

**NBER WORKING PAPER SERIES**

**WHY, INDEED, IN AMERICA?  
THEORY, HISTORY, AND THE ORIGINS  
OF MODERN ECONOMIC GROWTH**

**Paul M. Romer**

**Working Paper 5443**

**NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
January 1996**

Comments from Gavin Wright are gratefully acknowledged. This work was supported by the Canadian Institute for Advanced Research. This paper is part of NBER's research programs in the Development of the American Economy, and Economic Fluctuations. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

© 1996 by Paul M. Romer. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

WHY, INDEED, IN AMERICA?  
THEORY, HISTORY, AND THE ORIGINS  
OF MODERN ECONOMIC GROWTH

ABSTRACT

When they are used together, economic history and new growth theory give a more complete picture of technological change than either can give on its own. An empirical strategy for studying growth that does not use historical evidence is likely to degenerate into sterile model testing exercises. Historical analysis that uses the wrong kind of theory or no theory may not emphasize the lessons about technology that generalize. The complementarity between these fields is illustrated by an analysis of early industrialization. The key theoretical observation is that larger markets and larger stocks of resources create substantially bigger incentives for discovering new ways to use the resources. This simple insight helps explain why the techniques of mass production emerged in the United States during the first half of the 19th century. It also helps explain how a narrow advantage in the techniques of mass production for a small set of goods grew into broad position of industrial supremacy by the middle of the 20th century.

Paul M. Romer  
Department of Economics  
601 Evans Hall, #3880  
University of California, Berkeley  
Berkeley, CA 94720-3880  
and NBER

Whether new growth theory and economic history are a good match depends on the kind of question you address and the kind of answer you expect. I find that they complement each other when I try to answer questions about the world. Economists who believe that these lines of inquiry can go their separate ways are addressing entirely different kinds of questions or have a different notion of what it means to give a good answer.

#### I. Growth without history

Many recent attempts at testing models of growth proceed without making any reference to evidence from economic history. They rely on data series for many countries, typically for the last 30 or so years. They focus on questions about models instead of questions about the world. A representative conclusion is that the right model of economic growth is neoclassical in an extreme sense: It assumes that the technology is the same in all countries and concludes that exogenous differences in saving and education cause all of the observed differences in levels of income and rates of growth (Mankiw, 1995).

However, to take a specific case, differences in saving and education do not explain why growth was so much faster in the United States than it was in Britain around the turn of this century. In 1870, per capita income in the United States was 75% of per

capita income in Britain. By 1929, it had increased to 130%. In the intervening decades, years of education per worker increased by a factor of 2.2 in Britain and by a nearly identical factor of 2.3 in the United States. In 1929, this variable remained slightly lower in the U.S. (Data are taken from Maddison, 1995.)

In addition, differences in rate of investment in the two countries were not the result of exogenous differences in savings rates. The remarkable fact about the British economy during this period is how much of domestic savings was devoted to investment abroad. In the decade prior to 1913, net domestic investment was roughly equal to net foreign investment (Cairncross, 1953, p. 121). By 1914, net foreign assets were equal to 1.5 times GDP. To understand what happened in Britain, we must explain why investment abroad, especially in the United States, was so attractive to British savers.

It is difficult to look at the data for these two countries without wondering whether the well-documented technological developments in the United States aren't at least part of the story. Nevertheless, the standard model-testing exercise does not even consider this possibility. Nor does it seek out any direct evidence that would help us decide how important any differences in the technology might be. This would be a glaring flaw if the goal truly were to understand events in the world, but it is as natural as a null hypothesis if all you want to do is test models.

## II. History without theory

A second approach recognizes the value of economic history but denies the need for formal theory. It shows up each time someone proposes a new piece of mathematical

formalism. Only thirty years ago many economists still objected to a mathematical statement of the relationship between output and capital in terms of an aggregate production function and an aggregate stock of capital,  $Y=F(K,L)$ . Twenty years ago, a different group of economists objected when labor economists used mathematical equations and a new human capital variable  $H$  to capture the observation that a person's skills could be enhanced by investing in education or experience. Ten years ago, many economists readily acknowledged that output of knowledge must somehow be related to the inputs devoted to the production of knowledge but objected nevertheless when growth theorists suggested that economists make another try at capturing these relationships using mathematical expressions of the form  $dA/dt=G(H,A)$ .

Every time a familiar argument is translated for the first time from natural language into mathematics, the same objections arise. "These equations are so simplistic and the world is so complicated." This reflects a misapprehension of the role of formal theory. Set aside models. The key is to understand what it means to answer a question about the world. In the lead-up to his exposition of evolutionary theory, Richard Dawkins (1987) gives a refreshingly straightforward description of what constitutes a good answer to a such a question:

"If I ask an engineer how a steam engine works, I have a pretty fair idea of the general kind of answer that would satisfy me. Like Julian Huxley, I should definitely not be impressed if the engineer said that it was propelled by '*force locomotif*.' And if he started boring on about the whole being greater than the sum of its parts, I would interrupt him: 'Never mind about that, tell me how it *works*.' What I would want to hear

is something about how the parts of an engine interact with each other to produce the behavior of the whole engine. I would initially be prepared to accept an explanation in terms of quite large subcomponents, whose own internal structure and behavior might be quite complicated and, as yet, unexplained. The units of an initially satisfying explanation could have names like fire-box, boiler, cylinder, piston, steam governor. ... *Given* that the units each do their particular thing, I can then understand how they interact to make the whole engine move.

“Of course I am then at liberty to ask how each part works. Having previously accepted the *fact* that the steam governor regulates the flow of steam, and having used this fact in my understanding of the behavior of the whole engine, I now turn my curiosity on the steam governor itself” (p. 11).

The central element in this account of what Dawkins unapologetically calls *hierarchical reductionism* is a recognition that explanation operates on many levels that must be consistent with each other. What theories do for us is take all the complicated information we have about the world and organize it into this kind of hierarchical structure.

In building this structure, good theory tells us how to carve a system at the joints. At each level, theory breaks a system down into a simple collection of subsystems that interact in a meaningful way. Dawkins could have used a simple theory that makes a bad split of the engine into its front half and its back half. Instead, he uses a simple theory that makes a good split into the fire-box, the boiler, and so on. What growth theory must do is provide a good, simple split of the opportunities available to us in the physical world.

### III. Neoclassical versus new growth theory

Neoclassical growth theory explains growth in terms of interactions between two basic types of factors: technology and conventional inputs. At the next level, conventional inputs can then be subdivided into physical capital, labor, and human capital. The initial split into technology and conventional inputs is promising because technology does differ from all other inputs. However, for technical reasons, neoclassical theory mapped this split onto the theoretical dichotomy between public and private goods. This means that the theory leads to a dead end when you try to understand the details about technology in a second stage analysis analogous to Dawkins' investigation of the steam governor. Technology in the model does not correspond to anything in the world. You can understand capital in terms of things like machine tools that you can observe, but if you ask for a description of technology, neoclassical theory will tell you only about things that live in models, shifting production possibility frontiers and the like.

The obvious real world candidates for technology simply are not public goods. For example, a promising line of work in the 1960s studied embodied technological change. Implicitly, it modeled technology as designs for machines. This line of work lost its momentum, perhaps because of the difficulty people had in reconciling what we know about machine design with an initial cut that makes technology a public good. In their evolutionary alternative to neoclassical growth theory, Richard Nelson and Sidney Winter (1982) rejected the public good assumption and represented technology as routines followed within firms. Recent generations of neoclassical growth theorists have not

followed up either approach and have contented themselves with a '*force locomotif*' explanation: "Technological change causes economic growth."

New growth theory started on the technology-as-public-good path and worried about where technology came from, but soon backed up and reconsidered the initial split that economists make in the physical world. We now start by dividing the world into two fundamentally different types of productive inputs that we can call "ideas" and "things." Ideas are nonrival goods you could store in a bit string. Things are rival goods with mass (or energy). With ideas and things, you can explain how economic growth works. We use nonrival ideas to rearrange things, for example when we follow a recipe and transform noxious olives into tasty and healthy olive oil. Economic growth arises from the discovery of new recipes and the transformation of things from low to high value configurations.

This slightly different initial cut leads to insights that do not follow from the neoclassical model. It emphasizes that ideas are goods that are produced and distributed just as other goods are. It removes the dead end in neoclassical theory and links microeconomic observations on routines, machine designs and the like with macroeconomic discussions of technology.

In an analysis of American and British growth, the insight that is most relevant concerns scale. By definition, a nonrival idea can be copied and communicated, so its value increases in proportion to the size of the market in which it can be used. For example, if barriers to trade meant that a computer operating system written in Washington state could only be used in Washington state, it would be worth far less than



if it could be used all over the world. If there were only a few olive trees, no one would have bothered to figure out how to use the olives. If people can sometimes establish property rights over a nonrival good like an operating system or a recipe — this is the possibility that the public good approach precludes — differences in scale will change the rewards for producing new ideas.

#### IV. Why in America?

A great deal of historical analysis has addressed the performance of British and American economies around the turn of the century. For general discussions, see Nathan Rosenberg (1981), Nelson and Gavin Wright (1992), and Moses Abramovitz and Paul David (1995). From the beginning, observers have pointed to the abundance of natural resources in the United States as an early advantage, especially in agriculture. The surprising conclusion that emerges from recent historical scholarship is that resource abundance also interacted with scale to create a technological lead in manufacturing that persisted well into the 20th century.

The United States started as little more than an importer of European technology, but by the first decades of the 19th century, distinctively American technologies began to emerge. Entrepreneurs and inventors developed specialized machines that economized on human effort and made prolific use of the natural resources and energy that were available here (Rosenberg, 1981). Other nations in the new world also faced low prices for natural resources relative to labor. For example, Angus Maddison's data suggest that Australia had the highest level of GDP per capita from 1870 to 1900 because its stock of

resources was so large relative to its population. What made the United States unique was the combination of resource abundance and large markets (Abramovitz and David, 1995). In 1820, the population was 534,000 in Argentina, 33,000 in Australia, and 9.6 million in the United States. Moreover, even at this early date, the United States had a transportation system and a commercial infrastructure that effectively linked most of its citizens into a truly national market. By 1870, the population had grown to 1.8 million in Argentina, 1.6 million in Australia, and 40 million in the U.S., a third more than lived in the UK at that time.

As Rosenberg has observed (1963, 1981), large markets, which were also populated here by relatively homogeneous consumers, mattered because they encouraged firms to incur the design and set-up costs necessary for long production runs of standardized goods assembled from interchangeable parts. As he emphasizes, they also mattered because they induced large markets for specialized machines. The differences in incentives created by market size were presumably of great consequence when populations differed by a factor of 10 or 20 and flows of goods between nations were still relatively limited. More direct evidence that market size and incentives did matter for invention can also be inferred from Ken Sokoloff's evidence on the geographic pattern of patent awards in the United States (1988). His data show that inventive activity was concentrated around locations that had access to cheap transportation and expanded into new areas when the transportation system improved.

Resource abundance and scale effects were therefore key elements in the development of the production using specialized machinery, standardized goods, and

interchangeable parts. By the middle of the 19th century, when the British first started to take notice, this system was used in only a few industries, gun-making most famously. Other important industries in the United States such as iron-making still lagged behind their British counterparts. It took another half century or more for per capita output in the United States to move ahead of Britain. Scale effects continued to be crucial in this later period as well.

In the beginning, machinery was made in machine shops that were part of large manufacturing enterprises like textile mills. When markets grew, these shops eventually separated from their parent firms and began to operate as suppliers to many firms. However, the growth in potential markets came not just through growth in the industry of the parent firm. Most of it came from growth in other industries because of what Rosenberg has identified as a process of technological convergence (1963), which created an additional scale effect distinct from the one associated with population size. Firms engaged in the production of many different kinds of goods (including machine tools themselves) all used the same kinds of machinery to shape first wood, then metal. Thus, the former machine shop of the textile mill sold not just to other textile firms, but to all manner of manufacturing enterprise. As a result, the proliferation of specialized machine tools was limited only by the extent of what came to be a very large market.

Thus, scale acted through larger markets for both final goods and capital goods. Scale in this sense was determined by a large population, an integrated market, and technological convergence. A large quantity of natural resources acted initially by changing the price of materials relative to labor and encouraging the use of machinery.

Over time, abundant quantities of *potential* natural resources created an additional scale effect relating to the supply of the things that could be transformed by any particular new idea. This effect was most obvious in the development of uses for by-products (Rosenberg, 1985). For example, the quantity of animal waste grew with the expansion of the meat-packing industry. Its geographic concentration also increased as refrigeration and the railroad made it possible for meat-packing to be separated from the site of final consumption. This concentrated increase in the volume of animal by-products created incentives for firms to come up with new nonrival goods — literally, in this case, new recipes — for making use of raw materials that had previously been discarded as waste. This process ultimately led to the development of a by-products industry that was one of the early users of industrial chemistry.

The same motivation led to the investments that were needed to take advantage of other natural resources. Because of the quantities of resources that were available and the large markets for goods, large investments in basic technologies for extracting and processing these resources could be sustained. Because of these kinds of efforts, the U.S. became the world's leading supplier of virtually every industrial raw material, a fact that is reflected in high and increasing intensity of resources in U.S. exports from 1880 to 1930 (Wright, 1990). With the exceptions of wood and land, the U.S. achieved leadership in most raw materials because of its intensive use of its endowment, not because of the endowment itself (Wright, 1990). Because of the “congruence” (in the terminology adopted by Abramovitz and David, 1995) between the U.S. strength in intensive resource use and its early strength in manufacturing technologies, it developed a technological lead

over the rest of the world that expanded throughout the first half of this century (Nelson and Wright, 1992).

## V. Conclusion

Scale effects are clearly not the only interesting factor in this story. For example, new institutions like the United States Geological Survey, the private university, the large multidivisional firm, and the specialized research laboratory were important as well. Concerning the scale effects themselves, the arguments presented here will not tell historians anything they did not already know. The relatively modest contribution that new growth theory can make is to move the issue of scale up in our conceptual hierarchy. Scale effects should no longer be treated the way a growth accountant such as Dennison did, as a kind of afterthought that had something to do with plant size. They should be treated the way Adam Smith did, as one of the fundamental aspects of our economic world that follow from the nonrival character of ideas.

If new growth theorists have their way, the first distinction we will draw when we look at the physical world will be the one that separates rival things from nonrival ideas. Right from the start, this should be the way we carve up our physical world into a small number of interacting elements analogous to pistons and boilers. When the resulting theoretical framework is combined with the evidence and inferences from economic history, economists will be able to give a more convincing answer to the question of how industrial growth works and why it emerged first in America.

## References

- Abramovitz, Moses and David, Paul A. "Convergence and Deferred Catch-up: Productivity Leadership and the Waning of American Exceptionalism," in Ralph Landau, Timothy Taylor and Gavin Wright, eds., The Mosaic of Economic Growth. Stanford, CA: Stanford University Press, 1996, pp. 21-62.
- Cairncross, A. K. Home and Foreign Investment 1870-1913. Cambridge: Cambridge University Press, 1953.
- Dawkins, Richard. The blind watchmaker. New York: W. W. Norton, 1986.
- Maddison, Angus. Monitoring the World Economy, 1820-1992. Paris: OECD, 1995.
- Mankiw, N. Gregory. "The Growth of Nations." Brookings Papers on Economic Activity, 1995, (1), pp. 275-326.
- Nelson, Richard R. and Wright, Gavin. "The Rise and Fall of American Technological Leadership: The Postwar Era in Historical Perspective." Journal of Economic Literature, December 1992, 30, pp. 1931-64.
- Rosenberg, Nathan. "Technological Change in the Machine Tool Industry, 1840-1910." Journal of Economic History, December 1963, 23, pp. 414-43. Reprinted in: Rosenberg, Nathan. Perspectives on Technology. Armonk, NY: M. E. Sharpe, Inc., 1976.
- Rosenberg, Nathan. "Why in America?," in Otto Mayr and Robert C. Post, eds., Yankee Enterprise, the Rise of the American System of Manufactures. Washington D.C.: Smithsonian Institution Press, 1981. Reprinted in Exploring the Black Box, Cambridge: Cambridge University Press, 1995.

- Rosenberg, Nathan. "The Commercial Exploitation of Science by American Industry," in Kim B. Clark, Robert H. Hayes, Christopher Lorenz, eds., The Uneasy Alliance. Boston: Harvard Business School Press, 1985, pp. 18-51.
- Sokoloff, Kenneth L. "Inventive Activity in Early Industrial America: Evidence from Patent Records." Journal of Economic History, December 1988, 48, pp. 813-50.
- Wright, Gavin. "The Origins of American Industrial Success, 1879-1940." The American Economic Review, September 1990, 80(4), pp. 651-68.