4

Convergence?

4.1. Introduction

We now come to a central prediction of the Solow growth model. It states that two countries that are the same in all their parameters — savings rates, population growth rates, rates of technical progress, and so on — must ultimately exhibit similar levels of per-capita income. That's because capital per efficiency unit of labor must converge to a steady state value that's common to both countries. Indeed, this will happen *irrespective of the initial state of each of these economies*, as measured by their starting levels of per-capita income (or equivalently, their per-capita capital stock).

Does this sound trivial to you, or totally wild? The pro-trivial camp would say: we're assuming that all technological and behavioral parameters are similar. How could we expect anything *but* convergence? The pro-wild camp would counter: the assertion is far from obvious. Assumed to be the same are exogenous parameters of the model, true, but *not* the initial level of the capital stock or per-capita income. The claim of convergence is then based on the analysis that we conducted for the Solow model, which in turn is based on diminishing returns to accumulable inputs such as physical capital. its content is that. History, in the sense of different initial conditions, does not matter. All roads lead to a common steady state growth path.

Convergence is a provocative and important idea. In a sense, this theme (or often its negation) pervades the entire book. If true, convergence would imply that the market mechanism powerfully and progressively eliminates history through its workings, acting as a great leveler as it narrows inter-country disparities. In what follows, we parse this idea carefully, and look at the available evidence for or against convergence.

4.2. Convergence: Unconditional and Conditional

The prediction of progressive income equalization across countries relies both on the presumption of diminishing returns *and* the assumption that all parameters are the same across countries. It is an uncompromising prediction, built on strong assumptions. We therefore give it an equally uncompromising name: *unconditional convergence*. Later, we introduce a more accommodating sibling, *conditional* convergence, that permits country-level parameters to vary.

4.2.1. Unconditional Convergence. Figure 4.1 plots the logarithm of income against time, so that a constant rate of growth appears as a straight line; recall Chapter 3, Section 3.4.2. The line AB depicts the path of (log) per-capita income in steady state, with income per efficiency unit of labor perennially at $\hat{y}^* = f(\hat{k}^*)$. The path CD represents a country that starts below this steady-state trajectory. According to the Solow model, this country initially displays a rate of growth that *exceeds* the steady-state level. Its time path of (log) per-capita income moves asymptotically toward the AB line as shown, and its growth rate decelerates to the steady-state level. Likewise, a country that starts off at C' above the steady state, experiences a lower-than-steady-state rate of growth. Its time path $C'D'$ of (log) income flattens out to converge to AB from above. At any rate, that is what the hypothesis has to say.

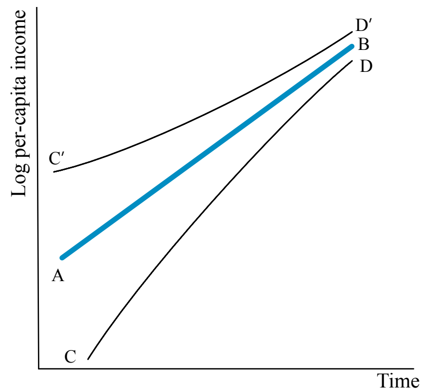


Figure 4.1. Unconditional convergence.

There's a particular growth regression equation that corresponds to this discussion. Figure 4.1 suggests that the growth rate of income per efficiency unit at any date t is proportional to the *gap* between $\ln \hat{y}(t)$ and $\ln \hat{y}^*$, where the former is the actual value of $\hat{y}$ at t and the latter is its steady state value. Converting these objects to per-capita incomes $y(t)$, the average growth rate at any date t is given by

$$g_i(t) = \alpha(t) + \beta \ln y_i(t) + \epsilon_i(t), \tag{4.1}$$

where g_i is the annualized growth rate of per-capita income, $\alpha(t)$ is an "intercept term" that contains the steady state information (assumed to be common across all countries) and ϵ_i is country-specific noise. Appendix A describes precisely how this equation is arrived at. You should note that the term $g_i(t)$ is not the steady state rate of growth of per-capita income, which is presumed to be the same across countries. Rather, it includes the speed of "recovery to the steady state" when a country is away from that steady state to start with.

Circumstantial evidence for unconditional convergence is indicated by a negative value of β, a claim often referred to as "β-convergence." That is, the higher is per-capita income, the lower its subsequent rate of growth. Indeed, because growth $g_i(t)$ over some interval $[t, t+1]$ can be approximately written as $g_i(t) \simeq \ln y_i(t+1) - \ln y_i(t)$,[34] a coefficient of $\beta \simeq 1$ means that all initial income differences are fully wiped out by subsequent convergence. In contrast $\beta \simeq 0$ means that initial incomes have no effect at all on subsequent growth rates.

4.2.2. Conditional Convergence. The prediction of unconditional convergence assumes that the level of and rate of change in technical knowledge, the rate of savings, the rate of population growth, and other parameters determining the steady state are the same across countries. (Or even if they're not, they too converge; see below.)

[34] As we already observed for our doubling time formula in footnote 14, $\ln(1+g) \simeq g$ when g is small, so $\ln y_i(t+1) - \ln y_i(t) = \ln y_i(t)(1+g) - \ln y_i(t) = \ln(1+g) \simeq g$.

While such similarities or differences have no effect on the Solow prediction that countries must converge to *their* steady states, those steady states could persistently vary across countries, so that there is no need for two countries to converge to *each other*. This qualification leads to the weaker notion of *conditional convergence*. Figure 4.2 illustrates. Given the parametric variation cross-nationally, different countries have their own steady-state paths, as illustrated by the lines AB and $A'B'$. If the rate of technical progress is the same across countries, then these paths will all be parallel to one another. The notion of conditional convergence typically maintains this equality of technical change as a working assumption, which is of course far weaker than the assumptions made for the unconditional variant.

Now imagine that the country with steady-state path AB starts at a point C above that path. The Solow model predicts that over time, this country will exhibit a slower rate of growth than the steady-state path as it slopes in to its steady-state path AB. This path is given by the curve CD. Likewise, a country that starts at point C' below its steady-state path $A'B'$ will exhibit a rate of growth higher than that of the steady state, with the resulting path $C'D'$ converging upward to its steady-state path.

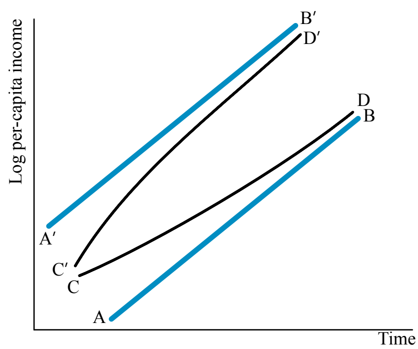


Figure 4.2. Conditional convergence.

Notice that C' is higher than C, so the negative relationship between initial incomes and subsequent growth rates — a hallmark of unconditional convergence enshrined in equation (4.1) — is thereby broken. So the conditional convergence hypothesis is more subtle. It states that if we, the analysts, could somehow infer the different locations of the steady state paths from country-specific parameters, we would be able to map *observed* income differences across countries to the *predicted* steady state values of income implied by those country-specific parameters. The widespread conditioning of predictions on parameters earns this concept the name *conditional* convergence. The growth regression equation (4.1) extends in a natural way to accommodate this idea:

$$g_i(t) = \alpha(t) + \beta \ln y_i(t) + \sum_{j=1}^{m} \gamma_j x_{ij} + \epsilon_i(t), \tag{4.2}$$

where we enter all kinds of parameters of the form x_{ij}, each recording the value of the jth parameter (say the savings rate in the base year) for country i. This regression equation is more accommodating than its sibling (4.1), because it permits parameters to vary across countries. Those parameters determine different steady states that are now country-specific, and so "control" or "condition" for the location of those steady states via the term $\sum_{j=1}^{m} \gamma_j x_{ij}$, and as shown graphically in Figure 4.2. Equation (4.2) *then* asks the question of whether there is β-convergence *after* the conditioning. See Section 4.4 and Appendix A for a more detailed development of this idea. With panel data for several countries over two or more sets of time intervals, one could also replace these parametric values by country fixed effects, which represents a more abstract way of conditioning for country-level differences that are presumably invariant over time.

As with its unconditional variant, conditional β-convergence allows for countries to be out of steady state, which in particular explains why the parameters x_{ij} could influence the growth rate in equation (4.2) *even if* the model is one of exogenous growth. Again, g_i is not a country's steady state growth rate, but gravitates to it.

4.2.3. Parametric Convergence. Unconditional convergence is a bold assertion, in large part because it presumes that all the so-called parameters of development are common to different countries. Among those parameters, we have so far emphasized those in the basic Solow model: savings rates, population growth rates and technical change. But of course, all sorts of parameters could affect the steady states of a more broadly-specified model: educational capacity, political institutions, legal arrangements, culture, religion; you name it. To the truly broad scholar of economic development, as we must all aspire to be, none of this should escape our attention simply by a restrictive declaration of exogeneity. In that sense, conditional convergence is a bit limited. It displays an unseemly haste to get important objects out of the way so that we can continue in our quest to show that the playing field is "conditionally" level. The real differences could well lie elsewhere — and they deserve investigation and explanation.[35] Rather than view them as impediments to the holy grail of "convergence," we should train our eye upon the determinants of savings, demographic change, technical progress, and the like; even culture, politics and institutions.

Parametric convergence refers to the question of inter-country convergence of the parameters themselves. Or put another way, it shifts the question of convergence to accompanying development outcomes other than incomes. Each such outcome — whether a measurable variable such as the saving rates or more nebulous objects such as norms and institutions — is worthy of study in itself, and indeed that is one way to view this entire book. For now, we use "parametric convergence" as a placeholder term where we put in the incredibly complex question of whether all or some or none of these parameters *also* converge to common values across societies. Viewed in this way, you can usefully think of unconditional convergence as a special case of conditional convergence, and parametric convergence as the force that unifies the two.

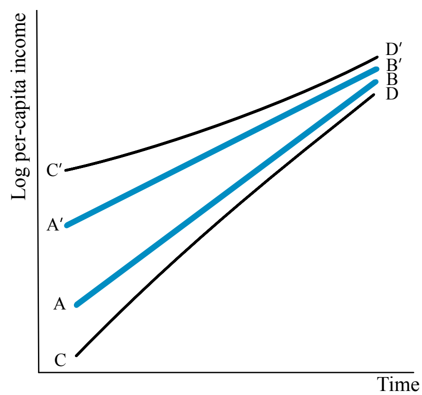


Figure 4.3. Parametric convergence.

Figure 4.3 summarizes graphically by combining ideas in Figures 4.1 and 4.2. There are two countries with steady state paths AB and $A'B'$. Those paths get ever closer to each other: the various parameters determining the two steady states *are themselves converging*. This is parametric convergence. On the same diagram, we overlay CD

[35] Perhaps some of these so-called "parameters" are really beyond the reach of economics and are relatively immutable, such as geography, though even here we know that factors such as transportation technologies could cause the influence of such factors to sharpen or weaken. The same is also true of "climatic factors" — think of induced climate change. But more emphatically, many "parameters" are plausibly direct outcomes of the development process.

and $C'D'$, one for a country starting at C with the steady state path AB, the other for a country starting at C' with the steady state path $A'B'$. Conditional convergence asserts the convergence of these paths to their respective steady states. Unconditional convergence is viewable as a combination of these two tendencies, predicting that CD and $C'D'$ tend to each other. It is in this sense that conditional and parametric convergence are modular components of unconditional convergence.

4.3. Act I: Convergence? 1870–1979

The first problem that arises when testing for convergence is the issue of time horizons. Ideally, we'd like to go back at least a century or more in history, but the systematic collection of data in developing economies is a modern phenomenon — certainly under a century old. When the first edition of this book was written in 1998, there were essentially two choices: cover a relatively small number of countries over a long period of time or cover a larger number of countries over a shorter time horizon. That we could even contemplate the former owed much to the economic historian Angus Maddison, who compiled per-capita income data on several countries going back to the mid-19th century (and earlier). The Maddison Project — see Maddison (1982, 1991, 2007) — continues today, nurtured by Maddison's colleagues and other eminent scholars.[36] As I write this in 2023, the 2020 Maddison database covers 169 countries up to 2018, with many going back to the 19th century, a large majority to at least 1950, and well over 60 to 1870. This is a laudable advance.

4.3.1. A First Pass. Our complaints about inadequate data held true *a fortiori* in 1986, when William Baumol published one of the first studies of unconditional convergence: a brave leap! At the time, there were just sixteen countries in Maddison's database for which "reliable" estimates of per-capita income existed (and I put "reliable" in quotes because this sort of historical detective work must always be taken with a large pinch of salt). These were, in order of poorest to richest in 1870: Japan, Finland, Sweden, Norway, Germany, Italy, Austria, France, Canada, Denmark, the United States, the Netherlands, Switzerland, Belgium, the United Kingdom, and Australia. They are among the richest countries in the world today.

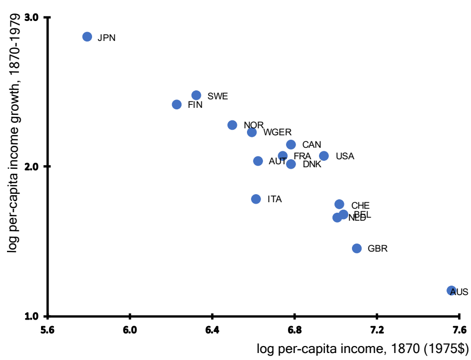


Figure 4.4. Growth and per-capita income for Baumol's 16 countries.

Figure 4.4 illustrates the exercise that Baumol conducted. Recalling the β- convergence growth equation (4.1), it plots 1870 log per-capita income for these sixteen countries on the horizontal axis, and the *growth rate* of that income over 1870–1979 (measured by the difference in the logs of per capita income) on the vertical axis. The convergence of these sixteen countries

[36] See the Maddison Project Database at `https://www.rug.nl/ggdc/historicaldevelopment/maddison/`, and consult Bolt and van Zanden (2020) for details on the Project.

to one another, starting from widely different levels of per-capita income in 1870, is unmistakeable. The value of β corresponding to Baumol's regression is a large negative 0.995. (Remember from our discussion of equation (4.1) that a slope of -1 means that by the terminal year, all the initial gaps in per capita income have been erased.)

Alas, there's a classic statistical pitfall lurking both in the picture and in the study. The sixteen countries are the first to have historical records for good reason: they are rich countries today! Yet in 1870 they were all over the economic map. Japan is a perfect case in point. It is there in large part because it is rich today and has kept records of its economic past, but in 1870, it was probably midway in the world's hierarchy of nations arranged by per-capita income.[37] In short, rich countries are likely to have come from a variety of backgrounds and will therefore *appear* to have converged. A true test of convergence would ask for no ex ante selection of countries at all. Or, to be a bit more relaxed about it, one would choose a set of countries that, *ex ante* would have a "forward shot" at convergence.

Bradford De Long (1988) addressed this question by adding, to Maddison's sixteen, seven other countries, which in 1870 had as much claim to membership in the "convergence club" as many included in Baumol's original set. These additional countries are Argentina, Chile, East Germany,[38] Ireland, New Zealand, Portugal, and Spain. Three of these countries (New Zealand, Argentina, and Chile) figure in the list of top ten recipients of British and French overseas investment per-capita as late as 1913, and a favorable perception regarding the growth prospects of these countries was widespread. *All* the new countries included had per capita GDP levels in 1870 higher than Finland, which was the second lowest in Baumol's sample. The lowest is Japan, the inclusion of which suffers from the opposite problem: if it's there in 1870, so should be half the world's countries, but accurate data for all those countries did (do) not exist.

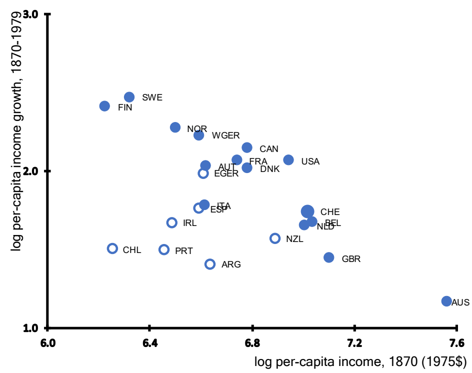


Figure 4.5. Growth and per-capita income for 22 countries. New countries shown by unfilled circles.

Figure 4.5 shows the modified pictures after De Long's countries are added and Japan is dropped from the initial sixteen. The earlier observations from Figure 4.4 appear as filled dots. Now matters don't look so good for convergence. Additionally, the 1870 data are likely to contain large measurement errors (relative to those in 1979), which make the various observations more scattered than they actually should be and makes any measurement of convergence more inflated than the case actually

[37]Here's an analogy: look at today's successful basketball stars. They would come from a variety of backgrounds; some poor and some rich. You could say that they "converged" to success, and in fact the rags-to-riches stories that we see often in the media bolsters such silly perceptions. After all, what about all the poor kids who *didn't* make it? Hindsight is no substitute for prediction.

[38]Note that the end of this period, 1979, predates the fall of the Berlin Wall by 10 years!

merits.[39] With reasonable measurement error accounted for (see De Long 1988), the β of equation (4.1) is now close to zero and insignificant.

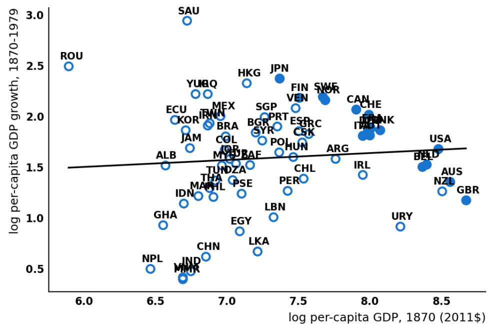


Figure 4.6. Growth and per-capita income, 1870–1979, for 66 countries from the Maddison Project. Original 16 in filled circles. Source: Maddison Project Database, version 2020, Bolt, Jutta and Jan Luiten van Zanden (2020).

Decades have passed since De Long wrote his critique, and now we have many more countries — over 60 of them — including several Latin American countries and Asian countries such as India or China, for which estimates of per-capita income in 1870 are available. The data continue to be overwhelmingly absent for Africa, so the selection problem hasn't gone away by any means. Yet, *despite that*, there is no sign of convergence over the 140 years from 1870 to 1979, even with the implicit selection bias arising from incomplete coverage. Figure 4.6 makes this clear by running the same exercise for the updated Maddison dataset. The filled data points represent Baumol's original 16, with their now-familiar downward sloping signature of convergence, though the data, now in 2011 USD and running over 1870-1979, are not the same. But as you can see, that convergence is nowhere to be found among the larger set of countries depicted in the Figure, even though these countries still constitute a selective subset for which income data are available for 1870.[40] Indeed, the best-linear regression (shown) exhibits divergence.

Yes, as Lant Pritchett (1997) has described it, it's really been "divergence, big time" over this period. This is especially so if we make a fair attempt to include missing estimates for the 1870s. Pritchett suggests a convincing way to do so. He makes the assumption that no country in the last century could have reasonably fallen below \$250 per capita in 1985 PPP USD. In Pritchett's words,

> "There is little doubt life was nasty, brutish and short in many countries in 1870. But even deprivation has its limit, and some per capita incomes must imply standards of living that are unsustainably and implausibly low. After making conservative use of a wide variety of different methods and approaches, I conclude that \$250 (expressed in 1985 purchasing power equivalents) is the lowest GDP per capita could have been in 1870. This figure can be defended on three grounds: first, no one has ever observed consistently lower living standards at any time or place in history; second, this level is well below extreme poverty lines actually set in impoverished

[39] Imagine that a set of families all have the same incomes in 1960 as well as in 2000. They are surveyed in both these years. However, the survey in 1960 is inaccurate, so there are errors of measurement, whereas the survey in 2000 is accurate, showing that they all have the same income. Then these families will appear to have "converged" to the same income from different starting levels.

[40] If Baumol's original study was so far off the mark, why include it here? The answer is that it is wrong in an interesting way. It illustrates an important pitfall in empirical reasoning, and the problems in connecting theory to data. The study is just as important for its error, as for the bold leap it takes with the available data.

> countries and is inconsistent with plausible levels of nutritional intake; and third, at a lower standard of living the population would be too unhealthy to expand."

In short, numbers below the \$250 bound — even if incomes were equally distributed — would place those countries below extreme nutrition-based poverty lines drawn at 2000Kcal per day. At any lower income, the country would systematically die off, with child and adult mortality rates climbing above unsustainably high levels.

But if that is so, "divergence, big time" is the only reasonable conclusion over this period. Per-capita income in the United States grew at a rate of 1.7% per annum over 1870–1960, which implies a four-fold climb over that 90-year period (apply your trusty doubling-time formula). But 42 out of 125 countries in the Penn World Tables (at the time) had per-capita income below \$1,000 in 1960. *They could not have grown four-fold* or more, for that would violate the biological lower bound of \$250 that we just discussed. That's a prima-facie case for divergence. Pritchett provides estimates of the extent of divergence, by extrapolating the incomes of these 42 countries backward, maintaining their relative rankings, and calibrating the extrapolation so that the poorest country in the sample just hits the \$250 lower bound in 1870; see Pritchett (1997) footnote 11 for more details. With these simulated numbers in hand, the extent of divergence can be further quantified; see Pritchett (1997, Table 2).

And yet, like all good stories, there is a twist in the tale, which we'll get to soon.

4.4. Interlude: Conditional Convergence

For now, let's see what's driving the divergence. Is it the lack of convergence even after conditioning for parametric differences, or is it that the parameters themselves are persistently different across countries? Figure 4.2 illustrated the concept of conditional convergence. A country that starts at point C is actually poorer than the country that starts at point C', but it also grows more slowly than its initially richer counterpart. A country that is below its own steady state is indeed predicted to grow faster than *its* steady-state growth rate, but to test this we have to also use the data to identify *where* those steady states are. (Unconditional convergence skips this step because it asserts that all the steady states "are in the same place" to begin with.)

4.4.1. Conditioning in the Solow Model.

Recall equation (3.12) from Chapter 3, Section 3.4.2, which describes steady state output relative to *effective* units of labor in the Cobb-Douglas case, and is reproduced here for easy reference:

$$\hat{y}^* \simeq A^{1/(1-a)} \left[\frac{s}{n+\pi+\delta}\right]^{a/(1-a)}. \tag{4.3}$$

We can express (4.3) in logarithmic form, so that

$$\ln \hat{y}^* \simeq \frac{1}{1-a} \ln A + \frac{a}{1-a} \ln s - \frac{a}{1-a} \ln(n+\pi+\delta). \tag{4.4}$$

Next, recall the idea of β-convergence from Section 4.2.1, which states that the growth rate of income in effective labor units is connected to the divergence of that income from its steady state value (in logarithms). See Appendix A for a precise derivation:

$$\hat{g}(t) = \beta[\ln \hat{y}(t) - \ln \hat{y}^*]. \tag{4.5}$$

In terms of per-capita income and its growth, we know that $\ln y(t) = \ln \hat{y}(t) + t \ln(1+\pi)$ and $g(t) \simeq \hat{g}(t) + \pi$, so we can combine equations (4.4) and (4.5) to get

$$g(t) = \alpha(t) + \beta \ln y(t) + \gamma_1 \ln s + \gamma_2 \ln(n + \pi + \delta), \tag{4.6}$$

where we've put the terms not involving s or n or y, including technical progress π which we continue to assume is the same across countries, into the "intercept term" $\alpha(t)$, and kept the core parameters s and n as controls.

This is one instance of how to derive a specific expression of the growth regression equation (4.2). We could go beyond it by invoking Chapter 3.5.2, in which we describe steady state income as a function of *three* parameters: the savings rates s_k and s_h in physical and human capital respectively, as well as the population growth rate n. See equation (3.24) for details. That equation, logged, would replace equation (4.4) as an extended description of the steady state in effective labor units:

$$\ln \hat{y}^* \simeq \frac{1}{1-a-c} A + \frac{a}{1-a-c} \ln s_k + \frac{c}{1-a-c} \ln s_h - \frac{a+c}{1-a-c} \ln(n + \pi + \delta), \tag{4.7}$$

where we've set δ_h and δ_k to the same value δ for simplicity. In exactly the way that we combined (4.4) and (4.5) earlier, we would now combine (4.7) and (4.5) to obtain another instance of the growth regression equation (4.2):

$$g_i(t) = \alpha(t) + \beta \ln y_i(t) + \gamma_1 \ln s_{ik} + \gamma_2 \ln(n_i + \pi + \delta) + \gamma_3 \ln s_{ih} + \epsilon_i(t). \tag{4.8}$$

This is how the basic conditioning exercises work. Now we take (4.8) to the data.

4.4.2. Conditional Convergence in the Data. It is hard to conduct a conditional convergence exercise over 1870–1979, as values of our parameters will typically be missing or unreliable. But doing this for more recent time periods is not a problem at all. We consider the period 1960–2019, and divide up this overall interval into three sub-intervals: 1960–1980, 1980–2000, and 2000-2019. For each of these three sub-periods we regress overall growth on baseline log income, using controls for savings rates and population growth rates at the start of each sub-period.

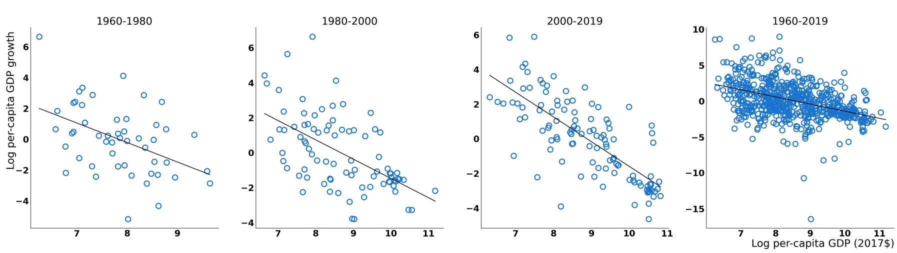


Figure 4.7. Residuals from the regressions of Table 4.1 after conditioning for covariates, shown against log per-capita income at baseline. For more information, see the caption to Table 4.1.

The results are interesting. We've already seen that there is no evidence of unconditional convergence over this period, with the exception of 2000–2019 to which we'll come presently. But Figure 4.7 and Table 4.1 display substantial evidence that the conditioning exercise "works". That is, controlling for just three variables in the data — the savings rate for physical capital formation, the savings rate for human capital, and the population growth rate — what's left over (or the residuals of the regression of

	Per-Capita Income Growth (Annualized Percentage)			
	1960–1980	1980–2000	2000–2019	Pooled
Constant	**16.46	2.013	***6.521	2.359
	(6.785)	(4.361)	(2.299)	(1.951)
LogGDPpc	**-1.267	***-1.104	***-1.426	***-1.008
	(0.5366)	(0.2464)	(0.1897)	(0.1189)
Log(Investment Rate)	0.3376	***1.854	**1.185	***0.9254
	(0.5194)	(0.6527)	(0.5014)	(0.2448)
Log(Ave. Years of Edu.)	***1.542	***1.651	***1.873	***1.124
	(0.5083)	(0.5004)	(0.3815)	(0.1788)
Log(0.05 + Pop. Growth)	1.669	***-3.566	***-2.698	***-3.121
	(1.731)	(1.342)	(0.8336)	(0.7048)
Observations	50	84	112	645
R^2	0.20545	0.38204	0.41048	0.16945
Adjusted R^2	0.13482	0.35075	0.38844	0.16426

Table 4.1. Conditional convergence, 1960–2019. Columns 1–3 regress twenty-year growth in per-capita GDP (Δ logGDPpc, annualized, in %) on initial log GDP per capita and other covariates in baseline years 1960, 1980 and 2000. Column 4 does the same in a pooled regression for baselines varying all the way from 1960–2000. Heteroskedasticity-robust standard-errors in parentheses. Significance codes are for 1% (***), 5% (**) and 10% (**8).
Data sources: Penn World Tables, World Development Indicators, Barro-Lee dataset.

per-capita income that are left "unexplained" by these three variables) does appear to converge across countries. Figure 4.7 shows how these residuals vary with the log of per-capita baseline income, and indeed, the relationship is strongly negative.

For each of the four panels in this Figure, the four columns of Table 4.1 report a corresponding regression, built with equation (4.8) firmly in mind. The first three panels report on 20-year growth periods starting from baseline values of per-capita income and the parameters from the extended Solow model. The last panel pools all these years between 1960–2019 in the same regression, and looks at 20-year growth rates starting from every year up to 2000. This is a strong set of relationships that survives obvious tweaks and variations, such as some time-averaging of baseline incomes, different time periods starting from that income, the use of contemporaneous and time-varying parameter controls, and so on.[41]

These results are also robust to a related way of thinking about conditional convergence, distinct enough that we can't call it a tweak or simple variation. The growth regression (4.8) presumes that countries are on their transition paths to their steady states, which is exactly in the spirit of testing for unconditional convergence (without controls), An alternative philosophy might rest — in the words of Mankiw, Romer and Weil (1992) — on "taking Robert Solow seriously" enough to presume that countries are *already* at their steady states, and that all ongoing per-capita growth is due to technical progress. That presumption yields a new steady state regression equation, an analogue

[41]For sibling regressions on which ours is based, see Mankiw, Romer and Weil (1992, Tables IV and V) and Barro and Sala-i-Martin (2004, 12.2 and 12.3). Barro and Sala-i-Martin's regressions contains a larger number of controls — tens of them in fact — but it's interesting that just three parameters serve to precipitate conditional convergence.

of (4.8), that can also be taken to the data, with similar findings. Appendix B describes the Mankiw-Romer-Weil approach.

How fast is the pace of conditional convergence? Our regressions can be conveniently interpreted in this respect. The coefficient on baseline per-capita income (Row 2 in Table 4.1) can be viewed as the annual percentage rate of convergence — of excess growth by relatively poorer countries. Using our convenient doubling-time formula, a coefficient of 1.2 then means that per-capita income differences between poor and rich countries will be halved every 60 years or so at the current pace of conditional convergence. That's positive, but not dramatic by any means,[42] especially considering that we are already conditioning for various parameters.

There is an amusing postscript to this exercise, perhaps to be taken with a pinch of salt. Look at the regressions in Table 4.1 for 1980–2000, 2000–2019 and the pooled version in the final column. In each of these, the sum of the regression coefficients on the investment rate and on rate of accumulation of human capital[43] is not far in absolute value from the coefficient on population growth. That equality in the sum is indeed a prediction of the growth regression equation (4.8), which is interesting even to a skeptic like me (when it comes to taking theories literally). In fact, we can go a bit further. The derivation of (4.8) from its immediate predecessor (4.7) shows that

$$\frac{a}{1-a-c} = -\frac{\gamma_1}{\beta_1} = -\frac{\text{coefficient on } \ln s_k}{\text{coefficient on } \ln y} \simeq 1.7,$$

for 1980–2000, with corresponding estimates of $\frac{c}{1-a-c} = 1.5$. These imply values for physical and human capital shares in national income of around 42% and 37%. The estimates for 2000-2019 are $a = 26\%$ and $c = 41\%$, while those for the pooled regression are $a = 30\%$ and $c = 28\%$. These vary from regression to regression, and each regression is extremely parsimonious, but they are not wildly off from acceptable estimates of the shares of physical and human capital in national income, which are about a third each (see, e.g., Mankiw, Romer and Weil 1992).

We are presently going to criticize conditional convergence for its cavalier insistence that objects such as savings rates are just, well, conditioning devices. They are deeply endogenous to the entire process and conditioning on them throws away plenty of useful information. That said, let's give conditional convergence its due. In the spirit of showing us useful correlations, it does a lot. It tells us that controlling for just three parameters gives us a convergence rate across countries that's not to be scoffed at. It is a hopeful exercise, but of course, there's an immense amount to be uncovered that drives the process, not to mention the differences in the parameters that we've controlled for.

4.4.3. Limits to Conditional Convergence. Conditional convergence can be a confusing concept. It is entirely consistent with divergence overall. There are deep social and behavioral roots that drive savings, educational investments, and reproductive decisions. Moreover, these three "Solow parameters" are just the start of a long list of religious, legal, political, cultural and institutional factors. Some of these

[42] Larger estimates are reported in Barro and Sala-i-Martin (2014) but they also contribute for a far larger of covariates.

[43] At first glance, the log of average years of education does not look like a "rate" at all. But average years of education *divided* by working life is indeed a rate, and taking logarithms yields the specification in Table 4.1, provided we presume that working life is roughly the same across countries.

inter-country differences (such as geography and climate) are possibly immutable — though technological innovations (e.g., in transportation) and induced climate change could cause them to effectively alter. Certainly, historical experience is immutable by definition. But barring such objects, other so-called "parameters" are genuinely endogenous outcomes of the development process.

So it's important to not go overboard with conditional convergence. When we include all these so-called parameters in a regression inspired by equation (4.8), we're effectively saying: let's control for them and see what happens "after that." To understand where the world as a whole is going, we must also train our eye upon these so-called parameters, and attempt to understand how these evolve. We owe it to ourselves to understand just *why* savings rates or population growth rates or technological know-how vary across countries, and not just control for those differences. We could continue conditioning till the cows come home — not just on the usual Solow parameters, but on norms, religions, institutions, laws, political systems, or even country fixed effects to take care of all time-invariant differences. But then, after a point, the term "convergence" — weakened by the need to condition for this or control for that — is just an inappropriate word. Perhaps, in the words of Pritchett (1997), *divergence, big time* is what it's all about. Perhaps, left to their own devices, economic forces separate rather than unite us. While conditional convergence is a good organizational device, it places insufficient emphasis on the process:

$$\text{“parameters"} \longrightarrow \text{outcomes} \longrightarrow \text{“parameters"} \longrightarrow \text{outcomes} \longrightarrow \ldots$$

ad infinitum, where I've placed "parameters" in quotes to get you to stop thinking of them as exogenous objects. This intertwining is central to the book.

4.4.4. Two Examples. As examples, one might ask: what about the reverse effect of income on savings, or on population growth rates? What might the explicit consideration of both directions of causality teach us?

Consider the savings rate. If your income is just enough to pay for the bare necessities of life, you won't be saving much. I don't mean that your *absolute* savings will be low, which is obviously true, but that your savings *relative to your income* is likely to be low as well. (Indeed, you might even be borrowing to make ends meet.) This observation extends to countries as well as individuals: poor economies are unlikely to display a high savings rate. As the economy grows, leaving subsistence levels behind, there is increased room for postponing current consumption, and savings rates rise. The rise may be additionally fueled by aspirations that begin to look attainable: savings rates can peak for the middle class, and by extension, in middle-income countries. This sharp rise in the savings rate could flatten out at even higher levels of income. Although the rich (and rich countries) can *afford* to save, the fact that they are ahead of many other individuals (or other countries)

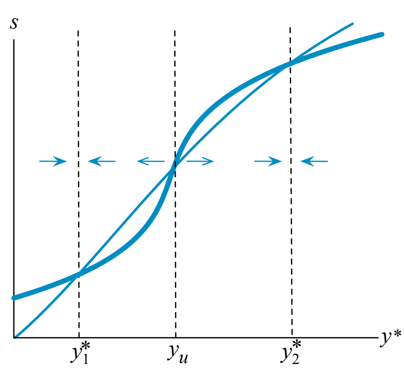


Figure 4.8. Endogeneity of the savings rate.

might blunt their need to accumulate wealth for their progeny, and consume more in the here-and-now. Figure 4.8 illustrates.

The thick curve describes how savings rates vary with income, starting off small and flat until some threshold is crossed, upon which the rate starts to increase but later flattens out again. The thin curve in the figure plots various combinations of savings rates and steady state levels of per-capita income, as described by the basic Solow model in Chapter 3, equation (3.8). The intersections of the two curves describe steady states of another “extended model," in which both the savings rate and the associated level of steady state income are jointly determined.

Figure 4.9 tells a similar story when population growth rates are endogenous. As Chapter 14 will discusses in detail, birth and death rates (and therefore net population growth) systematically vary with economic development. In poor countries, death rates are high. The greater incidence of famine, undernutrition, and disease, as well as difficult conditions of sanitation and hygiene, all contribute to this outcome. Birth rates are consequently high as well: families must procreate at a greater rate to reach some target number of surviving offspring. The combination of a high birth rate and a high death rate can mean that *net* population growth rate could be both high or low. With an increase in living standards, death rates start to fall, but for various reasons (see Chapter 14), birth rates adjust relatively slowly to this transformation in death rates. This causes the population growth rate to initially shoot up. The increase is all the more dramatic if the decline in death rates is rapid. In the longer run, and with further development, birth rates begin their downward adjustment and the population growth rate falls to a low level once again.

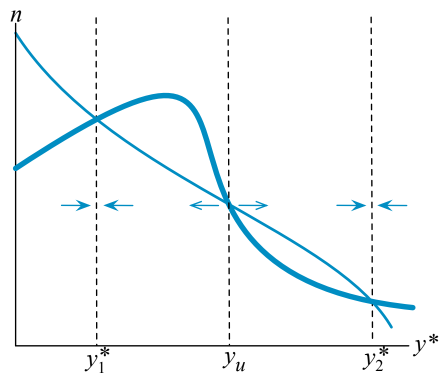


Figure 4.9. Endogeneity of population growth.

If this compressed description has raised several questions in your mind, don’t worry; it should. As promised, we will return to these issues later in the book. For now, concentrate on the implications of this “demographic transition" for per capita economic growth. The thick line in Figure 4.9 depicts demographic reactions: population growth rates first rise and then fall with per-capita income. The thin line connects population growth to steady state income as given by the Solow model.

Now we analyze both examples together. Notice that at the point y_1^*, the savings rate (Figure 4.8) or the population growth rate (Figure 4.9) is just right to regenerate y_1^* as a steady state, so the whole system hangs together and propagates itself. The value y_1^* is “locally stable" as well. Look at Figure 4.8 for instance. If there were to be a little push to the “right" of y_1^*, say, because of a temporary infusion of foreign aid, the higher savings rate that such a push would generate (follow the thick line) would *not* be enough to keep the economy at that higher level: over time, incomes would fall back to y_1^*. The same argument holds in reverse for a little leftward nudge from y_1^*.[44]

The same is true of the highest per-capita steady state y_2^*. It, too, is impervious to little nudges to left and right. And yet: from y_1^* there seems to be no way of getting to y_2^*, at least in small, incremental steps. The lower steady state y_1^* is a toy representation of a *development trap*: once in it, it is hard to escape its gravitational attraction.

Now verify that the same features are true of Figure 4.9 . To the right of y_1^*, there is a sharp increase in net population growth, which pushes per-capita incomes back down. One might think of y_1^* as a *Malthusian trap*, in (dubious) honor of the Reverend Thomas Malthus who warned that income improvements would be eroded in a rapid and presumably sinful orgy of reproduction. Similarly, y_2^* represents another stable point with low demographic change and a high level of per-capita income.

In both Figures, the bridge from y_1^* to y_2^* is broken by a *third* steady state — marked y_u, "u" for "unstable." It too is a point at which savings and per-capita income (or population growth and per-capita income) "hang together," but it is an unstable point, just as a pencil balanced precariously on its end is unstable, vulnerable to a small nudge. To the right of y_u, the economy enters into a virtuous cycle as savings rates rise (Figure 4.8) or population growth rates fall (Figure 4.9) to propel the economy to a permanently higher steady state at y_2^*. But to the left of y_u, just the opposite is true: the system careens towards the trap set at y_1^*. The unstable steady state y_u is like a dividing line which demarcates the two basins of attraction: one for the development trap at y_1^* and one for the enhanced steady state at y_2^*.

These examples are fully consistent with conditional convergence. But across two countries in the two different stable states, there is a failure of *un*conditional convergence. The two economies may be identical in the sense of having exactly the same structure, as described by the thick and thin lines in Figures 4.8 and 4.9. But they may exhibit profoundly different non-convergent outcomes. Such simple instances (some might say simplistic) are of course caricatures of reality. Some more realistic extensions are hidden in this footnote.[45] But they are not bad caricatures For instance, they capture well the possibility that the past can weight heavily on the present. It even hints at why the world's mobility matrix looks as it does in Chapter 2, sticky at the ends, slippery in the middle. Our examples have exactly this feature: the "middle" (around y_u) is unstable, while the "ends" (around y_1^* and y_2^*) are stable.

---

[44] In the face of such a nudge, the savings rate would fall, but not by enough to create a new self-justifying steady state: the system would work itself back up to y_1^*.

[45] For instance, there is no theoretical *necessity* that these curves must cut multiple times, or just three times even if they do so multiple times (though our accompanying descriptions do suggest that a "three-cut" isn't unreasonable). For instance, the population curve may not quite cut the thin line in panel B around y_1^*, but pass below it. Then the Malthusian trap would disappear, leaving only the salubrious steady state y_2^*. Nevertheless, as population growth peaks, the economy will pass through a prolonged period of slow growth (instead of slowing to a complete halt), and the ideas that we discussed in this section remain just as valid, although not in equally stark form. Among other simplifications, our population growth curve neglects the notion of demographic transition as a process in time. For instance, as per capita income *declines*, is it true that population growth reverts to the levels that it had seen earlier? Our diagram assumes that the answer is literally "yes," and this is certainly an exaggeration. Modeling such irreversibilities appropriately is certainly important, but they will not detract from the main points of what we have to say, and this is what makes a model capable of being relevant, even though it may be unrealistic.

4.4.5. Path Dependence. The arguments just made suggest a view of economic development that is powerful and provocative, and possibly correct in many situations. An economic outcome (say per-capita income) might influence a so-called parameter (savings rate, population growth rate, the distribution of asset ownership, the choice of technology), which then feeds back on that same economic outcome in a mutually interlocking circle. That interlocking could happen at different configurations even if two societies are governed by the same causal relationships — the same DNA, so to speak. The existence of such multiple outcomes hints at the possibility that economic development could be highly *path-dependent*. That is, modern societal arrangements could be influenced by historical outcomes, or even historically-held expectations regarding current outcomes. Some of these are relatively innocuous; e.g., driving on the left or the right side of the street, though it *is* important to coordinate on one of these arrangements so as to avoid havoc. But there are many other arrangements that could both interlock in the way described above, *and* significantly impact economic performance. In later chapters we return to this theme.

4.5. Act II: Convergence? 1980–2020

4.5.1. A New Trend. Figure 4.10 repeats the cross-country graph from Figure 4.6 for the more recent time interval 1960–2019. The picture looks quite bland, just as Figure 4.6 did, with no readily discernible pattern in any direction. But something remarkable does emerge if we break this period up into sub-intervals. Figure 4.11 does just that, by dividing the 60-year period between 1960–2019 into three roughly equal sub-periods: 1960–1980, 1980–2000, and 2000–2019. The first sub-interval continues to show some evidence of divergence, while the second appears to exhibit "neutrality": initial levels of income do not appear to map in any distinct way on to their subsequent rates of growth. The real change is in the sub-interval that spans 2000-2019. Higher initial incomes appear to be quite tightly correlated with lower economic growth per-capita. This is a distinctly 21st century phenomenon, which you and I will follow as the years go by — but we may well be on the cusp of something different.

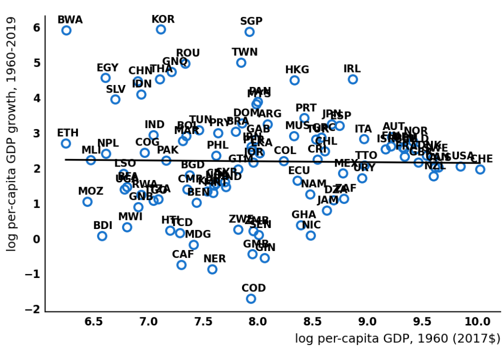


Figure 4.10. Growth and per-capita income, 1960–2019. Source: Penn World Tables Version 10.01.

Table 4.2 follows up by running the β-convergence regression in equation (4.1)

$$g_i = \alpha + \beta \ln y_i + \epsilon_i,$$

and tests for unconditional convergence by looking for a negative value of β. The table reports values of β for all the time periods considered in this chapter. The numbers are generally insignificant, except for 2000-2019, during which the coefficient turns sharply negative and significant. Is this the very first golden age of unconditional convergence?

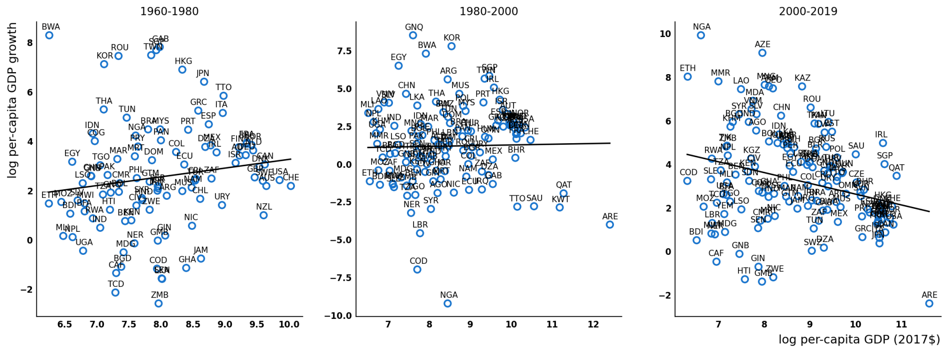


Figure 4.11. Growth and per-capita income in 20-year intervals over 1960–2019. Source: Penn World Tables Version 10.01.

	Growth Rate of GDP per capita				
	1870–1979	1960–2019	1960–1980	1980–2000	2000–2019
Coefficient β	-0.068	-0.025	0.354	0.048	-0.495
	(0.110)	(0.162)	(0.241)	(0.205)	(0.139)

Table 4.2. Regression coefficient β of per-capita GDP growth, measured as $100 \times [\ln(y(t) - \ln(y(s))]/(t-s)$ over different time periods $[s, t]$, on baseline log GDP per-capita $y(s)$. Standard errors in parentheses. Sources: Maddison database, Penn World Tables 10.01.

There is descriptive evidence in support of that viewpoint. We have already noted the dramatic growth in East and Southeast Asian countries in the last several decades. We know that developed countries on the frontier have been shaken by a series of recessions and slowdowns, notably spearheaded by the Financial Crisis of 2007-2008. But it isn't so much slowdown but distinct catchup. We have seen some African countries such as Rwanda begin a sustained take-off into economic growth. We have seen recovery in Latin America after their lost decades starting in the 1980s, benefiting particularly from a boom in the prices of fuels, chemicals and metals after 2000. And this boom has been driven in part by quite extraordinary rates of growth from China and more recently from India. It appears that all of that is beginning to add up. Note that that one country counts as one observation, so that the huge population sizes in India and China are not driving this new and intriguing tendency. It is more widespread. Roy, Kessler, and Subramanian (2016), Johnson and Papageorgiou (2020), Kremer, Willis and You (2021), Patel, Sandefur, and Subramanian (2021) and several others take note of this recent proclivity, with reactions ranging from careful skepticism to guarded positivity to more unbridled enthusiasm. My own reaction would fall into the guarded positivity camp. After all, this is a recent trend, and the global economic stage is incredibly complex.

However, even with the need for caution duly respected, our analysis so far can help make some sense of what might be happening. The next two sections consider two different takes on this theme. Both also serve as entry points into the rest of our book.

4.5.2. Parametric Convergence. We have already seen excellent evidence of conditional convergence, at least when controlling for investments in physical and human capital. We've noted that conditional convergence to this degree can sit side by side with unconditional *divergence* if the "parameters" that govern the Solow model, or some enriching of that model, remain persistently different across countries. The failure of unconditional convergence could in part be a failure of parametric convergence.

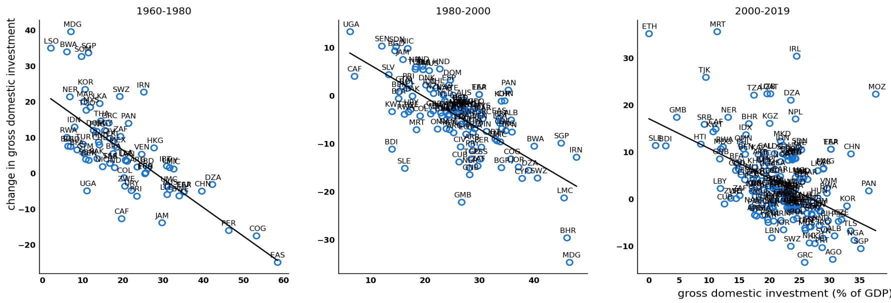
Turning that last sentence around, one might therefore expect to see some evidence of parametric convergence in the recent past. Of course, as already noted, "parameters" abound, but let's consider the variables that undergird the basic and augmented Solow models: investment (and savings) rates,[46] population growth rates, and educational enrollment. Figure 4.12 depicts the *rate of change* in these objects (over 1960–2019) as a function of their initial levels in 1960. For instance, Panel A shows the rate of growth of the rate of investment as a function of the initial rate of investment in 1985. The strong negative slope suggests that the rate of investment appears to be converging across countries. The same appears to be true of population growth rates and educational enrollment rates. Of course, we must always temper these observations with the possibility of measurement error at baseline as well as statistical mean reversion, both of which create an illusion of convergence where there may be none. But in these diagrams, all baseline years lie in fairly modern times (certainly the last panels of each do), so there is no a priori reason to suspect that measurement error plays more than a cursory role. Likewise, our choice of two-decade averaging is likely to swamp any kind of statistical regression to the mean, which would occur on smaller time scales. Thus in summary, we do seem to have parametric convergence at least for the small set of variables that make up the extended Solow model. In the words of Kremer, Willis and You (2021), we appear to be "converging to convergence."

There is, of course, an entire host of parameters — institutional, cultural, political, legal, religious — that are slow and sluggish to adjust. Perhaps these are more deserving of the title "parameter," being a bit more impervious to changes in the here-and-now. History weighs heavily on such variables. They are too many and too varied to all be captured in the net of our simple regression equations (see Kremer, Willis and You 2021 for more on such parameters), but they too are capable of endogenous change.

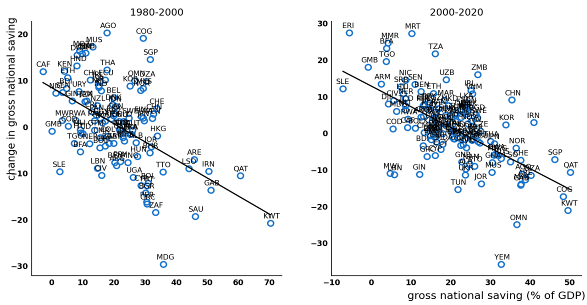
There is room for cautious optimism about these slow-moving parameters, because the operative word on the world stage today is *interconnectedness*. Simply put, there is today an unprecedented degree of global communication, spurred on by fundamental changes in technology as well as increasing globalization in economic activity. We are all ever more aware of our local foibles and deficiencies *in relation to global best practice*, and as policymakers, activists and governments and individual economic actors, those of us below the knowledge frontier are beginning to make a beeline for that frontier. However, these optimistic words must be tempered on two counts.

First, to make a run for the frontier a country has to be reasonably within striking distance of that frontier to begin with, a property that can be reasonably attributed to many *but not all* developing countries. While global interconnectedness can only help in this regard, it seems incumbent on governments to help bridge this knowledge gap. Second, neither the feasibility of global best practice nor the awareness of such

[46] In practice, these are separate measures. We know that savings and investment rates are not necessarily equal, given the existence of international capital flows.

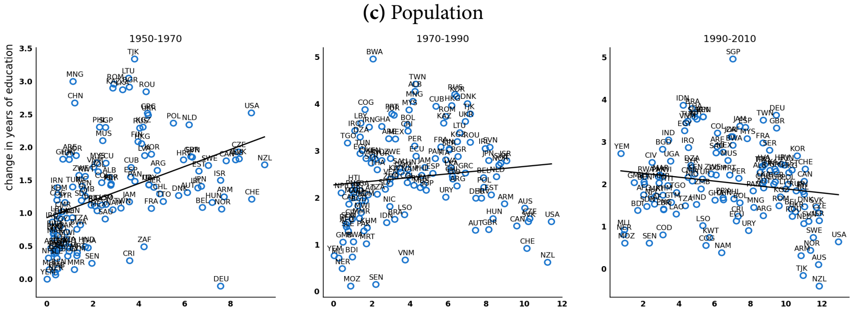


(a) Investment



(b) Savings

(c) Population



(d) Education

Figure 4.12. Parametric convergence for rates of investment (1960–2019), savings (1980–2020), population growth (1961-2019) and years of education (1950–2010). Sources: World Development Indicators, IMF World Economic Outlook Database, Barro-Lee Dataset.

practice guarantees a move towards "better parameters," if the earlier configuration is locked in by self-reinforcing relationships, as in the examples of Section 4.4.4 for savings and population growth rates. But what creates lock-in for some parameters and not others? Empirical patterns of convergence and divergence in incomes and parameters are not so much an explanation but a study of patterns, which can at best hint at something deeper at work. The Invisible Hand works well on some fronts, not all. To uncover these fronts is a central and ambitious task — perhaps too ambitious — but at least the questions that lie ahead are clear.

4.5.3. Within-Sector Convergence. A particular failing of the approach so far is that it is painfully aggregative. There is a focus on big-picture objects but less so on the more granular details, such as the study of different economic sectors, or questions of distribution, or the considerations of microeconomic forces that act on individual actors and may or may not create lock-in. All these will receive their share of attention in the chapters to come, but I do want to think here about some disaggregation into sectors, continuing this theme in far more detail in Chapter 6. As we've seen in Chapter 2, economies are composed — to greater and lesser degree — of at least three broad groups of output: agriculture, manufacturing and services. The box on "Sectoral Convergence" argues that we see evidence of convergence within manufacturing and services, though not in agriculture. Then the emergence of unconditional convergence *overall* is a sign that the intersectoral composition of production across economies is beginning to settle down. Like the investment rate or the population growth rate, this intersectoral composition could also be called a "parameter," but it is just as endogenous to economic development and has evolved with it.

Convergence Within Sectors

The unconditional convergence hypothesis can be recast in terms of *labor productivity* rather than per-capita income. That allows us to apply the hypothesis to sectors within the economy, rather than the economy as a whole. It would then read as follows: countries with lower initial labor productivity levels will experience faster growth rates of those levels. Such a statement can be made and assessed for different sectors within the economy. Dani Rodrik (2013) does just this for the manufacturing sector. In his words:

> "I show in this article that unconditional convergence does exist, but it occurs in the modern parts of the economy rather than the economy as a whole. In particular, I document a highly robust tendency toward convergence in labor productivity in manufacturing activities, regardless of geography, policies, or other country-level influences."

Rodrik uses cross-country data from the United Nations Industrial Development Organization to study growth rates of manufacturing productivity across 118 countries between 1965 and 2005. Under his baseline specification (see Table I, Column 1), the coefficient of unconditional convergence is 0.029, which means that a country which is at a tenth of the highest observed productivity will grow $0.029 \times \ln(10) \times 100$ percentage points faster every year. That translates to a sizable 6.7 percentage points per year. Best-practice technologies do diffuse — and they appear to diffuse quickly — in manufacturing.

Good complements to Rodrik's work are earlier contributions by Freeman and Yerger (2001), who study convergence in manufacturing labor productivity prior to 1970, but only among OECD countries, and Bénétrix, O'Rourke, and Williamson (2012), who study convergence in manufacturing (as measured by per-capita industrial output) over 1890–1972. across Latin America, the European periphery, the Middle East and North Africa, Asia, and sub-Saharan Africa. They document that less industrialized countries displayed higher per-capita manufacturing growth rates between 1920 and 1990. Yes, unconditional convergence — *conditional* on being in manufacturing to begin with — has been around for over a century.

In similar spirit, Kinfemichael and Morshed (2019) study the service sector for 95 countries over 1975–2012, using data from the United Nations Statistics Division and the International Labor Organization Statistics database (ILOSTAT). Following Rodrik (2013), they find evidence for unconditional convergence for the service sector: countries with low labor productivity experience faster labor productivity growth than their higher-productivity counterparts. This is true even of 11 out of 12 subsectors within the service sector. Under their baseline specification (see Table I, Column 1), the coefficient of unconditional convergence is 0.035, which means that a country which is at a tenth of the highest observed productivity will grow $0.35 \times \ln(10) \times 100 = 8$ percentage points faster every year.

There is little sign of unconditional convergence in agriculture (Kinfemichael and Morshed 2019, Dieppe and Matsuoka 2021), but even if there were, it would be of little relevance. Rather, what matters is the immense productivity gap *across* agriculture and other sectors in the economy, stemming from the preponderance of small-scale farms and a high labor-land ratio in many developing countries. It isn't at all hard to imagine that a major reason for divergence across countries in the greater part of the 20th century stems from differential rates of industrialization. As the great structural transformation from agriculture to industry proceeds, divergence is a likely outcome if the rate of that transformation varies across countries, with some countries leading the charge and other countries following behind. That divergence is independent of — and could happily coexist with — unconditional convergence *within every sector*. But, as all countries begin to complete their structural transformations, the forces of convergence within these latter sectors would begin to permeate the economy as a whole, causing the tilt towards unconditional convergence that we have begun to see. The need to look at these sectors through a finer lens becomes all the more apparent.

Within-sector convergence suggests that much of the divergence that we have seen historically could well be due to different rates of *structural transformation*. One such great transformation is the move from an agriculture-based economy to an economy with well-developed industrial and services sectors. Convergence *within* each of these sectors is no guarantee of convergence overall, because economies go through its structural transformation at different moments in history. (Just *why* structural transformation must be staggered in this way is a good question, but let's sidestep that one for now.) In doing so, they may well diverge from one another even though convergence — or even perfect equality just for the sake of argument — is a characteristic of each sector taken in isolation. Simon Kuznets, whose work we shall have occasion to return to, suggested that economic inequality rises and falls within countries as its population begins the long trek towards "modernization". The same observation, applied to countries as a whole as they make their staggered treks, could explain rising inequality (or divergence) across entire economies. Then, as those journeys are completed one by one, the within-sector forces reassert themselves, and there is convergence.

Just as in the case of parametric convergence, this is only a suggestive story, and its suggestions don't by any means extend to the claim that a golden age of convergence is irreversibly upon us. Economic development is not one structural transformation — it is an entire series of them (some overlapping). Each of them comes with its initial forces of disequalization, as some lucky or farsighted first-movers enter the new sector. They also make for later convergence, as the remainder climb on the wagon of that transformation. But pretty soon, another new upheaval could well shake things up, reigniting the twin forces of disequalization and subsequent catch-up. Economic development is an uneven process.

4.6. Summary

Robert Solow's idea that per-capita growth ultimately settles down to equal the rate of technical progress led us to the idea of *convergence*. Under an extreme version of this concept, known as *unconditional convergence*, relative income differences between countries must die away in the long run, with all countries honing in on a common growth path. This strong hypothesis rests on strong assumptions — that the underlying features of relevance, such as savings rates, demography, and technical change, are largely equal across countries. But these are only the parameters that govern the simple growth model as we know it. In more expansive form, there is a whole host of parameters that govern economic growth. They would all have to be in line — or nullify each other in some serendipitous way — for this hypothesis to work.

A weaker version of the convergence hypothesis, called *conditional convergence*, does away with this assumed sameness, and asserts that *controlling* for possible differences in cross-country parameters, initially poor countries do grow faster.

Convergence in either flavor is intimately connected to the notion of diminishing marginal productivity of capital, or accumulable inputs more broadly defined (including human capital). The idea is that a poorer country has a higher *marginal* return to those inpus and therefore exhibits a higher rate of per-capita growth.

It is, of course, possible that the background parameters are not the same, but they too are in the process of converging across countries just as per-capita income is hypothesized to do. We called this *parametric convergence*. We argued in this chapter that unconditional convergence can be viewed as the child of conditional convergence and parametric convergence. Unconditional convergence would be observed if there were a reasonable empirical tendency for both the latter forms of convergence to occur.

We then examined these concepts empirically. Time periods stretching back to the nineteenth century aren't easy to handle. The bias comes from the fact that countries which are rich today are more likely to have data going back that far, and if we restrict ourselves to just those countries, the resulting analysis is biased in favor of convergence. Lant Pritchett (1997) summarizes this bias well:

> "[T]he sample of countries for which historical economic data exists ...is severely nonrepresentative...Defining the set of countries as those that are the richest now almost guarantees the finding of historical convergence, as either countries are rich now and were rich historically, in which case they all have had roughly the same growth rate (like nearly

> all of Europe) or countries are rich now and were poor historically (like Japan) and hence grew faster and show convergence."

Some decades have elapsed since Pritchett made these observations. Thanks to initiatives like the Maddison Project, we now have historical data going back to the 19th century for more countries. Yet there is simply no tendency for unconditional convergence over 1870–1979 even if we consider this larger set of countries, despite the fact that substantial selection bias still remains. It is fair to say that over the late 19th century and most of the 20th century, unconditional convergence is dead in the water.

The weaker notion of conditional convergence is more eclectic and more subtle. It allows parameters to vary across countries, so that in principle, they all have different steady states. The link between high initial income and low subsequent growth is thereby broken. Rather, conditional convergence invites us to to allow for, say, possibly different rates of savings and population growth. We developed that conditioning exercise by deriving a regression model from the basic Solow theory, augmented by the human capital extension from Chapter 3, Section 3.5.2. There are two variants of that exercise, depending on whether we presume countries to be "already" on their steady state path or not. We focused on the steady state exercise.

There is a great deal of evidence for conditional convergence, even if we condition on a relatively small set of parameters. Just controlling for differential savings rates and population growth rates is not quite enough, though: the estimated coefficients for savings and population are too large to be reconciled with the simple Solow model. This is a clear indicator that other "parameters" are either co-moving with savings and demography, or that they are working to amplify their effects. The inclusion of human capital — and then allowing for "savings in human capital" or education — goes a substantial distance towards reconciling the Solow model with the data.

We then took note of the limits to conditional convergence as a useful concept, observing that it essentially removes so-called parameters from our analytical perspective. Such parameters may be endogenous variables, engaged in a two-way dance of causation from their values to GDP, and then back again, resulting in the possibility of multiple steady states even across otherwise functionally identical economies. Indeed, if we make the assumption that human beings are fundamentally the same the world over, in the sense of having the same relevant genetic material — and I have no intention or reason to depart from that assumption — then we are *obliged* to consider these and other explanations for parametric variations across countries. We must probe the differences in savings, demographic change, technical progress, and keep on going: to culture, religion, politics and institutions, rather than simply view them as "controls for convergence," to be conditioned upon but not seriously examined.

We ended with a description of a more recent tendency towards unconditional convergence, and the accompanying convergence of some parameters. This tendency is particularly visible over the period 2000-2020. Now, a lot has happened in the world in this short period — a great recession, a sustained boom in commodity prices, numerous conflicts, the rise of authoritarianism the world over. All these have correlated effects: for instance, it is natural to suppose that developing countries would benefit more from commodity exports, while developed financial systems might bear the initial brunt of a recession driven by a financial crisis. Such phenomena could conspire to create an illusion of recent convergence. The world is just one data point after all, and these upheavals are not isolated, independent events.

But all that said, there is cause for some hope. We have seen some crucial parameters — rates of investment, educational attainments, population growth — steadily converging across countries. We have seen a rapid spread in the *awareness* of best practice both in the economic and the broader institutional spheres — though I emphasize the word "awareness" to underline that that is not necessarily the same as the convergence of best practices themselves. And we've also taken note of a history of convergence in the manufacturing and formal services sector, which suggests that as more countries begin to predominantly focus on such sectors, such tendencies will begin to be reflected in their economies *as a whole*, rather than just in particular sub-areas. These might be chalked up as tentative indicators of hope, at least as far as aggregative indicators at the *cross*-country level are concerned. But there are also many indicators of despair, especially in the political sphere and in the emergence of ugly attitudinal divides, driven by *within*-country inequalities. These are early days yet.

Appendix A: Deriving the Growth Regression Equation

Here is the derivation of the growth regression equation (4.1) and its companion (4.2), which includes covariates. For simplicity, we use Solow model with Cobb-Douglas technology, but the argument can be extended to more general functional forms. The derivation here will require you to know about linear approximations of $f(x)$ around some $f(x^*)$, where x^* and x are close enough so that higher order terms (quadratic in the difference between x^* and x, and still higher) can be neglected:

$$f(x) \simeq f(x^*) + (x - x^*)f'(x^*), \tag{4.9}$$

where $f'(x^*)$ is the derivative of f at x^*. Now to business, which will look messy but in the end is no harder conceptually than the approximation (4.9). First recall our growth equation (3.7), reproduced here to include effective units as well as technical progress π just as described in the main text:

$$\frac{\hat{k}(t+1)}{\hat{k}(t)} \simeq \frac{(1-\delta) + s\Gamma(\hat{k}(t))}{1+n+\pi},$$

where you should remember that $\Gamma(\hat{k})$ is just the output-capital ratio $\hat{y}/\hat{k}$, and the "$\simeq$" reminds you that $1 + n + \pi$ is a close approximation of $(1 + n)(1 + \pi)$. In the Cobb-Douglas case, we know that $\hat{y} = A\hat{k}^a$. Moving terms around, it is obvious that $\hat{k} = (\hat{y}/A)^{1/a}$ and $\Gamma(k) = A^{1/a}\hat{y}^{(a-1)/a}$. Using these in the equation above, we see that

$$\hat{g}(t) \equiv \frac{\hat{y}(t+1) - \hat{y}(t)}{\hat{y}(t)} \simeq \left[\frac{(1-\delta) + sA^{1/a}\hat{y}(t)^{(a-1)/a}}{1+n+\pi}\right]^a - 1. \tag{4.10}$$

where the left hand side is the rate of growth of income in effective labor units, $\hat{g}$. Now recall the formula in (3.12) for the steady state level of effective-labor-unit income, which we reproduce here:

$$\hat{y}^* \simeq A^{1/(1-a)}\left[\frac{s}{n+\pi+\delta}\right]^{a/(1-a)},$$

and substitute into equation (4.10) to obtain (after some boring manipulation):

$$\hat{g}(t) \simeq \left[\frac{(1-\delta) + (n+\pi+\delta)(\hat{y}^* \hat{y}(t))^{(1-a)/a}}{1+n+\pi}\right]^a - 1 \equiv f(\hat{y}(t)), \tag{4.11}$$

where I have put in the f notation on the right hand side of (4.11) to make you think of it *as a function of* $\hat{y}(t)$, with $\hat{y}(t)$ converging of course over time to the steady state $\hat{y}^*$. Using equation (4.9), we approximate f linearly around the value $\hat{y}(t) = \hat{y}^*$. Note that when $\hat{y}(t) = \hat{y}^*$, the right hand side is zero, or in other words, $f(\hat{y}^*) = 0$! Therefore, in this case, using equation (4.9), we can simply write

$$\hat{g}(t) \simeq f(\hat{y}(t)) \simeq [\hat{y}(t) - \hat{y}^*] f'(\hat{y}^*), \tag{4.12}$$

where it should be obvious that $\hat{y}(t)$ is playing the role of x and $\hat{y}^*$ the role of x^* in (4.9). Taking derivatives of f with respect to $\hat{y}(t)$ in (4.11) and evaluating the result at $\hat{y}(t) = \hat{y}^*$, we get the mercifully simple answer

$$f'(\hat{y}^*) = (1-a)(n+\pi+\delta)/\hat{y}^*,$$

which we use in equation (4.12) to obtain

$$\hat{g}(t) \simeq (1-a)(n+\pi+\delta)\frac{\hat{y}(t) - \hat{y}^*}{\hat{y}^*}. \tag{4.13}$$

By a linear approximation, this time applied to the logarithmic function, we know that

$$\ln \hat{y}(t) - \ln \hat{y}^* \simeq [\hat{y}(t) - \hat{y}^*] \frac{d \ln y(t)}{dy(t)}\Big|_{\hat{y}(t)=\hat{y}^*} = \frac{\hat{y}(t) - \hat{y}^*}{\hat{y}^*},$$

and applying this to equation (4.9), we must conclude that

$$\hat{g}(t) \simeq (1-a)(n+\pi+\delta)[\ln \hat{y}(t) - \ln \hat{y}^*], \tag{4.14}$$

We can travel from (4.14) to the growth regression for per-capita income, by noting first that $\hat{g}(t) \simeq g(t) - \pi$, and then recalling that $\hat{y}(t) = y(t)/(1+\pi^t)$ and $\hat{y}^*(t) = S/(1+\pi^t)$, where S, the coefficient on the steady state path, is pinned down by various parameters of the growth model. Putting all this together, we see that

$$g(t) \simeq \alpha(t) + (1-a)(n+\pi+\delta)[\ln y(t) - \ln S], \tag{4.15}$$

where the term $\alpha(t)$ collects time-specific terms containing $t \ln(1+\pi)$. Our growth regression equations directly follow from (4.15), with the β coefficient predicted to be

$$\beta \simeq (1-a)(n+\pi+\delta), \tag{4.16}$$

at least if countries are close to steady state. For the unconditional regression, S is presumed to be the same across all countries and can be subsumed into the intercept term $\alpha(t)$, which leads to the unconditional regression equation (4.1). If S varies across countries, then the additional covariates as written in (4.2) are used to control for this variation. Indeed, in the Cobb-Douglas case, $\ln S$ can be expressed as a linear combination of the main parameters that govern the Solow model: savings rates and population growth rates. If savings in human capital is also included, as in Mankiw, Romer and Weil (1992), then the corresponding savings rate must also enter the list of covariates, and β must be adjusted to reflect the parameters of the three-input production function, and not just two, as in (4.16). This extension yields the regression specification (4.2).

Appendix B: Steady State Regression

There is an alternative to the growth regression equation that we've developed in the main text, which exploits the structure of the Solow model even more literally. That alternative presumes that countries are *already in steady state*, and attempts to "explain" observed variations in income levels across countries by conditioning on the parameters of the Solow model. I introduce you to the basics of such an approach.

Unlike the growth equation, a "steady state regression" equation presumes that countries are close to their steady states. Recall (3.12) in its logarithmic form (4.4), which states that

$$\ln \hat{y}^* \simeq \frac{1}{1-a}\ln A + \frac{a}{1-a}\ln s - \frac{a}{1-a}\ln(n+\pi+\delta). \tag{4.17}$$

Assuming that countries are at the ir individual steady states, we can unwrap $\hat{y}^*$ to recover per-capita income $y(t)$ at any date t:

$$\hat{y}^* = \frac{Y(t)}{L(t)e(t)} = \frac{Y(t)}{L(t)e(0)(1+\pi)^t} = \frac{y(t)}{e(0)(1+\pi)^t},$$

where $e(0)$ is technical knowledge at some baseline date (say 1960 or 1995), and t is counted in terms of years elapsed from the baseline date. Taking logarithms,

$$\ln \hat{y}^* = \ln y(t) - \ln e(0) - t\ln(1+\pi).$$

Substituting this expression into (4.17) and moving terms around, we obtain

$$\ln y(t) \simeq \alpha(t) + \frac{a}{1-a}\ln s - \frac{a}{1-a}\ln(n+\pi+\delta), \tag{4.18}$$

where $\alpha(t)$ is just the collection of terms $\frac{1}{1-a}\ln A + \ln e(0) + t\ln(1+\pi)$. The plan now is to regress $y(t)$ on the parameters exactly along the lines suggested by (4.18). We may call this a *steady state regression* because it relies on the presumption that every country in the regression is essentially on its steady state path, as described by 4.18).

The coefficients on $\ln s$ and $\ln(n+\pi+\delta)$ are unknowns and will be estimated by the best fit to the data. However — and this is the power of a theoretical prediction — the theory suggests that after we are done estimating the equation, the coefficients will be close in absolute value (they are *both* supposed to equal $a/(1-a)$), but of opposite sign. In fact, if we take the estimate of a from the previous section, $a/(1-a)$ should be around 0.5. We may therefore conduct our regression with the following expectations:

- The coefficient on $\ln s$ is positive and the coefficient on $\ln(n+\pi+\delta)$ is negative, but
- The two coefficients should have the same *absolute* magnitude of around 0.5.

Column 1 of Table 4.3 takes up this little challenge; see Mankiw, Romer and Weil (1992) and the subsequent extension by Bernanke and Gürkaynak (2002), whose numbers we report. The average of investment-GDP ratios over 1960–95 is their proxy for the savings rate, and the dependent variable is 1995 per-capita GDP in the year 1995. Close to *two thirds* of the worldwide variation in per-capita GDP in 1995 can be explained by s and n (correlation coefficient = 0.65). This is a powerful finding indeed.

Additionally, and as predicted, the coefficient of $\ln s$ is significant and positive, whereas that of $\ln(n+\pi+\delta)$ is significant and negative. At least in qualitative terms, the Solow model predicts broad relationships that do show up in worldwide data. But there's a bug: the coefficients are too large to be anywhere close to 0.5 (1.11 on

	Log GDP per working-age person, 1995	
	Standard Model	With Human Capital
$\ln s_k$	1.11	0.60
	(0.14)	(0.12)
$\ln(n+\pi+\delta)$	-2.54	-1.81
	(0.50)	(0.36)
$\ln s_h$	-	0.85
		(0.10)
$A(35)$	4.58	7.92
	(1.44)	(1.07)
$\bar{R}^2$	0.65	0.83

Table 4.3. Regressions of log GDP per working-age person, 1995, on baseline variables in 1960. s_k is measured by the ratio of investment to GDP, s_h by fraction of working-age population in secondary school, $\pi + \delta$ is taken to equal 0.05, $A(35)$ is the constant term for $t = 35$. Standard errors in parentheses. Source: Bernanke and Gürkaynak (2002, Tables 2 and 4), extending Mankiw, Romer and Weil (1992, Tables I and II). Sample: 72 non-oil countries from Penn World Tables 6.0 with population > 1m and data quality > D by Summers and Heston (1988).

savings and −2.54 on population). And they're far from being of similar magnitude, as predicted by the basic model. Population growth rates seem to generate a larger depressive effect on per-capita incomes than the upward kick from savings rates.

Yes, countries with larger savings rates and lower population growth rates have higher incomes. The Solow model nails this well. But it's too easy a nail — you don't need a model just to predict *this*. The more crucial matter is that per-capita incomes empirically move far more with the parameters than predicted, an echo of our earlier calibrations, where the empirical responsiveness of per-capita income to savings and population was larger than the theory would have us believe.

It turns out that the human capital extension of the Solow growth model studied in Chapter 3 goes a significant way towards reconciling and explaining these anomalies. Recall that in that model, we obtained a formula for the effective steady state $\hat{y}^*$ given by equation (3.24). We used its logged expression (4.7) in this chapter to obtain a conditional growth regression; we recall that logged expression here as

$$\ln \hat{y}^* \simeq \frac{1}{1-a-c} A + \frac{a}{1-a-c} \ln s_k + \frac{c}{1-a-c} \ln s_h - \frac{a+c}{1-a-c} \ln(n+\pi+\delta), \quad (4.19)$$

where recall that we've set $\delta_k = \delta_h$ to the same value δ for simplicity. We can open this up using the formula $y(t) = \hat{y}^* e(t) \simeq \hat{y}^* e(0)(1+\pi)^t$ — again, the presumption is that *countries are already in steady state* — and then take logarithms to see that

$$\ln y(t) \simeq a(t) + \frac{a}{1-a-c} \ln s_k + \frac{c}{1-a-c} \ln s_h - \frac{a+c}{1-a-c} \ln(n+\pi+\delta), \quad (4.20)$$

where $a(t) = \frac{1}{1-a} \ln A + \ln e(0) + t \ln(1+\pi)$ just as in equation (4.18). In fact, (4.20) is is the exact analogue of (4.18), but with human capital incorporated.

Notice that the theoretical predictions change significantly once human capital is in the model. As a rule of thumb, say that the three inputs have equal shares of national income, so that $a = c \simeq 1/3$. Then the coefficients on savings in physical capital and human capital, $a/(1-a-c)$ and $c/(1-a-c)$, are both predicted to be around 1, while the coefficient on log population is predicted to be -2. These are much larger predictions, and for good reasons that we've already mentioned in Chapter 3.5.2. An increase in s_k, for instance, not only increases the accumulation of physical capital, it amplifies the savings of *human* capital as well, creating an echo effect. The same is true of s_h. And an increase in population growth, quite apart from its direct effect in reducing per-capita income, has indirect effects on *both* human and physical savings, thereby generating a coefficient that is the sum of two savings coefficients.

Column 2 in Table 4.3 runs the "augmented" steady state regression corresponding to equation (4.20). Apart from measures of s_k and n, it now introduces a proxy for s_h, which is the percentage of the working-age population enrolled in secondary school.[47] This augmented model has an even larger correlation coefficient of 83%, and now the model displays a better fix on the numbers, and the population coefficient, while not *precisely* the sum of the coefficients on s_k and s_h, is not far off either. Investment rates in physical and human capital correlate well with per-capita income, and to a quantitative degree that isn't way out of line relative to the extended Solow model.

[47] We omit a discussion of the rationale behind these choices; see Mankiw, Romer and Weil (1992).