3

Solow's Surprise: Investment Is Not the Key to Growth

Politicians are the same all over. They promise to build bridges, even where there are no rivers.

Nikita Khrushchev

Nobel laureate Robert Solow published his theory of growth in a couple of articles in 1956 and 1957. His conclusion surprised many, and still surprises many today: investment in machinery cannot be a source of growth in the long run. Solow argued that the only possible source of growth in the long run is technological change. Solow in the 1957 article calculated that technological change accounted for seventh-eighths of U.S. growth per worker over the first half of the twentieth century.

While economists applied (and still apply) Solow's model of growth to many poor countries, many are reluctant to accept his view that technological change, not investment, drives long-run growth. While development practitioners slowly weaned themselves from the Harrod-Domar conclusion that growth was proportional to investment in the short run, they continued to believe that investment was the dominant determinant of growth in the long run.

Economists call the belief that increasing buildings and machinery is the fundamental determinant of growth *capital fundamentalism.* Whether capital fundamentalism holds is fiercely debated in the academic literature on growth; we will see in the next chapter what happens when the notion of "capital" is extended to include skills and education—human capital. In this chapter, we will see that capital fundamentalism is incompatible with "people respond to incentives."

But capital fundamentalism has few doubters in the international financial institutions. Paging through their recent reports, one finds statements like these: "The adjustment experience of sub-Saharan Africa has demonstrated that to achieve gains in real per capita GDP an expansion in private saving and investment is key" (International Monetary Fund, 1996).[1] Latin America too must meet "the challenge of sustaining the level of investment necessary for continued output growth" (Inter-American Development Bank, 1995).[2] In the Middle East, "Improving the investment performance—in both human and physical assets—is an important determinant of the ... region's ability to grow" (IMF, 1996).[3] In East Asia, "accumulation of productive assets is the foundation of economic growth" (World Bank, 1993).[4] In case you have any remaining doubts, you should know that "additional investment is the answer—or part of the answer—to most policy problems in the economic and social arena" (United Nations 1996).[5]

But the conventional wisdom that investment in buildings and machinery is the key to long-run development is another panacea that has not met expectations.

Solow's Shocker

To see how Solow arrived at his surprising conclusion that investment cannot be the source of growth, let's go back to his original vision of growth in his 1956 article, with the 1957 follow-up article. The more men and machines an economy had, the higher its production was. Over time production would grow as we invested in more machines and had more workers.

When we say "growth," what we mean is that each person's standard of living should keep increasing. The only way that we can have a higher standard living for each of us, on average, is if each of us produces more goods, on average. So what we are interested in is production per worker, sometimes called *labor productivity*.

We want production per worker to increase, and there are only two inputs into production: machines and workers. So you might think that the way to increase production per worker is to increase machines faster than the number of workers is increasing. In other words, the way to increase production per worker is to increase machines per worker.

But increasing machines per worker immediately runs into problems. As we increase machines per worker, eventually each worker will be using more than one machine at once, dashing madly from one machine to another, like Charlie Chaplin in the movie *Modern Times*. It's hard to believe that anything good will happen to production from giving one more machine to a worker who already has eight of them. This is diminishing returns.

Diminishing returns has a simple and unavoidable logic: increasing one ingredient of production relative to another ingredient indefinitely cannot increase production indefinitely. When you increase machines relative to workers, the return to each additional machine will get lower and lower.

To see diminishing returns in action, suppose for a moment that one ingredient is fixed, and you try to increase the other one.

The Flour Next Time

Today I am making my kids' favorite breakfast food, pancakes. My pancake recipe calls for one cup milk and two cups Bisquick flour. These proportions are not totally rigid. I think my pancake connoisseurs will still eat them if I make the pancakes thinner by using more milk than the recipe calls for.

Then I realize that I have just barely the right amount of Bisquick for pancakes sufficient for my three children. Suddenly my daughter Rachel reminds me that her friend Eve is coming over for brunch. I knew this but forgot. Concealing the bowl of pancake batter from her view, I slip another cup of milk into the bowl. Nobody will notice. Then my son, Caleb, reminds me that his friend, pancake-devouring Kevin, is coming over for brunch too. I slip some more milk into the batter. Maybe they won't notice. Then my co-parent comes in and reminds me that my preschooler Grace's friend Colleen is coming too. In desperation I dump yet more milk into the pancake batter. Fifteen minutes later, the eating audience rejects the world's thinnest pancakes in disgust.

This is diminishing returns in action: increasing one ingredient while the other ingredient is unchanged does not enable me to achieve sustained growth in production of pancakes. Diminishing returns sets in to the ingredient that I am trying to increase (milk) while the other ingredient (Bisquick) is unchanged. I indeed have

diminishing returns to milk. The effect of the first cup of milk on my pancake production was very favorable. Without that cup of milk I have nothing but dry Bisquick; with it I have at least a thick pancake. But when I have already dumped in three cups of milk for only two cups of flour, adding yet one more cup of milk has a pitiful effect on pancake production.

We can increase production of GDP for a given number of workers by increasing machines per person. If there were no machines to begin with, this is Okay; then an additional machine would increase output a lot. When there were already plenty of machines, an additional machine would increase output very little.

How severe these diminishing returns are going to be depends on how important capital is in production. The diminishing returns in my pancake experiment depended on how important the ingredient was that I tried to expand by itself. My failed attempt to expand pancake production by increasing one ingredient would have been even more disastrous if I had been increasing one of the more minor ingredients, like salt, holding everything else constant. I don't think my customers would like the results if I tried to double pancake production by adding more and more salt to an unchanging amount of flour and milk.

If a minor ingredient like salt had been the only ingredient in fixed supply, on the other hand, I would have had a lot more potential to expand pancake production. If I had run out of salt and still had plenty of flour and milk left, I would have been in fine shape for the demands of the children. I think I could have got away with it if I doubled flour and milk together, leaving salt unchanged. A lot of the debate about capital fundamentalism will turn on how important capital is as an ingredient to production.

The reason that Solow's diminishing returns to investment had particular fury was that buildings and machines are a surprisingly minor ingredient in total GDP. We can get a measure of the importance of capital in the United States by calculating the share of capital income in total income. *Capital income* means all the income that accrues to the direct or indirect owners of the buildings and machines: corporate profits, stock dividends, and interest income on loans (since loans finance part of investment). Solow estimated capital income to be about one-third of total GDP in the United States in his 1957 article.[6] It is still about one-third of total income today.[7] The other two-thirds of income is wage income, that is, income to workers.

Thus, capital accounts for only one-third of total production, and workers account for two-thirds of total production. If capital accounts for only one-third of output, then diminishing returns to investment are going to be severe. When machines are scarce, the additional output from one more machine will be high. When machines are abundant, additional output from one more machine will be low.

Not the Way to Grow

Diminishing returns all seems simple and obvious, but it led to Solow's surprise. Increasing machines was *not* a feasible way to sustain growth. If an economy tried to grow by buying more and more machines, then there might be extremely high growth at the beginning when machines were scarce. But diminishing returns means that growth would fall as machines become abundant relative to the labor force. If machines per person grew at a constant rate, eventually the growth of output per person would drop to zero.

Another surprising implication of Solow's view was that saving will not sustain growth. The saving diverts money from consumption today toward buying machinery for production tomorrow, but this does *not* raise the long-run rate of growth, because machinery cannot be a source of long-run growth. So high-saving economies would achieve no higher sustained growth than a low-saving economy would. Growth in both cases would drop to zero as the unavoidable diminishing returns to increasing machines set in. The high-saving economy would have higher income than the low-saving economy, but neither would be able to sustain growth.

Here was Solow's surprise: the simple logic of production suggested that growth of output per worker could not be sustained. Yet the United States and many other industrial economies had already sustained economic growth of 2 percent per worker for two centuries. How did we observe sustained growth of output per worker when such sustained growth is not logically possible?

It's Technology, Stupid

Solow's solution to his surprising paradox was technological change. Technological change would progressively economize on the ingredient in fixed supply: labor. In other words, technological change keeps making a given amount of labor go further.

Solow argued that technological progress happened for noneconomic reasons like advances in basic science. Judging by the steady advance of the technological frontier in the United States, it was plausible to assume a constant rate of technological progress. It was this rate of technological progress that determined long-run growth of income per person.

Think of technology as a blueprint that arranges the workers and machines. Technological change means these blueprints get better and better. Say that the workers first had blueprints telling each of them to follow the item being manufactured all the way through the production process. I haul the raw material from the pile out back, then carry it to the melting-down machine, and I melt it down. I next carry the molten slop over to the molding machine and mold the slop into a product. Then I take the molded product over to the finishing machine, and I finish it. Then I carry it over to the painting machine, and I paint it. I throw the product into the shipment truck. Then I get into the shipment truck and drive it over to the house of the customer who had ordered the product. I take the customer's money and go to the bank to deposit it and then drive back to the plant. Then I haul some more raw material from the pile out back, carry it over to the melting-down machine . . .

Then I get a new blueprint in the mail, courtesy of a certain Mr. H. Ford of Dearborn, Michigan. Mr. Ford suggests that it would be more efficient to have each worker stay at one machine and have the product rather than the workers move. Mr. Ford suggests installing a conveyor belt to carry the product from one machine to the next. So now I stay put at one machine, the painting machine. All of the time that I spent running from one machine to the next is eliminated. I also get very skilled at painting. I can use the extra time and skill to paint more products. Each of the other workers at the other machines also has extra time to produce more. The new labor-saving blueprint allows a given number of workers to produce more with the same machines.[8]

If the new blueprint comes along at the same time as new machines are added, then the technical leap forward will stave off diminishing returns. I am more effective because of the more intelligent way of arranging my labor time. The new blueprint effectively gives us more workers, so effectively labor and machinery have both increased, and there is no diminishing returns to machinery.

This example illustrates the general principle: technical change will avoid diminishing returns if it saves on the ingredient in fixed supply: labor. Each worker becomes more and more efficient thanks to better technology, so it seems as if there were more workers. The effective number of workers keeps up with the increasing number of machines, so diminishing returns never sets in.

In the long run, all of growth of production per worker has to be labor-saving technical change.

An Aside About the Luddite Fallacy

Some people believe labor-saving technological change is bad for the workers because it throws them out of work. This is the Luddite fallacy, one of the silliest ideas to ever come along in the long tradition of silly ideas in economics. Seeing why it's silly is a good way to illustrate further Solow's logic.

The original Luddites were hosiery and lace workers in Nottingham, England, in 1811.[9] They smashed knitting machines that embodied new labor-saving technology as a protest against unemployment (theirs), publicizing their actions in circulars mysteriously signed "King Ludd." Smashing machines was understandable protection of self-interest for the hosiery workers. They had skills specific to the old technology and knew their skills would not be worth much with the new technology. English government officials, after careful study, addressed the Luddites' concerns by hanging fourteen of them in January 1813.

The intellectual silliness came later, when some thinkers generalized the Luddites' plight into the Luddite fallacy: that an economy-wide technical breakthrough enabling production of the same amount of goods with fewer workers will result in an economy with—fewer workers. Somehow it never occurs to believers in Luddism that there's another alternative: produce more goods with the same number of workers. *Labor-saving technology* is another term for *output-per-worker-increasing technology*. All of the incentives of a market economy point toward increasing investment and output rather than decreasing employment; otherwise some extremely dumb factory owners are forgoing profit opportunities. With more output for the same number of workers, there is more income for each worker.

Of course, there could very well be some unemployment of workers who know only the old technology—like the original Luddites—and

this unemployment will be excruciating to its victims. But workers *as a whole* are better off with more powerful output-producing technology available to them. Luddites confuse the shift of employment from old to new technologies with an overall decline in employment. The former happens; the latter doesn't. Economies experiencing technical progress, like Germany, the United Kingdom, and the United States, do not show any long-run trend toward increasing unemployment; they do show a long-run trend toward increasing income per worker.[10]

Solow's logic had made clear that labor-saving technical advance was the only way that output per worker could keep increasing in the long run. The neo-Luddites, with unintentional irony, denigrate the only way that workers' incomes can keep increasing in the long run: labor-saving technological progress.

The Luddite fallacy is very much alive today. Just check out such a respectable document as the annual *Human Development Report* of the United Nations Development Program. The 1996 *Human Development Report* frets about "jobless growth" in many countries. The authors say "jobless growth" happens whenever the rate of employment growth is not as high as the rate of output growth, which leads to "very low incomes" for millions of workers. The 1993 *Human Development Report* expressed the same concern about this "problem" of jobless growth, which was especially severe in developing countries between 1960 and 1973: "GDP growth rates were fairly high, but employment growth rates were less than half this."[11] Similarly, a study of Vietnam in 2000 lamented the slow growth of manufacturing employment relative to manufacturing output.[12] The authors of all these reports forgot that having GDP rise faster than employment is called *growth of income per worker*, which happens to be the only way that workers' "very low incomes" can increase.[13]

Transitions

Increases in machinery per worker could not be a source of long-run growth, but they could be a source of growth in the transition to the long-run path. An economy that started with very few machines would have a very high return to each additional machine. Because of these high returns, investment would temporarily bring high growth. As the machines accumulated, diminishing returns would set in, and growth would fall. Eventually the economy would settle down to a

comfortable existence at the growth rate of labor-saving technological progress. So we *could* revive investment as an important source of growth if transitions are important relative to long-run growth.

However, there are problems with the idea that transitions are important relative to the long-run growth rate. If most growth comes from the transition to the long run, then there must have been very few machines originally. The return to those machines must have been very high, because they were so scarce. This means the return on machines—the interest rate—in the economy would be very high at the beginning. In fact, interest rates would have had to be ridiculously high; Robert King and Sergio Rebelo calculated that the U.S. interest rate would have had to be over 100 percent a century ago for transitional increases in capital per worker to explain U.S. growth. But the evidence we have on interest rates in the United States suggests that they have been relatively constant over time (certainly never 100 percent anyway); this confirms Solow's finding that U.S. growth was a long-run phenomenon, not a transitional movement from low to high capital.

There is also a logical problem with making transitions and investment important in explaining growth. The assumption is that all economies are starting far away from their long-run position. Then investment in machinery will allegedly help the ones that started below their long-run position to grow rapidly (after which they will grow at the rate of technological change). The ones that started above their long-run position will grow slowly or even decline, until they settle back down at their long-run position (after which they will grow at the rate of technological change).

But the proponents of investment as the engine of growth have not supplied a good reason that all countries would be so far away from their long-run position. In the absence of such a reason, the most logical assumption is that most countries are close to the long-run position. After all, what has the long run been doing all this time?

Solow in the Tropics

Solow never mentioned income differences between countries as something that he was trying to explain. He applied his theory only to growth in the United States, where the key fact was constant growth over a long period. He never mentioned tropical countries in any of his writings; in fact, he never applied his model to any other

country besides the United States Solow is not to blame for how his model was applied to the tropical countries. However, his model became the basic theory of growth taught in economics classes. Economists in the 1960s did apply the Solow framework to explaining a wide variety of growth experiences, including the poor tropical countries.

Here's how it would work in explaining cross-country differences. All countries are assumed to have access to the same technology and the same rate of technological progress. The thinking is that there is no reason that major technological breakthroughs that happen in one country cannot be implemented in other countries. (That doesn't mean that the countries *do* implement them; it means they *could* implement them). Once the blueprints are available in one country, the same blueprints could be used in any other country.

So we rule out differences in available technology. Then the only reason some countries are poorer than others is that they have started with very little machinery. Poor tropical countries will have higher returns to machines than will the rich temperate countries. Poor tropical countries will have strong incentives to grow more rapidly than the mature temperate economies that are growing at the rate of technical progress. Eventually the poor tropics will catch up to the rich temperate zone, and all will grow at the rate of technical progress.

Any country that starts out with low capital will offset this unlucky heritage with very high returns to capital. Since international finance capital flows to countries with the highest rate of return (people respond to incentives), international finance capital will flow to this high-return, low-capital country. The unlucky country will catch up to the more fortunate countries, erasing the memory of its unlucky beginnings. The incentives guarantee that the poor will grow faster than the rich. You can see how nicely this view fits with the postwar optimism about development I described in the previous chapter.

After the failure of growth in many poor countries, the problems with the application of Solow's vision to explain income differences across countries became apparent. Fellow Nobel laureate Robert Lucas pointed out one of the big problems with the naive application of the Solow vision to cross-country income differences. American income per person is fifteen times larger than Indian income per person. In the Solow framework, with technology the same across

countries, this income difference could arise only because U.S. workers have more machines than do Indian workers. How many times more machines would the U.S. workers be required to have to explain an income superiority of 15 times? Since machinery is not very important as an ingredient in production, the answer is: a lot. Lucas's calculation implied that each American worker would have to have around 900 times more machines than each Indian worker.[14] American workers do have many more machines, but not that much more. Those who have done the calculations find that American workers have only about twenty times more capital than Indian workers.

Why is it necessary that Indian workers have such an exorbitant superiority—900 times more machines—to explain an income difference of 15 times? It all goes back to the slight role of capital in production: capital accounts for only about a third of all production. Explaining income differences across countries with a relatively minor ingredient like capital doesn't work. Accounting for all cross-country income differences with Solow's model would require a gargantuan difference in machines per worker.

This should have been—but wasn't—foreseen. After all, Solow himself had shown why machines could not explain differences in income across time for the same country, like the increase in U.S. output per worker over forty years: because machines would have to have been more relatively scarce at the beginning than they really were. It is the same logic that shows why machines cannot explain large differences in income across countries rather than across time.

But the solution to the diminishing-returns problem that Solow advanced for growth in the long run in one country—technical progress determined by noneconomic causes like basic science—does not work across countries. It could make sense to assume that technology changes over time for noneconomic reasons like advances in science. But to say that countries have different growth rates because they have different rates of technological progress for some mysterious noneconomic reason is not very satisfying. This is just answering the question of why growth rates differ by saying that growth rates differ—which leads us back to economic incentives. Technology must vary across countries for economic reasons. If technology is so powerful as to explain sustained income growth over time in the same country, it is the logical candidate to explain big income differences between countries. And if technology differs

between countries, there must be strong economic incentives to get better technology. I take up the idea of technology responding to incentives in Part III.

Returns and Flows

We haven't even gotten to the worst part about the idea that machinery was the key to development. Lucas also calculated the implied rate of return to machines. Indian machinery should be 900 times scarcer than U.S. machinery if we explain all of the U.S.-India income difference with differences in machinery. Lucas used the Solow principle that machines have higher returns where they are scarce and calculated that the profit rate yielded by Indian machines should be 58 times larger if they are so much scarcer. These super-returns are the counterpart to King and Rebelo's calculation that the return to capital would have had to be over 100 percent a century ago if we explained U.S. growth with transitional capital accumulation. With such huge incentives to invest in poor countries, Lucas wondered, "Why doesn't capital flow from rich to poor countries?"

An answer might be that poor countries have disadvantages to the investor like political instability, corruption, and the risk of expropriation. But these differences in rates of return are too large to be canceled out by such factors. The foreign investor in India still comes out ahead even if he only can get out of the country two rupees, on average, of every one hundred rupees of profit. Nobody thinks that the probability of expropriation in India is 98 percent. Even spectacularly venal governments do not attain a theft rate, on average over many years, of ninety-eight cents on the dollar. Even allowing for reasonable Indian political risk, Lucas argued, one should observe capital fleeing from New York to New Delhi. People should respond to incentives.

That didn't happen. In the 1990s, the U.S. economy had a gross inflow of new loans and investments from the rest of the world equal to $371 for each and every American every year. Over the same period, the loans and investments coming into India worked out to an inflow every year for each and every Indian of—four cents. The incentives to invest in India were not there.

There was nothing peculiar about India's paucity of foreign capital for a poor country. In 1990, the richest 20 percent of world population received 92 percent of portfolio capital gross inflows; the poorest 20 percent received 0.1 percent of portfolio capital inflows. The richest 20 percent of the world population received 79 percent of foreign direct investment; the poorest 20 percent received 0.7 percent of foreign direct investment. Altogether, the richest 20 percent of the world population received 88 percent of private capital gross inflows; the poorest 20 percent received 1 percent of private capital gross inflows.

The Growth That Wasn't

The most important evidence against the Solow vision applied across countries was the failure of growth in many poor countries. With high returns to scarce capital, the poor countries had every incentive to grow faster than the rich countries. The poorer the country, the faster the growth should have been. The poor shall inherit the growth. It didn't work out that way.

Ironically, the first economists to recognize the failure of growth in many poor countries were not specialists in poor countries at all. Development economists who did follow poor countries were certainly aware that things were going badly wrong in Africa and Latin America, but they didn't seem to notice the challenge to the old growth paradigm. Instead it took a rich-country economist like Paul Romer to look up the data and point out that the old paradigm was not working.

Romer used data on over a hundred countries from the compilation of country incomes by Robert Summers and Alan Heston. At the time of his presentation at the National Bureau of Economic Research Macroeconomics Annual Conference in 1987, he had data for growth between 1960 and 1981. He showed that the poor countries were not growing any faster than the rich countries. He demonstrated that the Solow prediction applied to tropical countries had failed.

Romer was showing 1960–1981 data to illustrate the failure of the prediction that the poor grow faster. Ironically, these were the good years for poor countries. The poor countries did even worse both before and after these years that supplied the original damaging blow to the old Solow paradigm applied to the tropics.

The last year in Romer's data set, 1981, was also the last good year for many poor countries. As we will see in chapter 5, Latin America and sub-Saharan Africa had two lost decades for economic growth after 1981. The Middle East and North Africa went into the tank a little later. Since 1981, poor countries have not only not caught up to

rich countries; they have done worse than rich countries. They are losing ground.

The poorest three-fifths of countries have had nearly zero or slightly negative growth of income per person since 1981. The bottom two-fifths of countries, already doing badly over the 1960 to 1981 period, continued to do badly between 1981 and 1998. The middle fifth of countries, which had done well between 1960 and 1981, did badly between 1981 and 1998. The richest 20 percent of countries continue to have a positive growth rate of about 1 percent per person. The next richest fifth of countries, which includes the East Asian superstars, also had respectable growth on average.

Rich countries had some slowdown in growth. The United States had growth per person of 1.1 percent over the 1981 to 1998 time frame compared to 2.2 percent between 1960 and 1980. But this slowdown is nothing compared to Nigeria's change in per capita growth per year from plus 4.8 percent over the 1960–1980 period to minus 1.5 percent between 1981 and 1998.

Despite all the moaning and groaning by rich peoples about slow growth, they have done much better on average than the poor countries over the last half century. The ratio of the richest country's per capita income to that of the poorest country has risen sharply over that period. The rich have grown richer; the poor have stagnated (figure 3.1).

For the whole period 1960 to 1999, the poorest countries did significantly worse than the rich countries, with the poorest two-fifths barely mastering positive growth. The poorest four-fifths of countries in 1960 (including only those countries on which we have available data) roughly correspond to what later became known as the Third World. Seventy percent of these Third World countries grew more slowly over the whole period than the median growth of 2.4 percent per capita for the richest countries. They were falling behind, not catching up.

The Mark of History

Now that it was apparent that this prediction of faster growth of poor countries was not working out, economists started asking some pointed questions about poor countries in earlier periods. Economists had taken it as a given that poor countries were poor when they started applying the Solow model to the tropics in the 1960s.

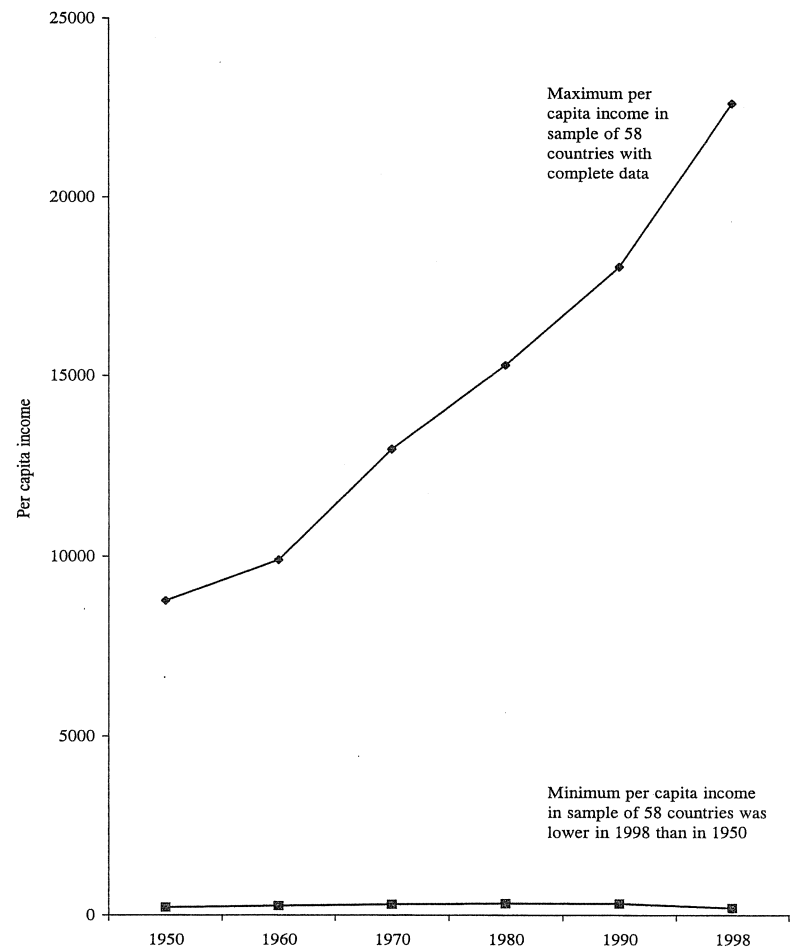

Figure 3.1
The maximum per capita income has grown strongly over the last half century, while the minimum per capita income has stagnated.

Nobody in the 1960s seemed to be asking how the poor nations had gotten to be so much poorer than the rich nations.

A moment's thought supplied the answer, although this moment of thought didn't come along until much later. The poor countries had gotten to be poorer than the rich countries by growing more slowly over some previous period. There had to be some primordial time, back between the Adam and Eve era and now, when the incomes of nations were much more equal. Since the incomes of nations are remarkably unequal now, there must have been a strong process of divergence of national incomes, contradicting the prediction of the Solow model applied across countries that nations' incomes would converge to each other.

Lant Pritchett of the Kennedy School of Government at Harvard crystallized this moment of thought in a recent article.[15] The reasoning is straightforward. The very poor nations today are just barely above the subsistence level in income per person. Subsistence means not starving to death. Therefore, the very poor nations today must have had about the same income a century or two ago as they do today. It couldn't have been less, because that would mean they were below subsistence a century or two ago, which is impossible since they lived to tell the tale. The very rich nations were also much closer to the subsistence level a century or two ago, since we do have data showing they have had substantial growth of income per person over the last century or two. Therefore, the gap between the very richest and the very poorest has grown over the past century or two.

If there's any remaining doubt, you can get data on today's poor countries. An indefatigable economic historian, Angus Maddison, has reconstructed data from 1820 to 1992 on a sample of twenty-six countries. Although the poor countries were underrepresented in Maddison's sample, it is apparent even so that there has been a lot of divergence. The ratio of the richest country—the United States—to the poorest country—Bangladesh—today is about thirty times. The ratio of the richest to poorest in 1820 was only about three times (figure 3.2). All of today's eight poor nations in the Maddison sample were also at or near the bottom in 1820. (The historically highest-ranked nation of today's eighth poorest, Mexico, was already the tenth poorest in 1820.) The countries that were at the bottom in 1820 largely stayed at the bottom; the richest countries increased their incomes by a factor of ten or more.

This is a remarkable outcome. For today's rich countries, more than 90 percent of today's incomes have been created since 1820. Yet

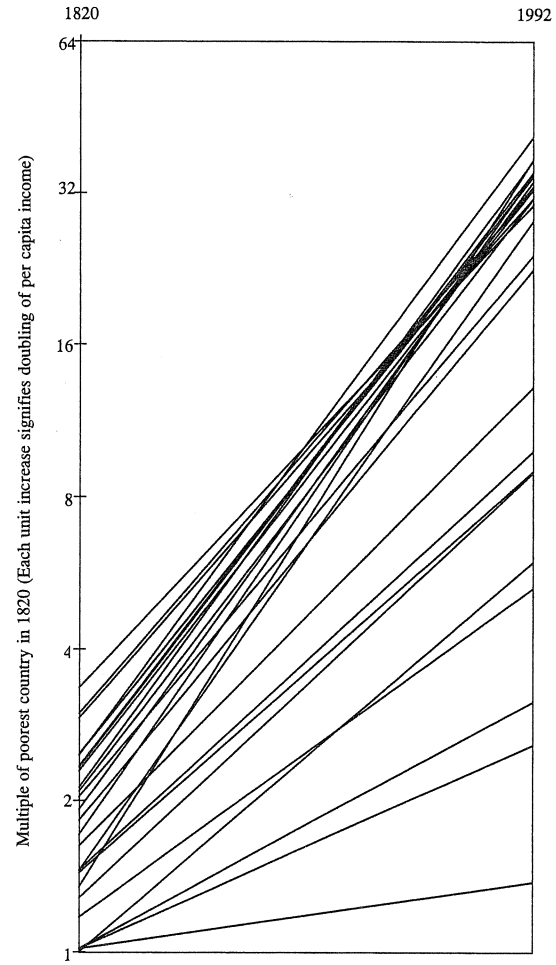

Figure 3.2
The rich got richer, 1820–1992

the income they had attained nearly two centuries ago was already a meaningful predictor whether they would become rich.

The Winners Write Economic History

So why was there a presumption in economic thought for so long that the poor catch up to the rich? William Baumol of Princeton, for example, had a famous paper in which he showed that a group of sixteen industrial countries had caught up to the leader over the past century. The poor among this group of countries had grown faster than the rich. Therefore, he argued that there was a general tendency toward convergence of national incomes.[16]

How had Baumol gotten such a different conclusion to what would later be the seemingly irrefutable argument of Pritchett? Baumol's conclusion, and similar ones that had floated around in economic thought for a long time, turns out to be based on an error. (It's an unmistakable error once you point it out, but not obvious before you point it out—and a nice illustration of how hard economists have to work to figure out even such an elementary question of whether the poor grow faster than the rich.) Brad de Long of Berkeley pointed out the error in Baumol's analysis by asking how Baumol had chosen his group of countries.[17] The countries that have easily available historical data are today's rich countries. It's the rich countries that can afford the economic historians who reconstruct long series of income statistics. Baumol understandably selected a sample of countries that had easily available data—and by doing this unintentionally predetermined the answer in favor of convergence. Naturally these countries, all rich today, wherever they began, will seem to converge to each other. Since the selection did not screen any out on the basis of where they started, they likely started from a variety of circumstances. Some of them likely started out already relatively rich and others relatively poor. Since they all wound up rich at the end—because that's the way Baumol implicitly chose the group—it's a lock that the initially poor in the group of rich-at-the-end countries will have grown faster than the initially rich.

This bias explains why Baumol went astray (as he graciously admitted once de Long pointed it out). More generally, this story helps explain why there was such a bias in economic discussions for so long to assume convergence of national incomes. Economists looked mainly at those that were winners at the end, because those were the countries that had the good-quality data. (Also, economists

from rich countries prefer to talk about and visit other rich countries.) The winners write economic history.

Even Maddison's sample suffered a lot from the selection bias toward winners, as it includes only eight countries that the World Bank today classifies as poor—less than a third of the sample. Since poor nations make up the vast majority of all countries in the world, this is still a severe bias in favor of those that have wound up rich today. The Maddison sample whose 1820 income can be guessed has no country from Africa, for example. This Africa data shortage has everything to do with Africa's poverty. Chad today does not support a lot of economic historians rooting around in their country's past. Already poor (and illiterate) Chad in 1820 did not have a government statistics department churning out figures. From the reasoning that today's poor countries cannot have grown much, it is clear that we would see even more evidence for the rich-getting-richer in a more complete sample.

Even my discussion of trends over the 1960 to 1999 period was biased toward the winners at the end. Virtually all winners at the end have good data; the countries that have run into disasters often do not have complete data. I can check this by looking at the World Bank classification of countries at the end of the period as either industrial (members of the Organization of Economic Cooperation and Development) or developing. My calculation of trends over the 1960 to 1999 period, which already showed the poor countries growing more slowly, used only the 100 countries that have data for 1960 and 1999. Only one industrial country lacks complete data: Germany, because of the difficulty of getting consistent data before and after unification. In contrast, half of the countries the World Bank classifies as developing in 1999 lack complete data. So my 1960 to 1999 sample was biased toward the winners at the end.

I already showed that a tendency for the poor countries to grow more slowly over the 1960 to 1999 period and the rich countries to grow faster. Now I know, because of the bias toward the winners, that even this conclusion was understated. There were likely even bigger disasters among poor countries that dropped out of the data altogether—such as Myanmar, Zaire (Congo), Liberia, Chad, and Haiti. Poor economic performance makes it hard to keep statistical offices running. For example, Zaire's statistical office had collapsed by 1999, but earlier data show long-run growth of −2.4 percent per year.

Growth Accounting Meets the Gang of Four

The most straightforward way to assess the importance of capital accumulation is to account for how much of output growth per worker is explained by capital growth per worker. The contribution of capital growth per worker to output growth per worker is equal to the share of capital in production times the growth rate of capital. As I have already noted, the share of capital in production is about one-third, so if capital per worker were growing at 3 percent, then the contribution of capital to growth would be one percentage point. If growth of output per worker were 3 percent, then we would say that capital accounted for one-third of the growth per worker. The part of growth that is unexplained by capital accumulation will be the part explained by technological progress. The contribution of labor-saving technological progress to growth is equal to the labor share (which is one minus the capital share) times the growth rate of technical change. So if labor-saving technological change were growing at 3 percent, then we would say technological change accounted for two percentage points of the 3 percent growth.

Alwyn Young of the Chicago Business School did this kind of calculation for the fast-growing East Asian economies—the so-called gang of four (Korea, Taiwan, Singapore, and Hong Kong). He reached the conclusion that most of the fast growth of East Asia was due to capital accumulation and a relatively small part due to technological progress. His most startling finding was for Singapore; there, technological progress occurred at a rate of only 0.2 percent per year. Paul Krugman later popularized this finding in *Foreign Affairs*. He drew an analogy between capital-intensive Singaporean growth and capital-intensive Soviet growth, setting off a cyclone of protest. Singapore's prime minister denounced Krugman publicly and announced that Singapore would henceforth have a goal of 2 percent per year technological progress.[18]

Scholars as well as prime ministers have criticized the Young-Krugman finding (justly in my view) on several grounds. First, it doesn't take into account our official motto: people respond to incentives. Robert Barro of Harvard and Xavier Sala-ì-Martin of Columbia pointed out in their textbook on growth that capital accumulation itself responds to technological change. If technology is improving, then the rate of return of capital is improving. If the rate of return on capital is improving, then more capital will be accumulated. In the long run, capital per worker, labor-saving technology, and output per worker will all grow at the same rate (as they did in the example). But we would say that the cause of growth is the growth in technology, to which both capital accumulation and output growth respond. When Peter Klenow and Andrés Rodríguez-Clare redid the Young calculations, taking into account the response of capital to technological change, they found that technological change accounted for a much higher share of output growth than Young had found for the gang of four.

Second, the finding that capital accumulation accounts for East Asian growth, even if it were true, does not address whether that experience can be replicated elsewhere. To address the latter question, we need to see how much the variation in capital growth rates across countries accounts for the variation of growth per worker across countries. The answer is not much. Klenow and Rodríguez-Clare attribute only 3 percent of the variation of growth per worker across countries to variations in capital growth per worker, while variations in technological progress accounted for 91 percent (human capital accounted for the puny remaining 6 percent).[19] Another study finds that variations in the growth of physical capital explain only 25 percent of the variations in growth performance across countries.[20]

To make things concrete, consider some East Asia and non–East Asia country examples. Both Nigeria and Hong Kong increased their physical capital stock per worker by over 250 percent over the 1960 to 1985 time frame. The results of this massive investment were different: Nigeria's output per worker rose by 12 percent from 1960 to 1985, while Hong Kong's rose by 328 percent. And consider another even more capital-intensive pair: the Gambia and Japan both increased their capital stocks per worker by over 500 percent between 1960 and 1985. The result in the Gambia was that output per worker rose 2 percent from 1960 to 1985, while in Japan it rose 260 percent.[21] These are among the worst comparisons that one can make, but the result holds for the whole sample: variations in capital growth do not explain much of the variations in output growth. (It may be that capital investment is measured incorrectly because not all of the measured "investment" really went into productive machines. I still would conclude that measured investment is not the key to growth.)

To give another example of failure of capital-led growth, capital per worker in Tanzania's manufacturing sector grew at 8 percent per

annum over the period 1976 to 1990, but manufacturing output per worker fell at 3.4 percent per annum over the same period. This is particularly striking because one would expect that manufacturing equipment and technological expertise could be purchased on the international market, and so the relationship between inputs and outputs in manufacturing should not differ much among countries.[22]

Third, the rates of return in East Asia did not behave the way they were supposed to if capital accumulation was the main source of growth. As we saw, the rate of return to capital must be high at the beginning if transitional capital accumulation is the main source of growth. Capital accumulation should lead to diminishing returns; the rate of return to capital should fall. A study in 1997 found that the rate of return to capital in Singapore actually increased over time.[23] This 1997 study concludes that technological progress was central to Singapore's high growth of output per worker. He reached similar conclusions for the other three members of the gang of four.

Conclusion

The World Bank helped finance the Morogoro Shoe Factory in Tanzania in the 1970s. This shoe factory had labor, machines, and the latest in shoe-making technology. It had everything except—shoes. It never produced more than 4 percent of its installed capacity. The factory, which had planned to supply the entire Tanzanian shoe market and then export three-quarters of its planned production of 4 million shoes to Europe, never exported a single shoe. The plant was not well designed for Tanzania's climate; it had aluminum walls and no ventilation system. Production finally ceased in 1990.[24]

Why machines in many developing countries are no more productive than tail fins on a Chevy has little to do with the machines themselves and everything to do with the environment in which producers used the machines. Morogoro Shoe Factory was owned by the government of Tanzania, a government that had failed at every big and small development initiative since independence.

Multiplying machines when incentives for growth were lacking was useless. Maybe the machines would produce things nobody wanted. Or maybe the machines were there but other crucial inputs were unavailable (a common problem in Tanzania and elsewhere was that imported raw materials and spare parts were often unavailable because of government controls on selling dollars to producers). Not only could machines not be a permanent source of growth, even their genuine productive potential often went to waste because governments messed up the market incentives to use machines efficiently.

Even when machines were used efficiently, Solow's original insight that capital could not be the ultimate source of growth was right on target. There is more capital in richer economies, but that is because technological progress offsets diminishing returns.

The facts contradict the capital fundamentalists. The imams of capital fundamentalism who applied the Solow model to the tropics turned this insight on its head. If transitional capital accumulation were the main source of growth differences, then countries should have very high rates of return to capital at the beginning. They do not. If transitional capital accumulation were the main source of growth differences, we would expect the poor capital-scarce countries to grow faster than the rich as they respond to these high returns to capital. They do not. If transitional capital accumulation were the main source of growth differences, we would expect financial capital to flow from rich to poor countries in response to the high returns to capital. It does not. If transitional capital accumulation were the main source of growth differences, we would expect capital accumulation to explain a lot of the cross-country differences in growth. It does not. Trying to grow by physical capital alone was another useless panacea.

That's not the end of the story, because there would be a determined effort to revive the application of the Solow model to poor countries by augmenting it with education of workers—human capital. A new group of scholars would claim that controlling for education and saving, poor countries did tend to grow faster than rich countries. To see if education proved to be the panacea for growth, let's turn to the next chapter.