

# Conditional Convergence Revisited: Taking Solow Very Seriously

Kieran McQuinn\* and Karl Whelan†  
Central Bank and Financial Services Authority of Ireland

March 2006

## Abstract

Output per worker can be expressed as a function of technological efficiency and of the capital-output ratio. Because technology is exogenous in the Solow model, all of the endogenous convergence dynamics take place through the adjustment of the capital-output ratio. This paper uses the empirical behavior of the capital-output ratio to estimate the speed of conditional convergence of economies towards their steady-state paths. We find that the conditional convergence speed is about seven percent per year. This is somewhat faster than predicted by the Solow model and is significantly higher than reported in most previous studies based on output per worker regressions. We show that, once there are stochastic shocks to technology, standard panel econometric techniques produce downward-biased estimates of convergence speeds, while our approach does not.

---

\*E-mail: [kmcquinn@centralbank.ie](mailto:kmcquinn@centralbank.ie)

†E-mail: [karl.whelan@centralbank.ie](mailto:karl.whelan@centralbank.ie). The views expressed in this paper are our own, and do not necessarily reflect the views of the Central Bank and Financial Services Authority of Ireland or the ESCB.

# 1 Introduction

The Solow (1956) model has provided the theoretical framework through which most cross-country studies of empirical growth have interpreted their findings. The model predicts that economies tend to converge towards a steady-state path, and that the growth rate of this path is determined by technological progress while its level can be shifted up and down by variables such as the savings rate and population growth rate. Mankiw, Romer, and Weil (1992) and Barro and Sala-i-Martin (1992) reported that this “conditional convergence” speed appeared to be significantly lower than Solow’s model predicts, based on simple cross-sectional regressions. This finding has had a considerable effect on thinking about growth and development. In particular, many have followed Mankiw, Romer, and Weil’s suggestion that the Solow model needs to be augmented with accumulable human capital, a modification that implies slower conditional convergence, and also much larger effects on output levels for policies that boost physical and human capital accumulation.

Since these early papers, there has been a large empirical literature on conditional convergence. Much of this research has focused on the fact that the cross-sectional regressions in the original convergence papers can produce downward-biased estimates of convergence speeds due to a failure to account for country-specific variables that do not change over time (i.e. fixed effects).<sup>1</sup> However, there is still little agreement on how best to deal with country-specific fixed effects in the context of dynamic panel regressions, and many of the commonly-used panel estimators have been shown to produce upward-biased estimates of convergence speeds.<sup>2</sup>

This paper presents a new approach to implementing the Solow model and estimating the speed of conditional convergence. Our approach has important methodological and substantive implications. In terms of methodology, we show that it is possible to consistently estimate the conditional convergence speed predicted by the Solow model without having to rely on techniques to deal with country-specific fixed effects. Thus, our approach gets around the principal econometric problems that have been associated with the cross-country growth literature, and as such, we believe it produces more reliable and credible estimates of convergence speeds than previously reported.

In terms of substance, our approach leads to a very different assessment of the speed of

---

<sup>1</sup>See, for instance, Islam (1995) and Caselli, Esquival, and Lefort (1996).

<sup>2</sup>Bond, Hoeffler, and Temple (2001) discuss the upward biases associated with panel data techniques such as the standard “within groups” or fixed effects estimator.

conditional convergence, and of the adequacy of the basic Solow model, than that suggested by most previous research. Our results point to a conditional convergence speed of about seven percent per year, well above the two-percent figure often cited as a stylized fact. We argue that, if anything, the basic Solow model errs in somewhat *underpredicting* rather than overpredicting the speed of convergence.

The basis for our approach is a property of the Solow model that is relatively well known but that has not been exploited previously in the empirical literature on conditional convergence. Output per worker can be expressed as a function of the level of technological efficiency and of the capital-output ratio: This decomposition has been used, for instance, by Hall and Jones (1997). Because technology is strictly exogenous in the Solow model, all of the endogenous convergence dynamics take place through the adjustment of the capital-output ratio towards its target or steady-state level.<sup>3</sup> In our empirical analysis, we use a cross-country data set to estimate the rate at which capital-output ratios tend to converge towards their steady-state values, which also tells us about the conditional convergence speed for output per worker.

We believe our approach gives different results from previous studies because it takes the Solow model seriously in ways that previous work has not.<sup>4</sup> We focus on the endogenous convergence dynamics predicted by the model, while other studies have only examined these dynamics indirectly. Previous research has focused on regressions for output per worker. We argue that these regressions should not be interpreted as estimating the Solow model per se, but rather a *joint* model that combines Solow with a specification of the process generating technological efficiency across countries. Indeed, because of the need to make assumptions about technology, it can be argued that this approach has run somewhat counter to the spirit of the Solow model, which treats technology as exogenous and, thus, makes no predictions about it.

In addition, to the extent that their underlying specification for technology may be incorrect, output per worker regressions can produce misleading results. We show that the standard specification of the technology process in previous studies is a highly inaccurate one. This approach—employed by essentially all of the panel data studies—assumes that

---

<sup>3</sup>See, for instance, Chapter 4 of Brad DeLong’s recent *Macroeconomics* textbook for a discussion of this property.

<sup>4</sup>We are aware, of course, of Mankiw, Romer, and Weil’s well-known opening statement that “This paper takes Robert Solow seriously.” We argue, however, that our approach remains truer to the spirit of Solow’s model than does the approach in the that paper.

technological efficiency grows at the same rate across all countries, implicitly assigning all differences in the level of technology to a country-specific fixed effect. This assumption, however, is clearly rejected by even a cursory examination of the evidence on total factor productivity: TFP growth tends to vary widely across countries at any point in time.

We use Monte Carlo simulations to illustrate the consequences for output per worker regressions of failure to model the technology process accurately. We document a new and important source of bias in these regressions: The failure to account for country-specific stochastic shocks to TFP growth is a significant source of downward bias for estimates of the speed of conditional convergence. Thus, estimates based on the standard within-groups regression technique are subject to both this downward bias, as well as the finite-sample upward bias documented by Nickell (1981). We show that for samples of the size used in our empirical work, the downward bias due to stochastic technology shocks dominates. Indeed, the empirical results obtained from within-group regressions for output per worker turn out to be roughly consistent with the faster convergence speeds suggested by our method. These results are of importance because it has previously been thought that within-groups estimates represent upper bounds on the likely convergence speed.

In contrast to the output per worker regressions, our approach does not require a specification of the technology process for each country. For this reason, there is no theoretical case for the inclusion of fixed-effects in our regression specification. Indeed, while we do report such estimates, hypothesis tests reject the presence of fixed effects, allowing the model to be estimated using pooled OLS. In light of the ongoing debates about the efficiency of various panel estimation techniques for dealing with country-specific effects, we believe our results provide a simple and intuitive alternative characterization of the conditional convergence process.

The plan for the paper is as follows. We start by reviewing the theoretical results concerning convergence dynamics in the Solow model and presenting our empirical estimates of conditional convergence. We then discuss the relationship between our econometric approach and the traditional approach based on output per worker regressions. We outline how the traditional regressions are subject to a number of important biases that do not affect our approach and how the results from these regressions are generally consistent with our conclusions. Finally, we discuss some of the implications of our results for growth and development economics.

## 2 Convergence Dynamics in the Solow Model

### 2.1 Theory

Our analysis is based on the standard Solow model assumptions, starting with a Cobb-Douglas production function

$$Y_t = K_t^\alpha (A_t L_t)^{1-\alpha} \quad (1)$$

where  $0 < \alpha < 1$  which implies diminishing marginal productivity of capital. Assuming continuous time, the capital stock accumulates according to

$$\dot{K}_t = sY_t - \delta K_t \quad (2)$$

where  $s$  is the investment share of output.

In our analysis, we will work with a reformulated version of the production function, derived as follows. Defining the capital-output ratio as

$$X_t = \frac{K_t}{Y_t} \quad (3)$$

output per worker can be expressed as

$$\frac{Y_t}{L_t} = A_t X_t^{\frac{\alpha}{1-\alpha}} \quad (4)$$

This decomposition has been used in a number of previous studies, most notably by Hall and Jones (1997). Relative to the more familiar decomposition of output per worker into technology and capital-per-worker terms, this decomposition has an important advantage. The long-run capital-output ratio can be shown to be independent of the level of  $A_t$ , something which is not true of capital-per-worker. Hence, this formulation completely captures the effects of  $A_t$  on long-run output, while the more traditional decomposition features a capital deepening term that depends indirectly on the level of technology.

**Capital-Output Ratio Dynamics:** The dynamics of the capital-output ratio are derived as follows. The growth rate of this ratio is

$$\frac{\dot{X}_t}{X_t} = \frac{\dot{K}_t}{K_t} - \frac{\dot{Y}_t}{Y_t} = (1 - \alpha) \left( \frac{\dot{K}_t}{K_t} - \frac{\dot{A}_t}{A_t} - \frac{\dot{L}_t}{L_t} \right) \quad (5)$$

Denoting the growth rates of technology and workers as

$$\frac{\dot{A}_t}{A_t} = g \quad (6)$$

$$\frac{\dot{L}_t}{L_t} = n \quad (7)$$

and using equation (2), the dynamics of the capital-output ratio become

$$\frac{\dot{X}_t}{X_t} = (1 - \alpha) \left( \frac{s}{X_t} - g - n - \delta \right). \quad (8)$$

These dynamics imply that the ratio converges to a steady-state level defined by

$$X^* = \frac{s}{g + n + \delta}. \quad (9)$$

With this definition in hand, the ratio's dynamics can be re-written in "error-correction" form as

$$\dot{X}_t = \lambda (X^* - X_t), \quad (10)$$

where

$$\lambda = (1 - \alpha)(g + n + \delta). \quad (11)$$

As noted by Jones (2000), this first-order differential equation has a solution given by

$$X_t = e^{-\lambda t} X_0 + (1 - e^{-\lambda t}) X^*. \quad (12)$$

For our analysis, it is useful to note that this result also holds approximately for the *log* of the capital-output ratio. In other words, letting  $x_t$  be the log of this ratio, then

$$\dot{x}_t = \frac{\dot{X}_t}{X_t} = \lambda \left( \frac{X^* - X_t}{X_t} \right) \approx \lambda (x^* - x_t) \quad (13)$$

which implies a solution of the form

$$x_t = e^{-\lambda t} x_0 + (1 - e^{-\lambda t}) x^*. \quad (14)$$

**Output-Per-Worker Ratio Dynamics:** These results also allow for a simple characterization of the dynamics of output per worker. Again letting lower-case letters represent logged variables, we have

$$y_t = a_t + \frac{\alpha}{1 - \alpha} x_t. \quad (15)$$

So the dynamics of output per worker are given by

$$\dot{y}_t = g + \frac{\alpha}{1 - \alpha} \dot{x}_t. \quad (16)$$

The steady-state path for output per worker is the level of output per worker consistent with the capital-output ratio being equal to its steady-state level:

$$y_t^* = a_t + \frac{\alpha}{1 - \alpha} x^*. \quad (17)$$

Using equation (13), output per worker dynamics can then be expressed as

$$\dot{y}_t = g + \lambda(y_t^* - y_t). \quad (18)$$

Thus, the convergence speed,  $\lambda$ , of the capital-output ratio is also the so-called conditional convergence speed of output-per-worker, i.e. it is the speed at which output per worker adjusts towards its steady-state level. Note, however, that the output growth equation has two components to it: Growth is determined by technological progress as well as the gap between  $y_t$  and  $y_t^*$ . In contrast, movements in the capital-output ratio are determined only by the gap between output and its steady-state level.

These results provide the basis for the empirical strategy followed in this paper, which is to use the convergence properties of the capital-output ratio to directly estimate the speed of conditional convergence. As we will discuss at greater length later, the fact that capital-output dynamics depend only on the gap between output and its steady-state level considerably simplifies the empirical estimation of convergence speeds relative to the estimation of output per worker regressions, where some specification of the nature of technological progress is necessary.

## 2.2 Generality of Results

Our derivations have been based upon a Cobb-Douglas production function, so a few points about the generality of these results are worth noting. The first relates to the generality of the representation of output per worker as a function of technology and of the capital-output ratio. For the economy to exhibit steady-state growth, the production function must be of the form  $F(K, AL)$ , implying labor-augmenting technological change.<sup>5</sup> If, in addition, there are constant returns to scale, then one can write

$$F\left(\frac{K}{Y}, A\frac{L}{Y}\right) = 1 \quad (19)$$

---

<sup>5</sup>See Jones and Scrimgeour (2005) for a discussion of this result.

This restriction implies the existence of an implicit function  $\frac{Y}{L} = G(A, \frac{K}{Y})$ , relating output per worker to technology and to the capital-output ratio.

Second, the formula for the steady-state capital-output ratio, equation (9), also holds for any model with a production function  $F(K, AL)$  and featuring diminishing marginal productivity of capital and labor. This latter assumption implies that the growth rates of  $K$  and  $AL$  must be the same along the steady-state growth path. Because these growth rates are given by  $s\frac{Y}{K} - \delta$  and  $g + n$  respectively, the formula for the steady-state capital-output ratio automatically follows.

Finally, in relation to the convergence dynamics derived here, it should be kept in mind that there is no general analytical solution for convergence dynamics in the Solow model. For this reason, most discussions of its convergence dynamics have proceeded via first-order log-linearizations of the production function.<sup>6</sup> In this sense, the formulas derived here based on the Cobb-Douglas production function can be viewed as a first-order approximation to the dynamics implied by more general production functions.

### 3 Empirical Results

In our empirical implementation, we estimate conditional convergence speeds by examining the rate at which capital-output ratios in a large panel of countries tend to converge to the steady-state levels predicted by the Solow model. Here, we describe the estimating equations used, the details behind the construction of the data, and our empirical results.

#### 3.1 Estimating Equation

Equation (14), which has been derived from a continuous time model, has strong predictions for the behavior of the capital-output ratio that one should observe when data are sampled at discrete intervals. If the data are sampled every  $r$  periods, then the  $r$ -period change in the capital-output ratio should be given by

$$x_t - x_{t-r} = \left(1 - e^{\lambda r}\right) (x^* - x_{t-r}) \quad (20)$$

If the determinants of the steady-state capital-output ratio were constant over time then this equation could be directly estimated using any time series for the capital-output

---

<sup>6</sup>For instance, the formula for the convergence speed  $\lambda$  (equation 11) is derived in this manner in the textbook treatments of Barro and Sala-i-Martin (1995) and Romer (2001).

ratio. In reality, however, investment rates and population growth rates do change over time. Thus, our empirical approach allows the “target” capital-output ratio to vary over time, implying an estimating equation of the form

$$x_{it} - x_{i,t-r} = \left(1 - e^{-\lambda r}\right) (x_{it}^* - x_{i,t-r}), \quad (21)$$

where  $x_{it}^*$  is defined as

$$x_{it}^* = \log(s_{it}) - \log(g_i + n_{it} + \delta). \quad (22)$$

We report results from applying this regression to data based on both one-year and five-year intervals, and with changes in the investment rate and population growth rate allowed to affect the target capital-output ratio.<sup>7</sup>

### 3.2 Data Construction

Our data come from version 6.1 of the Penn World Tables, as documented by Heston, Summers, and Aten (2002). We use data for the ninety-six countries listed in Appendix A over the period 1950-2000. Implementation of our approach required construction of time series for both the actual and target capital stocks. Of course, capital stocks are not “primary” data, but rather must be constructed from assumptions regarding depreciation rates and initial conditions. Here, we describe our approach to constructing these series.

**Depreciation Assumptions:** A number of the early papers on conditional convergence, such as Mankiw, Romer, and Weil (1992), assumed a depreciation rate of three percent per year. Mankiw (1995) explains that this is approximately the figure obtained from the US national accounts when the value of depreciation was divided by the value of the capital stock. However, the Department of Commerce has significantly revised its capital stock estimates since the mid-1990s, with its new estimates on updated empirical evidence on depreciation for various types of assets. With these revisions, this same calculation now produces a figure of about four and a half percent.<sup>8</sup>

In fact, we believe that the most appropriate depreciation rate for the application of the Solow model is actually somewhat higher again. Various types of capital depreciate at very

---

<sup>7</sup>One way to think of this approach is that it implies that equation (14) holds each period, with changes in  $x^*$  implying jumps in the “initial conditions” element of the solution.

<sup>8</sup>See Fraumeni (1997) for a discussion of the Commerce Department’s methodology for constructing capital stocks. The data for these calculations were downloaded from the BEA’s website at [www.bea.doc.gov/bea/dn/home/fixedassets.htm](http://www.bea.doc.gov/bea/dn/home/fixedassets.htm)

different rates: Structures tend to depreciate at rates below two percent per year, while equipment tends to depreciate at rates above ten percent. Mankiw’s calculation implicitly weights these depreciation rates according to the weight of each type of capital in the *current value* of the stock. Consider, however, a re-formulated version of the Solow model with multiple types of capital, so that the production function is

$$Y_t = S_t^\gamma E_t^{\alpha-\gamma} (A_t L_t)^{1-\alpha} \quad (23)$$

where  $S_t$  is structures and  $E_t$  is equipment. In Appendix B we show that our estimating equations can only be obtained in this case if the weights used to calculate the “aggregate depreciation rate” reflect the contribution of each asset to production. In other words, the correct weights in this example would be  $\frac{\gamma}{\alpha}$  and  $\frac{\alpha-\gamma}{\alpha}$ . Empirical calculations of this form usually point to approximately equal weights for equipment and structures in the production function, whereas value weights point to structures being far more important.<sup>9</sup> An equally-weighted average of a two percent structures depreciation rate and a ten percent equipment depreciation rate points to an overall depreciation rate of six percent.

For this reason, six percent is our preferred depreciation rate. However, our principal conclusion—that the estimated convergence speed is at least as fast as predicted by the Solow model—is not affected by this choice of parameter. Thus, we also report results for depreciation rates of four and five percent.

**Initial Conditions Assumption:** Given an assumed depreciation rate, we can construct time series for capital stocks once we have an initial value for each stock. While initial capital stocks cannot be observed, one can make an informed guess based on the observation that the ratio of investment to capital is given by

$$\frac{I_t}{K_{t-1}} = \frac{\Delta K_t}{K_{t-1}} + \delta \quad (24)$$

and thus,

$$K_{t-1} = \frac{I_t}{\frac{\Delta K_t}{K_{t-1}} + \delta} \quad (25)$$

Our approach has been to construct an initial capital stock for 1960 for each country in our sample according to this formula, using the average growth rate of investment over

---

<sup>9</sup>For instance, Greenwood, Hercowitz, and Krusell (1997) report Cobb-Douglas exponents of 0.17 for equipment and 0.13 for structures. Similar calculations reported by Whelan (2003) show 0.145 for equipment and 0.165 for structures.

the previous decade as our proxy for the growth rate of the capital stock. This initial assumption appears to be quite accurate. For instance, when we apply this method to construct a proxy for the year 2000 stock, the resulting estimates have a correlation of 0.99 with the figures based on the 1960 initial conditions assumption, even though this latter series is almost completely based on data rather than initial assumptions.<sup>10</sup>

**Construction of  $x_{it}^*$ :** As noted above, we allow the the estimated target capital-output ratio to vary over time with changes in each country’s investment rate and population growth rate. In this sense, we are following the approach in previous panel studies such as Islam (1995) and Caselli, Esquivel, and Lefort (1996), which included these variables due to their effects in changing the steady-state level of output per worker. Unlike those studies, however, which assumed that  $g + \delta$  was constant across all countries, we also allow  $g_i$  to correspond to our estimate of each country’s average growth rate of technology. These estimates were based on the average growth rate of time series for  $A_{it}$  obtained from a levels accounting exercise based on the assumption of a Cobb-Douglas production function, using our series on capital, measuring labor input as the number of workers, and the standard capital share value of  $\alpha = \frac{1}{3}$ .<sup>11</sup>

### 3.3 Results

Table 1 reports results from estimation of equation (21) over the period 1960-2000, with the interval  $r$  set to one year, and for data based on depreciation rates of four, five, and six percent.<sup>12</sup> The first column reports the estimated convergence speed parameter  $\lambda$  from pooled OLS estimation of the equation without any intercept term or country-specific effects (since these terms are not predicted by the theory). The second column reports results from the within-groups (i.e. least-squares dummy variable) estimator which allows for country-specific fixed effects. The final column reports the average convergence rate consistent with the Solow model for each depreciation rate. In other words, it reports

$$\lambda^{Solow} = (1 - \alpha)(\bar{g}_i + \bar{n}_i + \delta) \tag{26}$$

---

<sup>10</sup>For instance, for our preferred depreciation rate of six percent, the starting 1960 value of the capital stock receives a weight of  $(1 - 0.06)^{40} = 0.084$  in the 2000 stock.

<sup>11</sup>Our results reported here are barely changed by the replacement of the one-third assumption with the capital share values reported by Bernanke and Gurkaynak (2002), where such values were available.

<sup>12</sup>As noted in Section 2, these regressions could also be run for the *level* of the capital-output ratio, as opposed to the logged value. This approach produces essentially the same results as those reported here.

where  $\bar{g}_i$  and  $\bar{n}_i$  are the average values across countries for the growth rates of technology and population.

The principal result from Table 1 (indeed, the principal result of the paper) is that for each of the depreciation rates reported, the estimated convergence speeds are somewhat faster than those predicted by the Solow model. For instance, for our preferred depreciation rate of six percent, both estimation techniques point to a convergence speed of about seven percent, compared with a prediction of  $\lambda = 0.063$  from the model. As expected, the estimated convergence speed rises with the assumed depreciation rate. However, even for the low depreciation rate of four percent, the estimated convergence speeds, at about six percent, are significantly higher than those reported in most conditional convergence studies.

Table 2 repeats these regressions using a five-year interval. Starting with Islam (1995), this has been the most common interval used in panel studies of conditional convergence. The use of five-year intervals is seen as reducing the impact of cyclical fluctuations on the estimated long-run convergence coefficients, as well as smoothing away much of the measurement error that may be associated with the annual data. For these regressions, the target capital-output ratio  $x_{it}^*$  is based on the investment and population growth rates for the five-year interval including period  $t$ . These regressions produce very similar results to those based on the one-year regressions, with conditional convergence speeds still somewhat faster than predicted by the Solow model.

For each of our regression specifications, the within-groups panel estimator suggests a slightly higher convergence speed than the pooled OLS specification. However, our preferred estimates are those based on the pooled OLS estimation, because hypothesis tests reject the presence of country-specific fixed effects. This result is perhaps a little surprising because it implies that we have not omitted any important country-specific determinants of the long-run capital-output ratio. Indeed, it turns out that failure to deal with country-specific factors can easily overturn this rejection of fixed effects. For instance, when defining  $x_{it}$ , if we replace our country-specific estimates of the growth rate of technology,  $g_i$ , with the world average growth rate of 1.4 percent, then the hypothesis of no fixed effects is overwhelmingly rejected. That said, these regressions also implied convergence speeds close to those reported in Tables 1 and 2.

### 3.4 Robustness

Our key finding, that conditional convergence speeds are somewhat higher than predicted by the Solow model, turns out to be quite robust to changes in specification and samples.

Tables 3 and 4 report results for the case  $\delta = 0.06$  from specifications that are more general than our basic regression. Table 3 reports results from a specification that allows for a more complicated pattern of dynamic adjustment by including lagged changes in the capital-output ratio; this additional term is significant in the one-year specification but not in the five-year specification. Again, the assumption of fixed effects is rejected, and the estimated convergence speeds of about seven percent are little changed relative to the original estimates.

Table 4 loosens the implicit theoretical restriction imposing coefficients of equal magnitude on the lagged capital-output ratio,  $x_{i,t-1}$ , as well as on the two components of the target ratio,  $\log s_{it}$  and  $\log (g_i + n_{it} + \delta)$ . Most of the previous studies of conditional convergence, based on output per worker regressions, have not imposed this latter restriction that the investment and population variables have coefficients of equal magnitude. The presence of fixed effects is again rejected in these regressions, leaving the pooled OLS regression as the appropriate technique. The magnitude of the coefficients from these regressions turn out to be very close to each other, exactly as predicted by the model, and the estimated convergence speeds implied by the coefficients on the lagged capital-output ratios are again very close to seven percent.

Table 5 reports results from estimating our base regression specifications for smaller samples of countries. Following Mankiw, Romer, and Weil (1992) we considered two alternative samples. The first is an “intermediate” sample of 80 countries based on the exclusion of countries that received a grade D for data quality, as well as countries with populations of below one million in 1960. The second is a sample of 23 OECD countries. Again, the assumption of fixed effects can be rejected in each case. The intermediate sample results point to a seven percent convergence speed for both one-year and five-year intervals. The one-year regression for the OECD sample gives a convergence speed of five percent, the only deviation from the pattern of results reported elsewhere, but the five-year regressions, which are probably more reliable, again report a convergence speed of about seven percent.

For our final robustness check, we note that we have followed in the tradition of previous papers on conditional convergence in reporting a single convergence speed based on information across a large sample of countries. However, it is likely that convergence speeds may

vary across countries. Given that we have forty years of data for each of these countries, it is also feasible to estimate our one-year equation separately for each country. Summary statistics from these exercises are reported in Table 6. Though the average convergence speeds are a touch lower than those reported for the equivalent pooled regressions in Table 1, the overall pattern is in line with our previous conclusions. In each case, the average convergence speed either equals or is just above that predicted by the model. More generally, even those countries with estimated convergence speeds at the lower end of the distribution have convergence speeds faster than the commonly-cited “stylized fact” figure of two percent.

## 4 Relationship to Output Per Worker Regressions

In reporting convergence speeds that are consistently equal to or above those predicted by the Solow model, our approach has produced a very different picture of the process of conditional convergence than is suggested by conventional wisdom. For this reason, we think it is important to reconcile our estimates with those produced from previous studies. To do so, it is necessary to understand the relationship between our approach (based on capital-output regressions) and those from previous studies (based on output per worker regressions).

In the next few sections, we outline the relationship between the two approaches and explain why our approach is more likely to provide accurate estimates of the speed of conditional convergence. First, we discuss how the standard panel approach to output per worker regressions relies on a very specific formulation of the cross-country process for technology. We outline how, even if this assumption is correct, our method will have considerable advantages. In the next section, we show that the standard assumption about technology is highly inaccurate and that this leads to a new (previously undocumented) source of bias in panel data growth regressions.

### 4.1 Relationship Between the Two Approaches

We have documented how the convergence properties of the capital-output ratio provide the basis for the conditional convergence predictions of the Solow model. However, our approach of directly estimating this convergence speed has not been used in previous empirical studies. Instead, these studies have focused on the behavior of output per worker. Here we consider

the model's predictions for regressions based on output per worker. To understand this relationship, the first step is to re-write the dynamic capital-output equation (21) in levels form as

$$x_{it} = \left(1 - e^{-\lambda r}\right) x_{it}^* + e^{-\lambda r} x_{i,t-r}. \quad (27)$$

This can be turned into an equation for output per worker dynamics by using the fact that

$$x_{it} = \frac{1 - \alpha}{\alpha} (y_{it} - a_{it}), \quad (28)$$

so that the capital-output equation becomes

$$\frac{1 - \alpha}{\alpha} (y_{it} - a_{it}) = \left(1 - e^{-\lambda r}\right) x_{it}^* + e^{-\lambda r} \left(\frac{1 - \alpha}{\alpha}\right) (y_{i,t-r} - a_{i,t-r}) \quad (29)$$

This can then be re-arranged in more compact form to give

$$y_{it} = a_{it} - e^{-\lambda r} a_{i,t-r} + \frac{\alpha (1 - e^{-\lambda r})}{1 - \alpha} x_{it}^* + e^{-\lambda r} y_{i,t-r} \quad (30)$$

Re-expressing the steady-state capital-output ratio in terms of its determinants then gives

$$y_{it} = a_{it} - e^{-\lambda r} a_{i,t-r} + \frac{\alpha (1 - e^{-\lambda r})}{1 - \alpha} [\log(s_{it}) - \log(g_i + n_{it} + \delta)] + e^{-\lambda r} y_{i,t-r} \quad (31)$$

This equation describes the conditional convergence predictions of the Solow model as it relates to output per worker. Controlling for the determinants of the steady-state capital-output ratio and also for the evolution of technological efficiency, there is a negative relationship between the initial level of output per worker and the subsequent growth, with the conditional convergence speed being dictated by the dynamics of the capital-output ratio.

To better explain the relationship between our empirical work and those in previous studies, it is worth noting that previous papers have not examined convergence dynamics in terms of equation (31) because they do not include measures of current and lagged technological efficiency. Instead, they have made simplifying assumptions about the form of the technology process and these simplifications have implied a specific estimable regression format. For the vast majority of research in this area, the assumption has taken the form

$$a_{it} = a_{i0} + \sum_{m=1}^t g_m. \quad (32)$$

In other words, it is assumed that all differences across countries in technological efficiency are accounted for by a once-off fixed effect due to "initial conditions". After this initial

period, all countries experience the same rate of technological progress.<sup>13</sup> Mankiw, Romer, and Weil (1992) initially justified this assumption on the grounds that technological progress “reflects primarily the advancement of knowledge, which is not country-specific.”

With this assumption in hand, the technology term in the output per worker relationship can be written as

$$a_{it} - e^{-\lambda r} a_{i,t-r} = (1 - e^{-\lambda r}) a_{i0} + \sum_{m=1}^t g_m - e^{-\lambda r} \sum_{m=1}^{t-r} g_m \quad (33)$$

and the expression for the growth in output per worker can be written in a panel-data regression format as

$$y_{it} = \gamma_i + \omega_t + \frac{\alpha(1 - e^{-\lambda r})}{1 - \alpha} [\log(s_{i,t-r}) - \log(g + n_{i,t-r} + \delta)] + e^{-\lambda r} y_{i,t-r} + v_{it} \quad (34)$$

where

$$\gamma_i = (1 - e^{-\lambda r}) a_{i0} \quad (35)$$

$$\omega_t = \sum_{m=1}^t g_m - e^{-\lambda r} \sum_{m=1}^{t-r} g_m \quad (36)$$

and  $v_{it}$  is an error term.

## 4.2 Panel Data Estimation Biases: A Review

One of the potential problems with estimating the speed of conditional convergence using equation (34) is that the identifying assumption regarding technology—equation (32)—may be incorrect, and we will discuss this problem in the next few sections. However, even if the technology assumption is correct, there are a number of serious econometric difficulties associated with this type of regression.

The first and best-known difficulty stems from the presence of country-specific fixed effects (the  $\gamma_i$  terms).<sup>14</sup> These effects imply serious complications for each of the three most common estimators that have been applied to estimate cross-country growth models:

---

<sup>13</sup>Some studies describe their assumption about technology as being  $a_{it} = a_{i0} + gt$ , which assumes a constant rate of world technological progress, i.e.  $g_t = g$  at all times. Since these same studies usually use time-effects rather than time trends to capture technological progress, the actual assumption is significantly more general.

<sup>14</sup>Bond (2002) provides a useful detailed discussion of these econometric problems.

- Pure cross-section regressions that ignore the panel element of the data (such as Mankiw, Romer, and Weil's) tend to substantially understate the speed of convergence. They relegate the fixed effect to the error term, and this effect is necessarily positively correlated with the lagged output term,  $y_{i,t-r}$ . The upward bias in this coefficient implies a downward bias in the implied convergence speed. If fixed effects are present, then the same argument also applies to pooled-OLS estimators that include the panel element of the data but ignore the fixed effect.
- Panel techniques such as the within-groups estimator (as applied by Islam, 1995, and many others) acknowledge the existence of the fixed effect but do not eliminate the biases associated with them. Effectively, this technique transforms the model by subtracting off country-specific means. In this case, assuming  $r = 1$ , the transformed lagged output term is  $y_{i,t-1} - \sum_{m=1}^T y_{i,m-1}$  and the transformed error term is of the form  $v_{it} - \sum_{m=1}^T v_{im}$ . These terms are negatively correlated, implying an upward-biased estimate of the speed of conditional convergence, a bias that has been documented analytically by Stephen Nickell (1981).
- First-differencing the model eliminates the fixed effect from the specification but this transformation induces its own problems: The transformed error term  $v_{it} - v_{i,t-1}$  is negatively correlated with the transformed lagged dependent variable  $y_{i,t-1} - y_{i,t-2}$ , so OLS estimation gives upward-biased convergence speeds. Thus, Caselli, Esquivel, and Lefort (1996) and others have estimated the model using GMM, with lagged regressors acting as instruments for the first-differenced variables. However, these lagged levels are only valid instruments under restrictive assumptions regarding the autocorrelation structure of the error term. In addition, Bond, Hoeffler, and Temple (2001) argue that the relatively slow pace of convergence implies that lagged levels of output per worker are typically very weak instruments for first differences. This leads to coefficients that are biased upwards towards their OLS levels.

The second difficulty is endogeneity bias. If shocks to output per worker also affect investment or population growth rates, then the convergence regression will contain endogenous regressors and estimates of the convergence speed may be biased. Caselli, Esquivel, and Lefort have argued in favor of the first-difference GMM approach as the best way of dealing with this endogeneity problem. However, as we have just noted, this approach suffers from other weaknesses that are not easily remedied.

An important advantage of our approach is that these two econometric problems, which have plagued the literature based on output per worker regressions, appear to pose far less difficulty for our methodology based on capital-output regressions. Our econometric tests all reject the presence of country-specific effects, so the estimated convergence speeds from our pooled OLS regressions are free from the downward bias associated with these regressions in the presence of fixed effects.

In relation to the problem of endogeneity bias, our method of directly assessing conditional convergence via the dynamics of the capital-output ratio is less likely to suffer from the endogeneity biases of the traditional approach. Much of the endogeneity bias in the traditional regressions likely stems from the joint effects that shocks to technological efficiency have on both output per worker and the regressors. For instance, a positive technology shock can simultaneously boost both economic growth and the investment share of GDP. Even if the causation problems run from output per worker to the regressors, so that there is a line of causation going from the capital-output ratio to output per worker to the regressors in our model, this endogeneity is likely to be quite weak. This is because variations in capital intensity account for a small fraction of the high-frequency fluctuations in output growth.

### 4.3 A Puzzle?

Table 7 reports results for three different estimation methods. As expected, pure cross-sectional estimation of the output per worker regression, equation (34), produces very low estimated convergence speeds, which likely reflects the downward bias due to the correlation between country-specific fixed effects and the initial output variable. Our data report a convergence speed of only 1.0 percent for equation (34), and 1.4 percent if this specification is augmented with a measure of school enrollment, as suggested by Mankiw, Romer and Weil. This latter estimate exactly matches the figure reported in their paper.

The middle panels report the results from within-group estimation of equation (34) using both one- and five-year intervals. As expected, these estimates are higher than those for the pure cross-sectional regressions, with the one-year estimate being 4.0 percent, and the five-year estimate being 5.8 percent. These estimates raise an interesting question. Within-groups estimation is normally understood to produce upward-biased estimates of convergence speeds. Indeed, Bond (2002) argues that the convergence speeds from within-groups estimation can generally be considered an upper bound. If our capital-output ratio

method has the advantages that we have indicated, then it is puzzling that our preferred estimates are higher than those produced by within-groups estimation.

We believe the explanation for this puzzle lies in a problem relating to output per worker regressions that has not been documented previously. Specifically, the standard assumption about the cross-country process for technology provides a very poor approximation to the actual process, and this results in a downward bias for estimated convergence speeds that generally dominates the traditional upward bias due to fixed effects.

## 5 Implications of Stochastic Technology

In this section, we document the inadequacy of the standard assumptions about technology and provide a simple model that fits the evidence better. We then explain the bias that the traditional assumption induces when estimating convergence speeds from output per worker regressions.

### 5.1 The Case Against the Fixed Level Effects Model

Figure 1 shows time series for TFP for four countries, from the top (USA), middle (Brazil and Seychelles) and bottom (Tanzania) of the world income distribution. The charts also show the fitted values from a regression fitting a deterministic trend to these TFP series.<sup>15</sup> The figure illustrates a number of ways in which the traditional fixed effects model of technology appears to be inaccurate.

First, note that at each point in time, TFP growth can take on quite different values across countries. While year-by-year values of TFP growth may not be wholly reliable proxies for underlying technology growth, the observation that there are significant variations across countries still holds if one smooths out the series by averaging across multiple years. For instance, using five-year averages of TFP growth as an indicator for the growth rate of technology, the standard deviation of TFP growth across our 96 countries still averages about two and half per cent. Contrasted with a median value for TFP growth of about one percent, these figures make clear that cross-country variation in TFP growth is a large and empirically important phenomenon.

Second, the figure makes clear that random country-specific shocks to TFP growth

---

<sup>15</sup>All the calculations reported here are based upon an assumed depreciation rate for capital of six percent, but the results are robust to the use of other values.

are an important source of this cross-country variation. In other words, the variation is not simply because some countries have high trend growth rates and some have low trend growth rates. As illustrated in the figure, even if we allowed for the idea that each country might have a separate trend growth rate for technology, these deterministic trends leave a large fraction of the variance of TFP unexplained: Across the 96 countries, the average  $R^2$  from individual regressions of the log of TFP on a time trend is 0.60.

Indeed, the coefficients on country-specific deterministic trends in these regressions should be interpreted carefully. If TFP growth is subjected to stochastic country-specific shocks, then such regressions could suggest spuriously different deterministic trends across countries. Indeed, once one accounts for the role played by random country-specific shocks, there is little evidence for differences in trend TFP growth rates across countries. To assess this question, we estimated a regression of the form

$$\Delta a_{it} = \sum_{j=1}^N \beta_j D_j + \sum_{t=1}^T \beta_k D_k + \epsilon_{i,t} \quad (37)$$

where  $D_j$  and  $D_k$  are country and time-dummies. Perhaps surprisingly, only 7 of the 96 country dummies proved to be significant at the five percent level. Thus, the combination of time effects capturing world technology developments, and country-specific shocks, appears to provide a good model of the cross-country process for technology.

These considerations suggest a model for technology of the form

$$\Delta a_{it} = g_t + \epsilon_{i,t} \quad (38)$$

This approach maintains the Mankiw-Romer-Weil idea of a common world technology trend representing advancement of knowledge, while also allowing for the country-specific shocks required to explain the evidence on TFP growth. While this specification requires a relatively small change from the standard assumption about technology (which implies  $\Delta a_{it} = g_t$ ) it turns out to have very important implications for output per worker regressions.

## 5.2 A New Source of Bias: Stochastic Technology Shocks

Once one allows for stochastic country-specific shocks to TFP growth, the level of technology becomes

$$a_{it} = a_{i0} + \sum_{m=1}^t g_m + u_{it}, \quad (39)$$

where

$$u_{it} = u_{i,t-1} + \epsilon_{it} = u_{i,t-r} + \sum_{m=0}^{r-1} \epsilon_{i,t-m}. \quad (40)$$

With this assumption, the technology term in the convergence relationship can be written as

$$a_{it} - e^{-\lambda r} a_{i,t-r} = (1 - e^{-\lambda r}) a_{i0} + \sum_{m=1}^t g_m - e^{-\lambda r} \sum_{m=1}^{t-r} g_m + u_{it} - e^{-\lambda r} u_{i,t-r}. \quad (41)$$

Thus, one can again obtain a growth regression of the form

$$y_{it} = \gamma_i + \omega_t + \frac{\alpha (1 - e^{-\lambda r})}{1 - \alpha} (\log(s_{i,t-r}) - \log(g_i + n_{i,t-r} + \delta)) + e^{-\lambda r} y_{i,t-r} + \eta_{it}, \quad (42)$$

where  $\gamma_i$  and  $\omega_t$  are as in (35) and (36). However, an important difference in this case is that the error term now takes the form

$$\eta_{it} = (1 - e^{-\lambda r}) u_{i,t-r} + \sum_{m=0}^{r-1} \epsilon_{i,t-m}. \quad (43)$$

These calculations show that once one allows for stochastic country-specific shocks to TFP growth (as suggested by the evidence) then the error term in the standard growth regression will contain the term  $u_{i,t-r}$ , which describes the effect of these shocks on the level of technology. This term is positively correlated with  $y_{i,t-r}$  and this will be a source of upward bias in the coefficient on this variable, and thus downward bias in the estimated convergence speed.<sup>16</sup> Whether this effect dominates, so that the within-groups estimator will be downward biased, will depend on a number of factors such as the length of time element of the panel. This is because the upward bias in estimated convergence speeds due to the ‘‘Nickell effect’’ related to fixed effects tends to zero as  $T \rightarrow \infty$ . In contrast, the downward bias due to stochastic technology will not disappear in panels with a long time series.

---

<sup>16</sup>Also worth noting is that the presence of  $u_{i,t-r}$  means that the identifying assumptions underlying the GMM estimators of Caselli, Esquivel, and Lefort (1996) and Bond, Hoeffler, and Temple (2001) will not hold. These models are based on the assumption that all autocorrelations for the error term are zero beyond two or three periods.

## 6 Monte Carlo Evidence

To illustrate the performance under realistic conditions of both output per worker regressions and our capital-output methodology, we undertook a number of Monte Carlo simulations. In each case, we simulated ninety-six separate Solow model economies, each with a technology process of the form

$$\Delta a_{it} = g + \epsilon_{it} \tag{44}$$

where  $g = 0.014$  and  $\epsilon_{it} \sim N(0, \sigma^2)$  with  $\sigma = 0.01$ . This calibrates  $g$  using the median growth rate of our estimated  $A_{it}$  series, and sets the volatility at a relatively high level, in light of the apparent importance of these shocks. In addition, the initial values for the country-specific technology series  $a_{i0}$  were drawn from a normal distribution with a standard deviation set to match the ratio of standard deviation to mean in our own estimates of the 1960 distribution of this series. The results from the simulations were not very sensitive to variations in the parameters of the technology process.

Concerning the other parameters of the model, the depreciation rate was set to  $\delta = 0.06$  in line with our preferred estimate, and the standard labor share value of two-thirds was used. We then assumed that all countries had population growth rates of  $n = 0.031$  percent. This is somewhat larger than implied by a realistic calibration, but it ensures that we have an average convergence speed of  $\lambda = (1 - \alpha)(g + n + \delta) = 0.07$  in line with our preferred empirical estimates. Finally, we assumed that each of these economies has an investment share of  $s = 0.105$ , which implies an equilibrium capital-output ratio of one, and also assumed that initial capital stocks are centered around this equilibrium with a standard deviation equal to ten percent of the equilibrium level. Again, the results reported here did not depend on these specific modelling choices; in particular, similar results were obtained from simulations in which the investment and population growth rates varied across the countries.

Table 8 reports results from a Monte Carlo exercise in which this model was simulated 1000 times with  $N = 96$  and  $T = 40$  chosen to match our dataset. The table gives the average convergence coefficients obtained from applying to the simulated data both our preferred method (pooled OLS estimation of the capital-output equations) and within-groups estimation of the standard output per worker equation. The simulations provide an important endorsement of our methodology. Both capital-output methodologies—based on one-year and five-year intervals—produce an average estimated convergence speed that almost exactly equals the underlying DGP’s “true” value of seven percent: The average

speed from the one-year method is  $\lambda = 0.071$ , while the average from the five-year method is  $\lambda = 0.072$ .

The table also shows how our preferred convergence speed of seven percent can be reconciled with the lower estimates obtained from the within-groups estimation. The average convergence speed obtained from the within-group estimation of the simulated data is  $\lambda = 0.053$  for one-year intervals and  $\lambda = 0.064$  for five-year intervals. Note that our simulation matches the pattern of the estimates from the one-year and five-year intervals produced by the within-groups method, with the five-year estimate being about one-percent higher.<sup>17</sup> In addition, they show that for panels with time series of the length used in our study the downward bias effect due to stochastic technology dominates the well-known upward bias from the “Nickell effect.”

Indeed, the Monte Carlo evidence suggests that our within-groups output per worker regressions come as close as they do to the “correct” convergence speed thanks to a somewhat fortuitous combination of offsetting biases. Table 9 illustrates this by reporting simulation results for alternative values of  $T$ . For smaller samples, the Nickell bias is very large. For example, when  $T = 20$ , the average convergence speed from within-groups output per worker regressions is sixteen percent for five-year intervals and eleven percent for one-year intervals. The reduction in the effective sample, and the consequent strengthening of the Nickell bias, explains why the estimated convergence speeds for five-year regressions are higher than those for one-year regressions. When  $T = 40$ , as in our application, the two biases are almost offsetting. However, for values of  $T$  that are higher than we used, we see the Nickell bias declining and the downward bias due to stochastic technology being more dominant. For instance, for time series with  $T = 150$ , the average convergence speeds from simulated output per worker regressions fall to just over one percent.

In contrast, for each of the sample sizes used, our capital-output methodology gives average convergence speeds that are always extremely close to the true value of seven percent. Thus, while the estimated convergence speed from the five-year output per worker regression, at almost six percent, is very close to the seven percent value given by our method, this should not be taken as evidence that these two methods need generally give similar answers. More importantly, our calculations show that it would be incorrect to

---

<sup>17</sup>Technically, the explanation for this pattern can be seen from equation (43) which shows that the term  $u_{i,t-r}$  is multiplied by  $(1 - e^{-\lambda r})$ , which gets smaller as  $r$  increases. Because the downward bias due to stochastic technology depends on the correlation between  $y_{i,t-r}$  and the part of the error term that depends on  $u_{i,t-r}$ , this bias gets smaller as  $r$  increases.

consider the convergence speeds from within-groups estimation of traditional output per worker regressions as being “upper bound” estimates.

## 7 Implications for Growth and Development

We have found that the process of conditional convergence appears to be well approximated by the predictions of the traditional (non-augmented) Solow model. These results clearly have important implications for the interpretation of cross-country growth patterns. Two issues, in particular, are worth highlighting.

The first is the role played by human capital in the growth process. Thus far, we have ignored human capital altogether, with productivity increases due to human capital implicitly incorporated into our estimates of TFP. However, our analysis would not be changed if we had re-expressed the production function as

$$Y_t = K_t^\alpha (A_t h_t L_t)^{1-\alpha} \quad (45)$$

where  $h_t$  is a measure of the average level of human capital per worker. This formulation, used by Hall and Jones (1997) and others, illustrates how improvements in educational standards can increase labor productivity, but does not change our analysis of convergence dynamics in any substantive way. The only difference is that the term  $A_t h_t$  plays the role that  $A_t$  played in our analysis.

This example shows that one can account for a role for human capital without necessarily re-formulating the Solow model along the lines of Mankiw, Romer, and Weil (1992) with a production function of the form  $Y_t = K_t^\alpha H_t^\beta (A_t L_t)^{1-\alpha-\beta}$  and exogenously given savings rates for both types of capital. This approach implies a particular specification for the evolution of human capital, such that increases in output due to higher savings rates for physical capital automatically also generate higher levels of human capital per worker. In practice, the relationship between accumulation of physical and human capital need not be so tight. Indeed, since the MRW model implies convergence speeds considerably slower than estimated here, we are inclined to reject it as a model of cross-country growth dynamics.

The second issue concerns the importance of the conditional convergence mechanism for explaining cross-country patterns of growth. We have stressed that the Solow model’s predictions about conditional convergence match the data well. In this sense, we would argue that the model should be taken *very* seriously. However, this is quite different from

Mankiw, Romer, and Weil’s position that a Solow model with *common technology growth* across countries is capable of explaining the majority of variation in cross-country growth rates, i.e. that conditional convergence dynamics provide the principal explanation for why some countries grow faster than others. In fact, simple calculations suggest that the conditional convergence mechanism is actually of limited importance in explaining international variations in growth rates.

For instance, growth accounting calculations using our data show that at least two-thirds of the variation in cross-country growth rates over the period 1960-2000 are accounted for by variations in TFP growth. This leaves only a small fraction of the cross-country variation in growth rates to be explained by conditional convergence dynamics. Of course, the idea that variations in TFP growth are required to explain variations in output growth across countries is hardly inconsistent with the underlying message of the Solow model that technological progress is the ultimate source of all growth.

## 8 Conclusions

Solow’s (1956) model of economic growth still represents the starting point for most analysis of the processes underlying long-run economic growth. However, despite its popularity as a pedagogical tool, many believe that the model does not provide an adequate picture of the processes underlying long-run growth dynamics. In particular, the large literature on empirical growth regressions has generally suggested that real-world economies converge towards their steady-states at speeds considerably slower than predicted by Solow’s original model.

Our paper has questioned this conventional wisdom, on two fronts. First, we have questioned the evidence on convergence speeds from existing cross-country growth regressions. It is well known that there are a number of difficult econometric problems associated with estimating convergence speeds from these regressions, including the presence of country-specific fixed effects and endogenous regressors. We have also documented an additional source of econometric difficulty due to stochastic technology shocks, and shown how this bias is likely to be more important than the well-known finite-sample bias for panel estimators documented by Nickell (1981). Taken together, we think these results point against the usefulness of traditional panel regressions as tools for estimating convergence speeds.

Second, we have introduced a new methodology for estimating conditional convergence

speeds, based on the dynamics of the capital-output ratio. In light of the substantial econometric problems surrounding traditional growth regressions, we think the methodology suggested in this paper provides a simple and direct way of estimating the speed of conditional convergence. By focusing directly on the speed at which capital-output ratios converge towards their target values, our approach focuses on the precise form of convergence dynamics predicted by the Solow model. It also avoids the need to deal with the technicalities relating to country-specific fixed effects and has good properties when economies are characterized by stochastic technology shocks.

Our finding that convergence speeds are consistently close to or slightly higher than predicted by the Solow model should be of interest to those wishing to understand the factors underlying cross-country growth patterns. One possible explanation for this fast speed of convergence is that the existence of mobile international capital may allow economies to converge somewhat faster than implied by the closed-economy Solow model, a possibility that has previously been raised by Barro, Mankiw, and Sala-i-Martin (1995). In addition, our results imply that one does not need to adopt models with broader concepts of capital, and thus slower convergence speeds, to understand the processes generating growth patterns across countries.

## References

- [1] Barro, Robert and Xavier Sala-i-Martin (1995). *Economic Growth*, McGraw-Hill.
- [2] Barro, Robert, N. Gregory Mankiw, and Xavier Sala-i-Martin (1995). Capital Mobility in Neoclassical Models of Growth, *American Economic Review*, 85, 103-115.
- [3] Bernanke, Ben and Refet Gurkaynak (2002). "Is Growth Exogenous? Taking Mankiw, Romer and Weil Seriously," *NBER Macroeconomics Annual*, pp. 11-57.
- [4] Bond, Stephen, Anke Hoeffler, and Jonathan Temple (2001). GMM Estimation of Empirical Growth Models, CEPR Discussion Paper No. 3048.
- [5] Bond, Stephen (2002). Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice, UCL Working paper CWP09/02.

- [6] 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, 363-389.
- [7] DeLong, J. Bradford (2003). *Macroeconomics*, McGraw-Hill.
- [8] Fraumeni, Barbara (1997). "The Measurement of Depreciation in the US National Income and Product Accounts," *Survey of Current Business*, July, 7-23.
- [9] Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell (1997). "Long-Run Implications of Investment-Specific Technological Change." *American Economic Review*, 87, 342-362.
- [10] Hall, Robert and Charles I. Jones (1997). Why Do Some Countries Produce so Much More per Worker than Others?, *Quarterly Journal of Economics*, 114, 83-116.
- [11] Heston, Alan, Robert Summers, and Bettina Aten (2002). Penn World Table Version 6.1, Center for International Comparisons at the University of Pennsylvania.
- [12] Islam, Nazrul (1995). "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 110, 1127-1170.
- [13] Jones, Charles I. (2000). A Note on the Closed-Form Solution of the Solow Model, mimeo, Berkeley.
- [14] Jones, Charles I. and Dean Scrimgeour (2005). The Steady-State Growth Theorem: Understanding Uzawa (1961), mimeo, Berkeley.
- [15] Mankiw, N. Gregory, David Romer, and David Weil (1992). "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, 407-437.
- [16] Mankiw, N. Gregory (1995). "The Growth of Nations," *Brookings Papers on Economic Activity*, Volume 1, 275-326.
- [17] Nickell, Stephen (1981). "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49, 1417-1426.
- [18] Romer, David (2001). *Advanced Macroeconomics*, McGraw-Hill.
- [19] Solow, Robert (1956). "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, 70, 65-94.

- [20] Whelan, Karl (2003). "A Two-Sector Approach to Modeling U.S. NIPA Data." *Journal of Money, Credit, and Banking*, 35, 627-656.

## A Definitions of Data Sets

All data come from the Heston, Summers, and Aten (2002) Penn World Tables, Version 6.1. The following list defines the three data sets used in our regressions.

---



---

Argentina	Ecuador	Jordan	Philippines
Australia <sup>O</sup>	Egypt	Kenya	Portugal <sup>O</sup>
Austria <sup>O</sup>	El Salvador	Korea <sup>O</sup>	Romania
Bangladesh	Ethiopia	Lesotho <sup>NI</sup>	Rwanda
Barbados <sup>NI</sup>	Finland <sup>O</sup>	Luxembourg <sup>NI</sup>	Senegal
Belgium <sup>O</sup>	France <sup>O</sup>	Madagascar	Seychelles <sup>NI</sup>
Benin	Gabon <sup>NI</sup>	Malawi	South Africa
Bolivia	Gambia <sup>NI</sup>	Malaysia	Spain <sup>O</sup>
Brazil	Ghana	Mali	Sri Lanka
Burkina Faso	Greece <sup>O</sup>	Mauritius <sup>NI</sup>	Sweden <sup>O</sup>
Burundi	Guatemala	Mexico <sup>O</sup>	Switzerland <sup>O</sup>
Cameroon	Guinea	Morocco	Syria
Canada <sup>O</sup>	Guinea-Bissau <sup>NI</sup>	Mozambique <sup>NI</sup>	Tanzania
Cape Verde <sup>NI</sup>	Honduras	Nepal	Thailand
Chad <sup>NI</sup>	Hong Kong	Netherlands <sup>O</sup>	Togo <sup>NI</sup>
Chile	Iceland <sup>NI</sup>	New Zealand <sup>O</sup>	Trinidad and Tobago
China	India	Nicaragua	Turkey <sup>O</sup>
Colombia	Indonesia	Niger <sup>NI</sup>	Uganda <sup>NI</sup>
CComoros <sup>NI</sup>	Iran	Nigeria	United Kingdom <sup>O</sup>
Congo	Ireland <sup>O</sup>	Norway <sup>O</sup>	Uruguay
Costa Rica	Israel	Pakistan	USA <sup>O</sup>
Cote D'Ivoire	Italy <sup>O</sup>	Panama	Venezuela
Denmark <sup>O</sup>	Jamaica	Paraguay	Zambia
Dominican Republic	Japan <sup>O</sup>	Peru	Zimbabwe

---



---

**Note:** *NI* means a country is *not* a member of the Intermediate sub-sample, while *O* means a country *is* a member of the OECD sub-sample.

## B A Solow Model with Two Types of Capital

Here, we consider a Solow model with a two types of capital, equipment ( $E_t$ ) and structures ( $S_t$ ).

$$Y_t = S_t^\gamma E_t^{\alpha-\gamma} (A_t L_t)^{1-\alpha},$$

The two types of capital accumulate according to

$$\begin{aligned}\frac{\dot{S}_t}{S_t} &= \frac{s_s Y_t}{S_t} - \delta_s \\ \frac{\dot{E}_t}{E_t} &= \frac{s_e Y_t}{E_t} - \delta_e\end{aligned}$$

The growth rate of output in this case is

$$\frac{\dot{Y}_t}{Y_t} = \gamma \frac{\dot{S}_t}{S_t} + (\alpha - \gamma) \frac{\dot{E}_t}{E_t} + (1 - \alpha) \left( \frac{\dot{A}_t}{A_t} + \frac{\dot{L}_t}{L_t} \right).$$

If we define an aggregate capital stock as

$$K = S_t^\gamma E_t^{\alpha-\gamma},$$

then the growth rate of this measure of capital is given by

$$\frac{\dot{K}_t}{K_t} = \frac{\gamma \dot{S}_t}{\alpha S_t} + \frac{\alpha - \gamma}{\alpha} \frac{\dot{E}_t}{E_t}$$

Thus, the output equation can be re-written in the usual form as

$$\frac{\dot{Y}_t}{Y_t} = \alpha \frac{\dot{K}_t}{K_t} + (1 - \alpha) \left( \frac{\dot{A}_t}{A_t} + \frac{\dot{L}_t}{L_t} \right)$$

Now turning to the capital growth equation, this is given by

$$\begin{aligned}\frac{\dot{K}_t}{K_t} &= \frac{\gamma s_s Y_t}{\alpha S_t} + \frac{\alpha - \gamma}{\alpha} \frac{s_e Y_t}{E_t} - \left( \frac{\gamma}{\alpha} \delta_s + \frac{\alpha - \gamma}{\alpha} \delta_e \right) \\ &= \frac{s Y_t}{K_t} - \left( \frac{\gamma}{\alpha} \delta_s + \frac{\alpha - \gamma}{\alpha} \delta_e \right)\end{aligned}$$

where

$$s = \left[ \frac{\gamma}{\alpha} \left( \frac{E_t}{S_t} \right)^{\frac{\alpha-\gamma}{\alpha}} \right] s_s + \left[ \frac{\alpha - \gamma}{\alpha} \left( \frac{E_t}{S_t} \right)^{-\frac{\gamma}{\alpha}} \right] s_e$$

These calculations show that the equations for our model can also represent a model with multiple types of capital. However, in this case the appropriate depreciation rate in the capital-growth equation is a weighted average of the underlying rates with the weights determined by the exponents  $\alpha$  and  $\alpha - \gamma$ .

Table 1: Dynamic Capital-Output Ratio Model: One-Year Intervals

Estimator	OLS	WG	$\lambda^{Solow}$
<hr/> $\delta = 0.06$ <hr/>			
$\lambda$	0.072 (0.003)	0.074 (0.004)	0.063
P-Value of Fixed Effects	0.965		
<hr/> $\delta = 0.05$ <hr/>			
$\lambda$	0.065 (0.003)	0.067 (0.004)	0.056
P-Value of Fixed Effects	0.962		
<hr/> $\delta = 0.04$ <hr/>			
$\lambda$	0.059 (0.002)	0.062 (0.003)	0.049
P-Value of Fixed Effects	0.923		

**Note:** Standard errors in parenthesis. WG denotes the within-groups (least squares dummy variables) estimator, while the p-value is from the null hypothesis of no fixed effects.  $\lambda^{Solow}$  refers to the convergence rate suggested by the Solow model as given by  $(1 - \alpha)(\bar{g}_i + \bar{n}_i + \delta)$ .

Table 2: Dynamic Capital-Output Ratio Model: Five-Year Intervals

Estimator	OLS	WG	$\lambda^{Solow}$
<hr/>			
$\delta = 0.06$			
$\lambda$	0.069 (0.002)	0.075 (0.002)	0.063
P-Value of Fixed Effects	0.517		
<hr/>			
$\delta = 0.05$			
$\lambda$	0.063 (0.002)	0.069 (0.003)	0.056
P-Value of Fixed Effects	0.529		
<hr/>			
$\delta = 0.04$			
$\lambda$	0.056 (0.002)	0.062 (0.003)	0.049
P-Value of Fixed Effects	0.505		
<hr/>			

**Note:** Standard errors in parenthesis. WG denotes the within-groups (least squares dummy variables) estimator, while the p-value is from the null hypothesis of no fixed effects.  $\lambda^{Solow}$  refers to the convergence rate suggested by the Solow model as given by  $(1 - \alpha)(\bar{g}_i + \bar{n}_i + \delta)$ .

Table 3: Capital-Output Ratio Model with Additional Dynamics ( $\delta = 0.06$ )

Estimator	OLS	WG
<u>1 Year Interval</u>		
$\lambda$	0.067 (0.003)	0.071 (0.004)
$(x_{i,t-1} - x_{i,t-2})$	0.085 (0.029)	0.072 (0.029)
P-Value of Fixed Effects	0.993	
<u>5 Year Interval</u>		
$\lambda$	0.071 (0.002)	0.077 (0.003)
$(x_{i,t-5} - x_{i,t-10})$	-0.021 (0.041)	-0.075 (0.045)
P-Value of Fixed Effects	0.499	

**Note:** Standard errors in parenthesis. WG denotes the within-groups (least squares dummy variables) estimator, while the p-value is from the null hypothesis of no fixed effects.  $(1 - \alpha)(\bar{g}_i + \bar{n}_i + \delta)$  for the depreciation rate (6 per cent) is as in Table 2.

Table 4: Unrestricted Dynamic Capital-Output Ratio Model ( $\delta = 0.06$ )

Estimator	OLS	WG
<b>1 Year Interval</b>		
$x_{i,t-1}$	-0.069 (0.003)	-0.085 (0.004)
$\log s_{it}$	0.063 (0.003)	0.063 (0.004)
$\log[g_i + n_{it} + \delta]$	-0.066 (0.003)	-0.042 (0.017)
P-Value of Fixed Effects	0.525	
<b>5 Year Interval</b>		
$x_{i,t-5}$	-0.287 (0.012)	-0.377 (0.020)
$\log s_{it}^5$	0.263 (0.013)	0.265 (0.018)
$\log[g_i + n_{it}^5 + \delta]$	-0.276 (0.030)	-0.398 (0.122)
Implied $\lambda$	0.068 (0.002)	0.095 (0.003)
P-Value of Fixed Effects	0.353	

**Note:** Standard errors in parenthesis. WG denotes the within-groups (least squares dummy variables) estimator, while the p-value is from the null hypothesis of no fixed effects.  $s_{it}^5$  and  $n_{it}^5$  refer to five-year moving average values of the investment rate and population growth rate.

Table 5: Results for Smaller Country Samples ( $\delta = 0.06$ )

Estimator	Intermediate Sample		OECD Sample	
	OLS	WG	OLS	WG
<u>1 Year Interval</u>				
$\lambda$	0.070 (0.003)	0.071 (0.004)	0.051 (0.004)	0.040 (0.005)
P-Value of Fixed Effects	0.975		0.291	
<u>5 Year Interval</u>				
$\lambda$	0.070 (0.002)	0.077 (0.003)	0.067 (0.003)	0.064 (0.004)
P-Value of Fixed Effects	0.634		0.764	

**Note:** Standard errors in parenthesis. WG denotes the within-groups (least squares dummy variables) estimator, while the p-value is from the null hypothesis of no fixed effects.

Table 6: Results from Running 96 Country-Specific Regressions

	$\lambda^{Solow}$	$\lambda^{Average}$	Std. Deviation	10 <sup>th</sup> Percentile	90 <sup>th</sup> Percentile
<u><math>\delta = 0.06</math></u>	0.063	0.063	0.030	0.033	0.103
<u><math>\delta = 0.05</math></u>	0.056	0.057	0.028	0.028	0.094
<u><math>\delta = 0.04</math></u>	0.049	0.051	0.026	0.021	0.086

**Note:**  $\lambda^{Solow}$  refers to the average convergence rate suggested by the Solow model as given by  $(1 - \alpha)(\bar{g}_i + \bar{n}_i + \delta)$ .  $\lambda^{Average}$  is the average of the 96 country-specific  $\lambda$ 's estimated along with the associated standard deviations and results for the 10<sup>th</sup> and 90<sup>th</sup> percentile.

Table 7: Estimated Convergence Rate:  $K/Y$  and  $Y/L$  Regressions

Dependent Variable	Estimator	Time Interval	Estimated $\lambda$
Output per Worker	Cross-Sectional Without School Enrollment	40 Years	0.010 (0.001)
Output per Worker	Cross-Sectional With School Enrollment	40 Years	0.014 (0.001)
Output per Worker	Within-Groups	1-Year	0.040 (0.006)
Output per Worker	Within-Groups	5-Year	0.058 (0.009)
Capital Output Ratio	Pooled OLS	1-Year	0.072 (0.003)
Capital Output Ratio	Pooled OLS	5-Year	0.069 (0.002)

**Note:** Standard errors in parentheses. The within-groups Output per Worker regression replicates the method used by Islam (1995), while the cross-sectional Output per Worker regression replicates the approach of Mankiw, Romer, and Weil (1992).

Table 8: Monte Carlo Results For True  $\lambda = 0.07$ , (N=96,T=40)

Dependent Variable	Estimator	Time Interval	Average Simulated $\lambda$
Capital Output Ratio	Pooled OLS	1-Year	0.071
Capital Output Ratio	Pooled OLS	5-Year	0.072
Output per Worker	Within-Groups	1-Year	0.053
Output per Worker	Within-Groups	5-Year	0.064

**Note:** ‘N’ is the number of countries in the sample while ‘T’ is the size of the time period. See the text for details of the simulations performed.

Table 9: Monte Carlo Results For True  $\lambda = 0.07$  and Alternative Sample Sizes

Dependent Variable	Time Interval	Simulated $\lambda$
<hr/> T=20 <hr/>		
Capital Output Ratio	1-Year	0.069
Capital Output Ratio	5-Year	0.068
Output per Worker	1-Year	0.110
Output per Worker	5-Year	0.164
<hr/> T=80 <hr/>		
Capital Output Ratio	1-Year	0.071
Capital Output Ratio	5-Year	0.072
Output per Worker	1-Year	0.022
Output per Worker	5-Year	0.027
<hr/> T=150 <hr/>		
Capital Output Ratio	1-Year	0.070
Capital Output Ratio	5-Year	0.070
Output per Worker	1-Year	0.012
Output per Worker	5-Year	0.013

**Note:** 'N', the number of countries is kept constant at 96 in all exercises. See the text for details of the simulations performed.

**Figure 1: Actual and Trended TFP for Select Countries 1960-2000**

