore, the estimated conditional convergence rate around 2.6% per year in Table 5, columns 2 and 4, may be only a small overestimate of the true convergence rate.

We can put the long-term results from Table 5 together with the shorter-term results from Table 1. In Table 1, the best estimate of the conditional convergence rate probably comes from column 3, which excludes country fixed effects and includes an array of X variables. The estimated conditional convergence rate here is around 1.7% per year, with a two-standard-error band of 1.3% to 1.9%. This result may be a small underestimate of the true convergence rate—to the extent that the system still omits long-run growth determinants that are conditionally correlated with per capita GDP. From Table 5, the best estimates likely come in columns 2 and 4 from equations with country fixed effects. The estimated convergence rate here is around 2.6% per year, with a two-standard-error band of 1.8% to 3.4%. This result may overestimate the true convergence rate—to the extent that the Hurwicz-Nickell bias still operates in this long-term sample.

When viewed together, the results from the two panel data sets suggest that the conditional convergence rate is, in terms of point estimates, bounded between 1.7% and 2.6% per year. Interestingly, this interval contains the iron-law rate of 2% per year.

B. Polity indicator of democracy

Columns 5 and 6 of Table 5 return to the modernization hypothesis by considering long-term panel regressions for the Polity indicator of democracy. Column 5 excludes country fixed effects. The estimated convergence rate, based on the estimated coefficient on the lagged Polity indicator, is -0.0390 (s.e.=0.0048). This coefficient is lower than that, -0.0505 (s.e.=0.0046), found with post-1970 data in Table 4, column 5.

From the standpoint of the modernization hypothesis, the important results in Table 5, column 5, are the significantly positive coefficient on the log of lagged per capita GDP, 0.0068 (s.e.=0.0022), and the significantly positive sum of the estimated coefficients on years of female and schooling (p-value=0.028). The GDP and schooling variables are jointly statistically significantly positive with a p-value of 0.000. This last finding parallels the result for post-1970 data in Table 4, column 5.

Table 5, column 6, adds country fixed effects. As usual, this change raises the magnitude of the estimated convergence rate—to -0.0510 (s.e.=0.0056). However, this rate is much lower than that (-0.1108, s.e.=0.0090) found with country fixed effects in the post-1970 data (Table 4, column 6). This result is expected because of the much longer time series employed in Table 5. With the variables observed over 140 years, the Hurwicz-Nickell bias in the presence of country fixed effects should be small (see the Monte Carlo results in the appendix).

In Table 5, column 6, the standard errors of the estimated coefficients of the per capita GDP and schooling variables go up, as before, due to the inclusion of country fixed effects. However, the estimated coefficient of the log of per capita GDP, 0.0108 (s.e.=0.0045), remains significantly positive. The sum of the estimated coefficients on the schooling variables is now close to zero. However, the coefficient on GDP and the sum of the coefficients on the schooling variables are jointly statistically significantly positive with a p-value of 0.056. Hence, even with the inclusion of country fixed effects, the long-term sample supports the modernization hypothesis with regard to the overall relationship of GDP and schooling to the Polity indicator for democracy.[30]

VII. Summary Observations

For a large panel of countries since 1960, the estimated conditional convergence rate for per capita GDP is around 1.7% per year in systems with a rich array of explanatory variables but no country fixed effects. The results feature statistically significant influences on economic growth from institutional quality, measured by maintenance of law and order and democracy. Analogous settings support the modernization hypothesis, in the sense of a significantly positive overall relationship of per capita GDP and schooling to indicators for maintenance of law and order and democracy.

The inclusion of country fixed effects produces much higher convergence rates per capita GDP, eliminates statistically significant effects of the institutional measures on economic growth, and removes the statistical support for modernization. I argue that these problematic findings reflect econometric issues associated with the inclusion of country fixed effects in panels with moderate time dimension.

In a panel for a limited number of countries observed since 1870, the estimated conditional convergence rate for per capita GDP in the presence of country fixed effects is about 2.6% per year. This setting also supports the modernization hypothesis, in the sense of a statistically significant, positive relationship between per capita GDP and democracy. These "reasonable" results with country fixed effects likely arise because the econometric problems posed by country fixed effects may not be serious in samples with long time frames.

[30]Acemoglu, et al. (2008, Table 7) argue that lagged per capita GDP lacks significant explanatory power for the Polity indicator over long samples for 25 countries starting in 1875. However, their analysis relies on the questionable Maddison data and—because of limitations in the availability of the Maddison data—considers effects only with a lag of 25 years. (They also fail to find significant effects at 50-year lags for a larger sample of countries.)

A combination of the results from the post-1960 and post-1870 panels suggests that the conditional convergence rate for per capita GDP is in the neighborhood of the “iron-law” rate of 2% per year. The similar findings from these two very different empirical settings suggest that a conditional convergence rate around 2% per year may be a robust empirical regularity. The two settings considered jointly also provide strong support for the modernization hypothesis.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91, December, 1369-1401.

Acemoglu, Daron, Simon Johnson, and James A. Robinson (2002). "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, 117, November, 1231-1294.

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared (2005). "From Education to Democracy?" *American Economic Review Proceedings*, 95, May, 44-49.

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared (2008). "Income and Democracy," *American Economic Review*, 98, September, 808-842.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt (2005). Competition and Innovation: An Inverted U Relationship," *Quarterly Journal of Economics*, 120, May, 701-728.

Aghion, Philippe, and Peter Howitt (1992). "A Model of Growth through Creative Destruction," *Econometrica*, 60, March, 323-351.

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

Arrow, Kenneth J. (1962). "The Economic Implications of Learning by Doing," *Review of Economic Studies*, 29, June, 155-173.

Aristotle (1932). *Politics*, translated by H. Rackham, Cambridge MA, Harvard University Press.

Barro, Robert J. (1991). "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, 106, May, 407-443.

Barro, Robert J. (1997). *Determinants of Economic Growth: A Cross-Country Empirical Study*, Cambridge MA, MIT Press.

Barro, Robert J. (1999). "Determinants of Democracy," *Journal of Political Economy*, 107, November, part 2, S158-S183.

Barro, Robert J. (2002). "Eastern Germany—A Lesson in Welfare Dependency," in R.J. Barro, *Nothing Is Sacred*, Cambridge MA, MIT Press.

Barro, Robert J. and Jong-Wha Lee (2013). "A New Data Set of Educational Attainment in the World, 1950-2010," *Journal of Development Economics*, 104, September, 184-198.

Barro, Robert J. and Jong-Wha Lee (2015). *Education and Modernization Worldwide: From the 19th through the 21st Century*, New York, Oxford University Press, forthcoming.

Barro, Robert J. and Rachel M. McCleary (2003). "Religion and Economic Growth," *American Sociological Review*, 68, October, 760-781.

Barro, Robert J. and Xavier Sala-i-Martin (1992). "Convergence," *Journal of Political Economy*, 100, April, 223-251.

Barro, Robert J. and Xavier Sala-i-Martin (2004). *Economic Growth*, 2nd ed., Cambridge MA, MIT Press.

Becker, Gary S. (1991). "The Demand for Children," chapter 5 in *A Treatise on the Family*, Cambridge MA, Harvard University Press.

Blomstrőm, Magnus, Robert E. Lipsey, and Mario Zejan (1996). "Is Fixed Investment the Key to Economic Growth? *Quarterly Journal of Economics*, 111, February, 269-276.

Blundell, Richard and Stephen Bond (1998). "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, November, 115-143.

Bollen, Kenneth A. (1980). "Issues in the Comparative Measurement of Political Democracy," *American Sociological Review*, 45, June, 370-390.

Bruno, Giovanni S.F. (2005). "Estimation and Inference in Dynamic Unbalanced Panel-Data Models with a Small Number of Individuals," *The Stata Journal*, 5, 4, 473-500.

Caselli, Francesco and Wilbur John Coleman (2001). "Technology Diffusion: The Case of Computers," *American Economic Review*, 91, May, 328-335.

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

Cass, David (1965). "Optimum Growth in an Aggregative Model of Capital Accumulation," *Review of Economic Studies*, 32, July, 233-240.

Coe, David T. and Elhanan Helpman (1995). "International R&D Spillovers," *European Economic Review*, 39, May, 59-887.

Dahl, Robert A. (1991). *Polyarchy: Participation and Opposition*, New Haven CT, Yale University Press.

De Long, J. Bradford (1988). "Productivity Growth, Convergence, and Welfare: Comment," *American Economic Review*, 78, December, 1138-1154.

Engerman, Stanley L. and Kenneth L. Sokoloff (1994). "Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies: A View from Economic Historians of the United States," Historical paper no. 66, National Bureau of Economic Research, December.

Feyrer, James and Bruce Sacerdote (2009). "Colonialism and Modern Income: Islands as Natural Experiments," *Review of Economics and Statistics*, 91, May, 245-262.

Glaeser, Edward L., Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer (2004). "Do Institutions Cause Growth?" *Journal of Economic Growth*, 9, September, 271-303.

Glaeser, Edward L., Giacomo A.M. Ponzetto, and Andrei Shleifer (2007). "Why Does Democracy Need Education?" *Journal of Economic Growth*, 12, June, 77-99.

Griliches, Zvi (1973). "Research Expenditures and Growth Accounting," in Bruce R. Williams, ed., *Science and Technology in Economic Growth*, New York, Macmillan.

Griliches, Zvi (1988). "Productivity Puzzles and R&D: Another Explanation," *Journal of Economic Perspectives*, 2, Fall, 9-21.

Griliches, Zvi and Frank Lichtenberg (1984). "R&D and Productivity Growth at the Industry Level: Is There Still a Relationship?" in Zvi Griliches, ed., *R&D, Patents, and Productivity*, Chicago, University of Chicago Press.

Grossman, Gene M. and Elhanan Helpman (1991). *Innovation and Growth in the Global Economy*, Cambridge MA, MIT Press.

Hall, Robert E. and Charles I. Jones (1999). "Why Do Some Countries Produce So Much More Output than Others?" *Quarterly Journal of Economics*, 114, February, 83-116.

Hauk, William R. and Romain Wacziarg (2009). "A Monte Carlo Study of Growth Regressions," *Journal of Economic Growth*, 14, June, 103-147.

Huntington, Samuel P. (1991). *The Third Wave: Democratization in the Late Twentieth Century*, Norman Oklahoma, University of Oklahoma Press.

Hurwicz, Leonid (1950). "Least-Squares Bias in Time Series," in Tjalling C. Koopmans, ed., *Statistical Inference in Dynamic Economic Models*, Wiley, New York.

Islam, Nazrul (1995). "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 110, November, 1127-1170.

Jorgenson, Dale W. and Zvi Griliches (1967). "The Explanation of Productivity Change," *Review of Economic Studies*, 34, July, 249-280.

Judson, Ruth A. and Ann L. Owen (1999). "Estimating Dynamic Panel Data Models: A Guide for Macroeconomists," *Economics Letters*, no. 1, 9-15.

Kiviet, Jan F. (1995). "On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models," *Journal of Econometrics*, 68, July, 53-78.

Knack, Stephen, and Philip Keefer (1995). ''Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures,'' *Economics and Politics*, 7, November, 207–227.

Knight, Frank H. (1944). "Diminishing Returns from Investment," *Journal of Political Economy*, 52, March, 26-47.

Knight, Malcolm, Norman Loayza, and Delano Villanueva (1993). "Testing the Neoclassical Theory of Economic Growth," *IMF Staff Papers*, 40, September, 512-541.

Koopmans, Tjalling C. (1965). "On the Concept of Optimal Economic Growth," in *The Econometric Approach to Development Planning*, Amsterdam, North-Holland.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert W. Vishny (1998). "Law and Finance," *Journal of Political Economy*, 106, December, 1113-1155.

Lee, Jong-Wha (1993). "International Trade, Distortions, and Long-Run Economic Growth," *IMF Staff Papers*, 40, June, 299-328.

Lipset, Seymour Martin (1959). ''Some Social Requisites of Democracy: Economic Development and Political Legitimacy,'' *American Political Science Review*, 53, March, 69–105.

Lucas, Robert E. (1988). "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, July, 3-42.

Maddison, Angus (1982). *Phases of Capitalist Development*, New York, Oxford University Press.

Maddison, Angus (2003). *The World Economy: Historical Statistics*, Paris, OECD, available at www.ggdc.net/maddison.

Malthus, Thomas R. (1798). *An Essay on the Principle of Population*, London, W. Pickering.

Mankiw, N. Gregory, David Romer, and David N. Weil (1992). "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, May, 407-437.

Manuelli, Rodolfo and Ananth Seshadri (2009). "Explaining International Fertility Differences," *Quarterly Journal of Economics*, 124, May, 771-807.

Marx, Karl (1904). *A Contribution to the Critique of Political Economy*, New York, International Library Publishing Company.

Mauro, Paolo (1995). ‘‘Corruption and Growth,’’ *Quarterly Journal of Economics*, 110, August, 681–712.

Nelson, Richard R. and Edmund S. Phelps (1966). "Investment in Humans, Technological Diffusion, and Economic Growth, *American Economic Review*, 56, May, 69-75.

Nerlove, Marc (2000). "Growth Rate Convergence: Fact or Artifact?" in Jaya Krishnakumar and Elvezio Ronchetti, eds., *Panel Data Econometrics: Future Directions*, Amsterdam, North-Holland, 3-34.

Nickell, Stephen (1981). "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49, November, 1417-1426.

Przeworski, Adam and Fernando Limongi (1993). "Political Regimes and Economic Growth," *Journal of Economic Perspectives*, 7, Summer, 51-69.

Ramsey, Frank (1928). "A Mathematical Theory of Saving," *Economic Journal*, 38, December, 543-559.

Rebelo, Sergio (1991). "Long-Run Policy Analysis and Long-Run Growth," *Journal of Political Economy*, 99, June, 500-521.

Rodrik, Dani (2013). "Unconditional Convergence in Manufacturing," *Quarterly Journal of Economics*, 128, February, 165-204.

Romer, Paul M. (1986). "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, 94, October, 1002-1037.

Romer, Paul M. (1987). "Growth Based on Increasing Returns due to Specialization," *American Economic Review Proceedings*, 77, May, 56-62.

Romer, Paul M. (1990). "Endogenous Technological Change," *Journal of Political Economy*, 98, October, part II, S71-S102.

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

Solow, Robert M. (1957). "Technical Change and the Aggregate Production Function," *Review of Economics & Statistics*, 39, August, 312-320.

Swan, Trevor W. (1956). "Economic Growth and Capital Accumulation," *Economic Record*, 32, November, 334-361.

Ursúa, José F. 2011. "Macroeconomic Archaeology: Unearthing Risk, Disasters, and Trends," unpublished, Harvard University, April.

Uzawa (1965). "Optimal Technical Change in an Aggregative Model of Economic Growth," *International Economic Review*, 6, January, 18-31.

Table 1 Growth-Rate Regressions for Cross-Country Panel
Five-year periods: 1960-65, …, 2005-10
(all equations estimated by OLS and include time effects)

	(1)	(2)	(3)	(4)
	No Fixed Effects	**Country Fixed Effects**	**No Fixed Effects**	**Country Fixed Effects**
Log(lagged per capita GDP)	0.0024* (0.0010)	-0.0335** (0.0039)	-0.0170** (0.0021)	-0.0458** (0.0045)
1/(life expectancy at birth)	--	--	-3.09** (0.58)	-1.03 (1.04)
Log(fertility rate)	--	--	-0.0277** (0.0043)	-0.0301** (0.0073)
Law & order (rule of law)	--	--	0.0157** (0.0054)	0.0051 (0.0088)
Investment ratio	--	--	0.031* (0.012)	0.053** (0.020)
Female school years	--	--	0.0024 (0.0014)	0.0054 (0.0033)
Male school years	--	--	-0.0028 (0.0015)	-0.0066* (0.0029)
Government consumption ratio	--	--	-0.026 (0.023)	-0.090* (0.042)
Openness ratio	--	--	0.0056* (0.0025)	0.0175* (0.0071)
Terms-of-trade change	--	--	0.117** (0.026)	0.113** (0.027)
Democracy indicator	--	--	0.029 (0.015)	-0.015 (0.019)
Democracy squared	--	--	-0.028* (0.014)	0.009 (0.017)
Inflation rate	--	--	-0.0180** (0.0042)	-0.0213** (0.0041)
R-squared	0.059	0.298	0.329	0.501
s.e. of regression	0.0371	0.0339	0.0242	0.0221
No. countries; observations	151; 1430	151; 1430	89; 841	89; 841

*Significant at 5% level. **Significant at 1% level.

Notes to Table 1

The dependent variable is the annual growth rate of real per capita GDP for the ten five-year periods: 1960-65, …, 2005-10. The sample criterion in columns 1-2 is to include countries only if they have data starting by the 1970-75 period (151 countries). The equations in columns 3-4 also require data on the array of X variables (89 countries, shown in Table 2). Lagged per capita GDP, the reciprocal of life expectancy at birth, the total fertility rate, and female and male years of school attainment for persons aged 15 and over are 5-year lags (for 1960, …, 2005). The ratios of investment and government consumption to GDP, the openness ratio, the indicator for law and order, and the democracy indicator are five-year averages of values lagged one to five years. The growth rate of the terms of trade and the inflation rate are for the same periods as the dependent variable. Standard errors of coefficient estimates are in parentheses. For calculating standard errors, the error terms are allowed to be correlated over time within countries. Joint p-values for the two democracy variables are 0.14 in column 3 and 0.56 in column 4.

Definitions and sources:

PPP-adjusted real per capita GDP is from Penn World Tables (www.pwt.econ.upenn.edu), version 7.0, in units of 2005 international dollars. Data for 2010 are from version 7.1. Also from version 7.0 are the ratios to GDP of investment (private plus public) and government consumption and the openness ratio (exports plus imports relative to GDP). These ratio variables use current-price information.

Life expectancy at birth and the total fertility rate are from the World Bank's *World Development Indicators* (*WDI*).

The law-and-order indicator is from Political Risk Services, *International Country Risk Guide*. The data were converted from seven categories to a 0-1 scale, with 1 representing the highest maintenance of law and order.

Average years of school attainment for females and males aged 15 and over at various levels of schooling are from Barro and Lee (2013), with data available at www.barrolee.com. These data are at 5-year intervals.

The terms-of-trade change (growth rates over five years of export prices relative to import prices) is from International Monetary Fund, *International Financial Statistics*, and *WDI*. This variable is interacted with the openness ratio.

The democracy indicator is the political-rights variable from Freedom House (www.freedomhouse.org). The data were converted from seven categories to a 0-1 scale, with 1 representing the highest rights. Data on an analogous concept for 1960 and 1965 are from Bollen (1980).

The inflation rate (averaged over 5-year intervals) is calculated from retail-price indexes from International Monetary Fund, *International Financial Statistics*, and *WDI*.

Table 2 Sample of 89 Countries Used in Regressions in Table 1, columns 3-4			
Country	**Starting period**	**Country**	**Starting period**
Argentina	1960-65	Jordan	1965-70
Australia	1960-65	Japan	1960-65
Austria	1960-65	Kenya	1960-65
Belgium	1960-65	South Korea	1965-70
Bangladesh	1965-70	Sri Lanka	1960-65
Bahrain	1970-75	Luxembourg	1960-65
Bolivia	1965-70	Morocco	1960-65
Brazil	1960-65	Mexico	1960-65
Botswana	1965-70	Mali	1965-70
Canada	1960-65	Malta	1970-75
Switzerland	1960-65	Malawi	1965-70
Chile	1960-65	Malaysia	1960-65
China	1960-65	Niger	1960-65
Cote d'Ivoire	1960-65	Nicaragua	1960-65
Cameroon	1965-70	Netherlands	1960-65
Congo, Republic	1960-65	Norway	1960-65
Colombia	1960-65	New Zealand	1960-65
Costa Rica	1960-65	Pakistan	1960-65
Cyprus	1960-65	Panama	1965-70
Denmark	1960-65	Peru	1965-70
Dominican Republic	1960-65	Philippines	1960-65
Algeria	1960-65	Papua New Guinea	1960-65
Ecuador	1960-65	Portugal	1960-65
Egypt	1960-65	Paraguay	1960-65
Spain	1960-65	Sudan	1970-75
Finland	1960-65	Senegal	1960-65
France	1960-65	Singapore	1965-70
Gabon	1965-70	Sierra Leone	1965-70
United Kingdom	1960-65	El Salvador	1960-65
Germany	1970-75	Sweden	1960-65
Ghana	1960-65	Syria	1970-75
Gambia	1965-70	Togo	1965-70
Greece	1960-65	Thailand	1960-65
Guatemala	1960-65	Trinidad	1960-65
Guyana	1970-75	Tunisia	1965-70
Honduras	1960-65	Turkey	1965-70
Haiti	1960-65	Taiwan	1960-65
Hungary	1970-75	Tanzania	1970-75
Indonesia	1965-70	Uganda	1965-70
India	1960-65	Uruguay	1965-70
Ireland	1960-65	United States	1960-65
Iceland	1960-65	Venezuela	1965-70
Israel	1970-75	South Africa	1960-65
Italy	1960-65	Zambia	1965-70
Jamaica	1960-65		

Table 3 Additional Growth-Rate Regressions for Cross-Country Panel (all equations include time effects and exclude country fixed effects)				
	(1)	**(2)**	**(3)**	**(4)**
	2SLS	**OLS Ten-year periods**	**OLS Two sets of fixed effects**	**OLS Polity variable**
Log(lagged per capita GDP)	-0.0173** (0.0022)	-0.0166** (0.0020)	-0.0745** (0.0066)	-0.0150** (0.0022)
1/(life expectancy at birth)	-3.32** (0.63)	-3.29** (0.58)	-2.32 (1.43)	-2.81** (0.60)
Log(fertility rate)	-0.0303** (0.0045)	-0.0242** (0.0042)	-0.0298** (0.0099)	-0.0260** (0.0044)
Law & order (rule of law)	0.0159** (0.0059)	0.0173** (0.0053)	0.0149 (0.0122)	0.0121* (0.0054)
Investment ratio	0.005 (0.014)	0.031* (0.012)	0.051* (0.024)	0.023 (0.013)
Female school years	0.0025 (0.0014)	0.0014 (0.0014)	0.0066 (0.0043)	0.0013 (0.0014)
Male school years	-0.0030* (0.0015)	-0.0019 (0.0014)	-0.0059 (0.0039)	-0.0017 (0.0015)
Government consumption ratio	-0.031 (0.026)	-0.014 (0.023)	-0.112* (0.052)	-0.037 (0.024)
Openness ratio	0.0072** (0.0028)	0.0042 (0.0027)	0.0224** (0.0085)	0.0053 (0.0028)
Terms-of-trade change	0.129** (0.027)	0.129* (0.038)	0.116** (0.029)	0.134** (0.027)
Democracy indicator	0.033 (0.018)	0.036* (0.016)	-0.005 (0.021)	0.022 (0.018)
Democracy squared	-0.035* (0.017)	-0.032* (0.014)	0.002 (0.019)	-0.023 (0.016)
Inflation rate	-0.0088 (0.0091)	-0.0171** (0.0055)	-0.0135** (0.0050)	-0.0172** (0.0042)
R-squared	0.322	0.448	0.616	0.316
s.e. of regression	0.0243	0.0179	0.0206	0.0237
No. countries; observations	89; 820	89; 407	89; 841	86; 755

*Significant at 5% level. **Significant at 1% level.

Notes to Table 3

See the notes to Table 1.

Column 1 corresponds to column 3 of Table 1, except that the instrument list replaces some of the explanatory variables with longer lags. For the log of lagged real per capita GDP, the instruments are the values for 1959, 1964, …, 2004. For the ratios of investment and government consumption to GDP, the openness ratio, the indicators for law-and-order and democracy, and the inflation rate, the instruments are 5-year lags. The p-value for the two democracy variables jointly is 0.094.

Column 2 corresponds to column 3 of Table 1, except that the dependent variable is the growth rate of per capita GDP for the five “10-year” periods 1960-70, …, 2000-10. The regressors are defined analogously for the 10-year periods.

Column 3 corresponds to column 4 of Table 1 but has one set of fixed effects for the five 5-year growth-rate observations from 1960-65 to 1980-85 and another set for the five observations from 1985-90 to 2005-10.

Column 4 is the same as column 3 of Table 1, except that the democracy indicator (the Freedom House measure of political rights) is replaced by the Polity measure of democracy less autocracy. The joint p-value for the two Polity variables is 0.35. The sample drops to 86 countries because of missing Polity data for Iceland, Luxembourg, and Malta.

Sources: The Polity indicator is for democracy less autocracy (converted from a -10 to +10 scale to a 0-1 scale, with 1 representing highest democracy), from Polity IV (www.systemicpeace.org). Other sources are given in the notes to Table 1.

Table 4 Regressions for Changes in Indicators of Law & Order and Democracy

Five-year periods. All equations estimated by OLS and include time effects.

	(1)	(2)	(3)	(4)	(5)	(6)
	Law & Order: 1985-90, …, 2005-10		Political Rights (Freedom House): 1970-75, …, 2005-10		Democracy/Autocracy (Polity): 1970-75, …, 2005-10	
	No Fixed Effects	Country Fixed Effects	No Fixed Effects	Country Fixed Effects	No Fixed Effects	Fixed Fixed Effects
Lagged indicator	-0.0577** (0.0056)	-0.130** (0.010)	-0.0581** (0.0052)	-0.1184** (0.0086)	-0.0505** (0.0046)	-0.1018** (0.0090)
Log(per capita GDP)	0.0051** (0.0018)	-0.0030 (0.0090)	0.0040* (0.0019)	-0.0051 (0.0082)	0.0012 (0.0018)	-0.0078 (0.0087)
Female school years	-0.0037** (0.0015)	0.0056 (0.0054)	0.0036* (0.0017)	0.0011 (0.0056)	0.0037* (0.0016)	0.0013 (0.0059)
Male school years	0.0056** (0.0015)	-0.0048 (0.0053)	-0.0007 (0.0017)	-0.0026 (0.0053)	-0.0013 (0.0017)	0.0019 (0.0056)
p-value for school-years variables > 0	0.020	0.82	0.001	0.67	0.003	0.34
p-value for GDP>0, school-years variables > 0	0.000	0.82	0.000	0.67	0.000	0.34
R-squared	0.392	0.584	0.173	0.356	0.187	0.331
s.e. of regression	0.0241	0.0223	0.0352	0.0333	0.0339	0.0330
No. countries; observations	89; 445	89; 445	89; 712	89; 712	86; 653	86; 653

*Significant at 5% level. **Significant at 1% level.

Notes to Table 4

See the notes to Table 1.

The dependent variable is $(Y_t-Y_{t-5})/5$, where Y_t is the level of an indicator variable in year t. Columns 1 and 2 use for Y_t the indicator for law and order from *International Country Risk Guide*, observed at the five dates: 1990, 1995, 2000, 2005, and 2010. Columns 3 and 4, use the Freedom House measure of political rights, observed at the eight dates: 1975, …, 2010. (Values for 1970 are interpolated based on data for 1965 (from Bollen [1980]) and 1972.) Columns 5 and 6 use the Polity measure of democracy less autocracy, observed at the eight dates: 1975, …, 2010. The explanatory variables apply to 1985, …, 2005 in columns 1-2 and 1970, …, 2005 in columns 3-6. Country fixed effects are included in columns 2, 4, and 6.

The p-value for school-years variables>0 is a test of the hypothesis that the sum of the coefficients of the school-years variables is zero against the alternative of being positive. The p-value for GDP>0, school-years variables>0 is a test of the hypothesis that the coefficient on the GDP variable is zero and the sum of the coefficients of the school-years variables is zero against the alternative that each of these is positive.

Table 5 Regressions for Long-Term Panels

Five-year periods: 1870-75, …, 2005-10

Estimation by panel least-squares; all equations include time effects.

	(1)	(2)	(3)	(4)	(5)	(6)
	Growth rate of per capita GDP				Polity Indicator	
	No Fixed Effects	Country Fixed Effects	No Fixed Effects	Country Fixed Effects	No Fixed Effects	Country Fixed Effects
Log(lagged per capita GDP)	-0.0044** (0.0013)	-0.0258** (0.0041)	-0.0090** (0.0020)	-0.0262** (0.0041)	0.0068** (0.0022)	0.0108* (0.0045)
Female school years	--	--	-0.0009 (0.0018)	-0.0026 (0.0025)	0.0025 (0.0019)	-0.0013 (0.0027)
Male school years	--	--	0.0020 (0.0019)	-0.0009 (0.0026)	-0.0008 (0.0021)	0.0015 (0.0028)
Democracy indicator (Polity)	--	--	-0.0296 (0.0161)	-0.0323 (0.0188)	-0.0390** (0.0048)	-0.0510** (0.0056)
Democracy indicator squared	--	--	0.0332* (0.0144)	0.0341* (0.0168)	--	--
p-value for democracy and democracy squared	--	--	0.025	0.082	--	--
p-value for school-years variables > 0	--	--	0.126	1.0	0.028	0.88
p-value for GDP>0, school-years variables > 0	--	--	--	--	0.000	0.056
R-squared	0.228	0.267	0.211	0.255	0.146	0.183
s.e. of regression	0.0281	0.0279	0.0267	0.0264	0.0285	0.0285
No. countries; observations	28; 764	28; 764	28; 727	28; 727	28; 700	28; 700

*Significant at 5% level. **Significant at 1% level.

Notes to Table 5

The sample criterion is to include countries only if they have GDP data starting by 1896 and also have data for most of the period on years of schooling and the Polity indicator of democracy. This criterion selected 28 countries: Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, China, Denmark, France, Germany, Italy, Japan, Mexico, Netherlands, New Zealand, Norway, Peru, Portugal, Russia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States, Uruguay, and Venezuela. Standard errors of coefficient estimates are in parentheses. In calculating standard errors of coefficient estimates, the error terms are allowed to be correlated over time within countries.

Columns 1-4: The dependent variable is the annual growth rate of real per capita GDP for the 28 countries for 28 periods: 1870-75, 1875-80, …, 2005-10. For the independent variables, the log of lagged per capita GDP, average years of female and male school attainment for persons aged 15 and over, and the Polity indicator are five-year lags, referring to 1870, 1875, …, 2005.

Columns 5-6: The dependent variable is $(Y_t\text{-}Y_{t\text{-}5})/5$, where Y_t is the level of the Polity indicator in year t. This variable applies over the same periods as in columns 1-4. The independent variables are defined as above.

Sources: GDP is from "Barro-Ursúa Macroeconomic Data," available at www.rbarro.com/data-sets. The source of the Polity indicator of democracy is given in the notes to Table 3. The data at 5-year intervals since 1950 on female and male average years of school attainment for persons aged 15 and over are discussed in the notes to Table 1. Data from 1870 to 1945 at five-year intervals are unpublished estimates to be described in Barro and Lee (2015).

Appendix

Monte Carlo Analysis of Dynamic Estimation with and without Country Fixed Effects

The dynamic model with fixed effects and no time-varying X variables follows Hurwicz (1950), Nickell (1981), Arellano and Bond (1991), and Nerlove (2000):

(A1) $$y_{it} = \eta_i + \gamma \cdot y_{i,\,t-1} + \varepsilon_{it},$$

where i = 1, …, N represents countries; t = 1, …, T represents periods; y_{it} is per capita GDP or some other country-time variable; $0<\gamma<1$; and ε_{it}~N(0, σ_ε^2) is an i.i.d. shock. I think of the fixed effect, η_i, as being drawn once for each country at the beginning of time, where η_i ~ N(0, σ_η^2) is an i.i.d. disturbance. As stressed by Nerlove (2000), the initial sample value y_{i0} cannot be viewed as independent of η_i, because y_{i0} comes from the cumulation of equation (A1) from the indefinite past; hence, y_{i0} depends on η_i and past realizations of the ε_{it}. Specifically, we have y_{i0} ~ N($\frac{\eta_i}{1-\gamma}$, $\frac{\sigma_\varepsilon^2}{1-\gamma}$). Therefore, a country that gets a high draw for η_i tends also to have a high y_{i0}.

The model with the addition of a time-varying, exogenous X variable is:

(A2) $$y_{it} = \eta_i + \gamma \cdot y_{i,\,t-1} + \alpha \cdot X_{it} + \varepsilon_{it},$$

(A3) $$X_{it} = \rho \cdot X_{i,t-1} + u_{it},$$

where η_i, γ, and ε_{it} are defined as before, α is a constant, $0<\rho<1$ governs the persistence of the X variable, and u_{it}~N(0, σ_u^2) is an i.i.d. shock. The initial sample value of the X variable is given accordingly by X_{i0} ~ N(0, $\frac{\sigma_u^2}{1-\rho}$). The initial sample value of y reflects its dependence on η_i and past shocks to ε_{it} and u_{it} from equations (A2) and (A3), and this dependence implies a relationship with X_{i0} (which depends on past shocks to u_{it}). Specifically, we can write

(A4) $$y_{i0} = \varphi \cdot X_{i0} + w_{i0},$$

where we can derive

(A5) $$\varphi = \alpha/[(1+\rho)(1-\gamma\rho)]$$

and w_{i0} ~$N(0, \sigma_w^2)$, independently of X_{i0}, where

(A6) $$\sigma_w^2 = \frac{\sigma_\varepsilon^2}{1-\gamma} + \frac{\alpha^2\sigma_u^2}{(1-\gamma\rho)^2}\left[\frac{\rho}{(1-\rho)(1+\rho)^2} + \frac{\gamma^2}{(1-\gamma)(1+\gamma)}\right].$$

Equations (A4) and (A5) imply that, if $\alpha>0$, a country that gets a high draw for X_{i0} tends also to have a high y_{i0}.

If periods are of length τ (years), the persistence coefficient, γ, relates to the convergence rate, $\beta>0$, from $\gamma = e^{-\beta\tau}$. I assume that $\beta\tau$ is much less than 1, so that $\gamma \approx 1-\beta\tau$ and, hence, $\beta \approx (1-\gamma)/\tau$. For annual periods ($\tau$=1 year), the convergence rate per year is $\beta \approx 1-\gamma$.

A. Model with Country Fixed Effects

Consider first the model in equation (A1) with a country fixed effect and no X variable. The standard deviation of the time-series shock, σ_ε, can be normalized to 1, so that σ_η represents the dispersion of the country (cross-sectional) shock, relative to the time-series shock. In Nickell (1981), the analysis applied as the number of cross sections, N, tended to infinity. In my Monte Carlo analysis, N (number of countries) = 20 seems large enough to approximate this asymptotic setting. I use N=100 below.

I contrast the results using OLS without country fixed effects with those using country fixed effects (OLS with dummy variables). As discussed in the text, the estimated coefficients γ and β from the two methods involve a tradeoff between two types of biases. I begin with a heuristic description of the results.

For OLS without country fixed effects, the Hurwicz-Nickell bias is unimportant—even if the time dimension, T, is small—but an omitted-variables bias applies. The omitted variables are the country fixed effects, η_i, which are positively correlated with the y_{it}. This effect biases up the estimated γ and, therefore, biases down the estimated β. This omitted-variables channel is more important the larger σ_η. For small enough σ_η, the omitted-variables effect is minor, and OLS without country fixed effects produces nearly unbiased estimates.

For OLS with country fixed effects, there are no omitted variables,[31] but the Hurwicz-Nickell bias may be large. This effect biases down the estimated γ and, therefore, biases up the estimated β. The size of the Hurwicz-Nickell bias depends on the length of the time series, T. If T is small—even 20 or 50 years—the bias is substantial. However, for large enough T, the bias becomes small, so that OLS with country fixed effects produces nearly unbiased estimates. However, even a sample of 140 years (the largest time frame considered in the text) is not sufficient to make the bias negligible. Unlike the case without country fixed effects, the size of σ_η is unimportant for the bias—because the estimation takes account of the variations in the η_i.

The upper part of Table A1 applies to OLS estimation with country fixed effects. Two values of the convergence rate, $\beta = 1-\gamma$, are considered, 0.02 and 0.10 per year. The emphasis in this part of the table is on the effect of the time dimension, T, on the mean of the estimated $\hat{\beta}$, shown in column 5. The bias is always upward, reflecting the Hurwicz-Nickell channel. For example, if β=0.02 per year, the mean of $\hat{\beta}$ when T=20 (line 1) is 0.151; that is, the upward bias is dramatic. The mean of $\hat{\beta}$ falls to 0.070 when T=50 (line 2) and 0.036 when T=140 (line 5). Hence, although the bias is much reduced compared to that applying to short time series, even T=140 years is insufficient to make the bias negligible. Changing the standard deviation σ_η of

[31]If the equation for y_{it} includes X variables—not all of which are included in the empirical specification—there would still be an omitted-variables problem.

the fixed effect does not impact the results (lines 4 and 9 of the table) because the fixed effects are estimated; that is, there is no omitted-variables problem here.

The results for $E(\hat{\beta})$ from Nickell's (1981) formula, shown in column 4 of Table A1, are similar to the Monte Carlo results for the mean of $\hat{\beta}$ in column 5. Some differences arise because Nickell's formula applies as the number of cross sections, N, approaches infinity. In addition, Nickell treated the initial sample values, y_{i0}, as given, rather than allowing for a relation with the fixed effect, η_i. This issue is more important in the estimation without country fixed effects (not considered by Nickell [1981]).

The lower part of Table A1 applies to OLS estimation without country fixed effects. The focus in this part of the table is on the impact of the standard deviation, σ_η, for the fixed effects. The bias in the estimate $\hat{\beta}$ is always downward, reflecting the omitted-variables channel. For example, when β=0.02, T=100, and σ_η=1 (line 11), the mean of $\hat{\beta}$ in column 5 is 0.0003. Hence, the bias is sharply downward in this case. However, the mean of $\hat{\beta}$ rises to 0.011 when σ_η=0.1 (line 12) and 0.020 when σ_η=0.01 (line 14). Thus, because the Hurwicz-Nickell bias is unimportant, the estimate $\hat{\beta}$ is virtually unbiased for a σ_η that is small enough to make the omitted-variables bias unimportant.

The time dimension, T, has a nonzero but moderate effect on the results without country fixed effects (lines 13 and 17 of Table A1). For example, reducing T from 100 to 20 (lines 12 and 13) actually lowers the bias: the mean of the estimate $\hat{\beta}$ goes from 0.0111 to 0.0126 (but the standard error of the estimate roughly doubles).

B. Model with Country Fixed Effects and an X Variable

Table A2 has the results for the model in equations (A2)-(A6), which adds an exogenous X variable. The conditions σ_ε=1 (for the time-series shock) and σ_u=1 (for the shock to the X variable) are normalizations. The other assumptions are β (convergence rate) = 0.02, σ_η=0.1 (for the fixed effect), ρ (persistence coefficient for the X process) = 0.98, and N (number of countries) =100. Results are shown for α (coefficient of the X variable) equal to 0.05 or 0.25 and T (time-series length) equal to 100 or 20.

Results from OLS without country fixed effects, shown in the lower part of Table A2, are straightforward. If the X variable is included in the regression (lines 9-12), so that omitted-variables effects are minor, the OLS estimates of the convergence rate, β (and also for α, the coefficient of the X variable) have small biases. These results follow because the Hurwicz-Nickell bias is again unimportant when fixed effects are excluded. Moreover, these findings apply for the different values considered for α and T. However, if the X variable is excluded from the regression (lines 13-16), the estimate $\hat{\beta}$ is seriously biased downward because of the omitted-variable effect described before. The bottom line is that satisfactory implementation of OLS without country fixed effects depends on the inclusion of the X variable.

The results for OLS with country fixed effects are shown in the upper part of Table A2. The conclusions with respect to $\hat{\beta}$ depend on the time-series length, T, and the coefficient α on the X variable. For example, if T=100, the mean of $\hat{\beta}$ is 0.032 (line 1) when α=0.05, and the mean is 0.021 (line 2) when α=0.25. Hence, the bias is small with a 100-year sample when a lot of the variation in the y variable reflects the observed X variable. However, if T=20, the mean of $\hat{\beta}$ is 0.103 (line 3) when α=0.05 and 0.029 (line 4) when α=0.25. That is, the Hurwicz-Nickell channel can lead to a strong upward bias when T=20.

When the X variable is excluded, the mean of $\hat{\beta}$ is 0.023 (line 5) when α=0.05 and T=100. That is, the bias can be small with a long time series if α is not too large. However, the mean of $\hat{\beta}$ is 0.007 (line 6) when α=0.25. That is, the bias is substantially downward because of the important omitted-variables effect with this higher value of α. When T=20, the bias is upward for each value of α because of the Hurwicz-Nickell effect. The mean of $\hat{\beta}$ is 0.106 (line 7) when α=0.05 and 0.033 (line 8) when α=0.25.

The results suggest circumstances when the regressions generate estimates of the convergence rate, β, without large bias. The first is for OLS with country fixed effects when the sample is large in the time domain. For example, when β=0.02, the fixed-effects model yields a mean of $\hat{\beta}$ of 0.036 when T=140 (Table A1, line 5). For the model that adds an exogenous X variable, the mean of $\hat{\beta}$ is 0.021 when T=100 and α=0.25 (Table A2, line 2). Even with the X variable excluded, the mean of $\hat{\beta}$ is 0.023 when T=100 and α=0.05 (line 5). The second case is for OLS without country fixed effects in the X-model when the (exogenous) X variable is included in the regression. When β=0.02, this framework yields a mean for $\hat{\beta}$ between 0.016 and 0.020 (Table A2, lines 9-12). OLS without country fixed effects also does well in the fixed-effects model when the variation in the fixed effects is small (Table A1, row 14).

The results can be used to get reasonable bounds on the convergence rate, β, as carried out in the text. The upper bound came from the estimate with country fixed effects without X variables, based on the long-term data back to 1870. Because T is large, the upper bound may be reasonably tight (as in Table A1, line 5, or Table A2, line 5). The lower bound came from the estimate without country fixed effects using the shorter-term data since 1960 with a rich array of X variables (as in Table A2, lines 9-12).

Table A1
Monte Carlo Results for Estimated Convergence Rates
Model with Country Fixed Effects

	(1)	(2)	(3)	(4)	(5)
OLS with country fixed effects					
	β	**T**	**σ_η**	$E(\widehat{\beta})$ (Nickell formula)	$mean(\widehat{\beta})$ (Monte Carlo)
(1)	0.02	20	1	0.165	0.151 (0.014)
(2)	0.02	50	1	0.076	0.070 (0.006)
(3)	0.02	100	1	0.046	0.043 (0.004)
(4)	0.02	100	0.1	0.046	0.043 (0.004)
(5)	0.02	140	1	0.038	0.036 (0.003)
(6)	0.10	20	1	0.231	0.206 (0.014)
(7)	0.10	50	1	0.147	0.140 (0.008)
(8)	0.10	100	1	0.122	0.120 (0.005)
(9)	0.10	100	0.1	0.122	0.120 (0.004)
(10)	0.10	140	1	0.115	0.113 (0.004)
OLS without country fixed effects					
(11)	0.02	100	1	--	0.00026 (0.00015)
(12)	0.02	100	0.1	--	0.0111 (0.0015)
(13)	0.02	20	0.1	--	0.0126 (0.0029)
(14)	0.02	100	0.01	--	0.0200 (0.0021)
(15)	0.10	100	1	--	0.0053 (0.0009)
(16)	0.10	100	0.1	--	0.0858 (0.0045)
(17)	0.10	20	0.1	--	0.0884 (0.0092)
(18)	0.10	100	0.01	--	0.1006 (0.0036)

Notes: $\beta=1-\gamma$ is the convergence rate in equation (A1), T is the length of the time series, and σ_η is the standard deviation of the country fixed effect. The Monte Carlo results in column 5 show the mean and standard deviation of the estimate $\hat{\beta}$ from 100 iterations of the model in equation (A1). The upper part refers to OLS estimation with country fixed effects. The lower part refers to OLS estimation without country fixed effects. All results shown use $\sigma_\varepsilon=1$ (for the time-series shock), N (number of countries) =100, and an observation period of $\tau=1$ year. Results are similar with 50 iterations, N=20 are similar, or $\tau=5$. For example, with country fixed effects, the mean of $\hat{\beta}$ when $\beta=0.10$, T=100, $\sigma_\eta=1$, and $\tau=5$ is 0.125 (0.006). Without country fixed effects, the mean of $\hat{\beta}$ for this case is 0.0047 (0.0008). Column 4 shows the expected value of $\hat{\beta}$ from the formula from Nickell (1981, p. 1422) in equation (1). This formula applies to OLS estimation with country fixed effects.

Table A2 Monte Carlo Results for Estimated Convergence Rates Model with Country Fixed Effects and X Variable						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(8)**
	β	**α**	**T**	**incl. X?**	$\boldsymbol{mean(\widehat{\beta})}$ **(Monte Carlo)**	$\boldsymbol{mean(\hat{\alpha})}$ **(Monte Carlo)**
OLS with fixed effects						
(1)	0.02	0.05	100	y	0.0315 (0.0025)	0.0534 (0.0034)
(2)	0.02	0.25	100	y	0.0208 (0.0005)	0.2516 (0.0032)
(3)	0.02	0.05	20	y	0.1029 (0.0120)	0.0450 (0.0144)
(4)	0.02	0.25	20	y	0.0288 (0.0034)	0.2465 (0.0135)
(5)	0.02	0.05	100	n	0.0234 (0.0025)	--
(6)	0.02	0.25	100	n	0.0069 (0.0024)	--
(7)	0.02	0.05	20	n	0.1056 (0.0131)	--
(8)	0.02	0.25	20	n	0.0326 (0.0054)	--
OLS without fixed effects						
(9)	0.02	0.05	100	y	0.0156 (0.0013)	0.0446 (0.0026)
(10)	0.02	0.25	100	y	0.0196 (0.0004)	0.2472 (0.0031)
(11)	0.02	0.05	20	y	0.0166 (0.0024)	0.0470 (0.0040)
(12)	0.02	0.25	20	y	0.0198 (0.0024)	0.2489 (0.0049)
(13)	0.02	0.05	100	n	0.0041 (0.0010)	--
(14)	0.02	0.25	100	n	0.0004 (0.0009)	--
(15)	0.02	0.05	20	n	0.0041 (0.0024)	--
(16)	0.02	0.25	20	n	-0.0024 (0.0023)	--

Notes: For the model in equations (A2)-(A6), $\beta=1-\gamma$ is the convergence rate, α is the coefficient on the X variable, ρ is the persistence coefficient for the X variable, and T is the length of the time series. The Monte Carlo results in columns 5 and 6 show the mean and standard deviation of the estimates $\hat{\beta}$ and $\hat{\alpha}$ from 100 iterations of the model. The upper part refers to OLS estimation with country fixed effects. The lower part refers to OLS estimation without country fixed effects. All results normalize to have $\sigma_\varepsilon=1$ (for the time-series shock) and $\sigma_u=1$ (for the X shock). Other assumptions are $\sigma_\eta=0.1$ (for the country fixed effect), ρ (coefficient for the persistence of the X process) = 0.98, N (number of countries) = 100, and use an observation period of $\tau=1$ year. Column 4 indicates whether the regressions include the X variable.