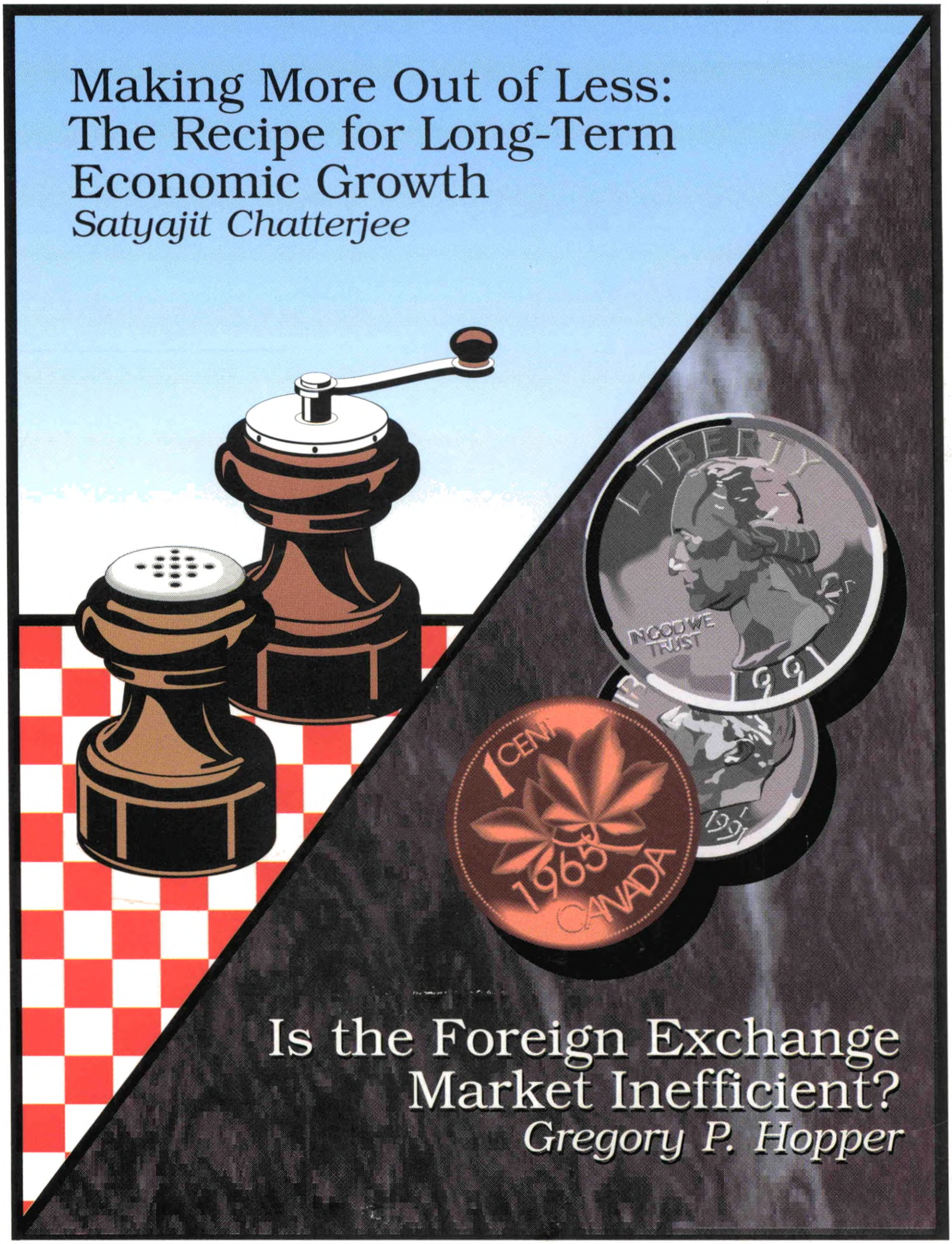
Business Review

Federal Reserve Bank of Philadelphia

May • June 1994 ISSN 0007-7011



Business Review

The BUSINESS REVIEW is published by the Department of Research six times a year. It is edited by Sarah Burke. Artwork is designed and produced by Dianne Hallowell under the direction of Ronald B. Williams. The views expressed here are not necessarily those of this Reserve Bank or of the Federal Reserve System.

SUBSCRIPTIONS. Single-copy subscriptions for individuals are available without charge. Institutional subscribers may order up to 5 copies.

BACK ISSUES. Back issues are available free of charge, but quantities are limited: educators may order up to 50 copies by submitting requests on institutional letterhead; other orders are limited to 1 copy per request. Microform copies are available for purchase from University Microfilms, 300 N. Zeeb Road, Ann Arbor, MI 48106.

Please send subscription orders, back orders, changes of address, and requests to reprint to Publications, Federal Reserve Bank of Philadelphia, Department of Research and Statistics, Ten Independence Mall, Philadelphia, PA 19106-1574, or telephone (215) 574-6428. Please direct editorial communications to the same address, or telephone (215) 574-3805.

MAY/JUNE 1994

From 1950 to 1991, the real value of output produced by an hour's work rose steadily in the United States. This increase in productivity is the hallmark of economic growth. But what factors caused this growth in productivity? The answer to that question must take into account the increase in the amount of capital used by each worker and technical progress. In basic terms, however, Satyajit Chatterjee shows that economic growth is the result of the human ability to make more out of less.

If interest rates for government securities in the Canadian market are at 5 percent, but they're only 4 percent in the U.S. market, does it make sense to invest in the higher-rate securities? In other words, are there exploitable profit opportunities in the foreign exchange market? Some economists say yes; others say no. Greg Hopper looks at the arguments on both sides and discusses factors such as statistical problems, rational expectations, and risk premia.

Making More Out of Less: The Recipe for Long-Term Economic Growth

*Satyajit Chatterjee**

References to the pace of economic growth in the United States and elsewhere are commonplace in the news media. However, it's much less common to find informed discussions of the forces that shape the economic growth of nations. This article describes the salient facts gathered by economists on the sources of economic growth in the United States and other countries. Although these facts do not cover every aspect of this complex phenomenon, they do shed considerable light on the mainsprings of economic growth and on the kinds of growth-related government policies that might prove beneficial to the economy.

The phenomenon of economic growth has many aspects, but the central one is that the real value of output produced by an hour's work in the U.S. has risen over the years: in 1991, the value of output for an hour of work was about twice the value of output for an hour's work in 1950 (Figure 1). This increase in productivity or in the economic worth of work-time is the hallmark of economic growth. The question to which we seek an answer is: what are the main reasons for increases in productivity? The key parts of the answer are an increase in the amount of capital used by each worker and technical progress.

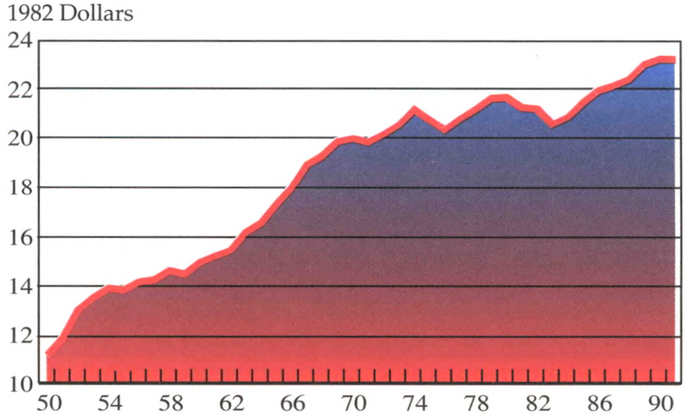
* Satyajit Chatterjee is a senior economist in the Research Department of the Philadelphia Fed.

SOURCES OF GROWTH IN OUTPUT PER HOUR WORKED

Economists approach the question of the sources of economic growth by relating the total value of output produced per hour to variations in the use of the two primary factors of production: labor-time and the capital stock. Since these terms will have specific meaning in this article, it's best to begin by defining what they mean.

FIGURE 1

Output per Hour Worked 1950 - 1991



Labor-time means the total amount of time, say, in a year, that members of a nation spend in the production of goods and services valuable to households and firms. In principle, it includes the total time spent earning a living (through honest means!) and also the total time spent doing chores around the house. Since measuring the time spent on housework is difficult, the practical definition of labor-time (and the one that we will use) is the time spent in the production of goods and services for the marketplace and in the production of government services.[1]

Capital stock means the vast stock of every kind of building, structure, and machinery used in conjunction with labor-time. It includes all factories and office buildings and the equipment therein, as well as infrastructure: facilities such as roads, railroads, bridges, canals, harbors, and docks. This enormously diverse collection of man-made things can be measured only in terms of its total economic or monetary value. Therefore, the capital stock of a nation means the monetary value of all the buildings, structures, and machinery used in conjunction with labor-time.

How, then, might variations in labor-time and the capital stock explain variations in output per hour? As an extreme but illuminating example, consider how a building is constructed in the United States versus how one is constructed in a country like India. In the United States, a construction worker has at his disposal a large array of sophisticated tools: pneumatic drills, jackhammers, electric screwdrivers, forklifts, cranes, and all kinds of heavy earth-moving equipment to help carry out the building tasks. In contrast, a construction worker in India works with nothing more than ordinary hammers, chisels, and shovels. As a result, if the construction of a similarly sized building is to be completed in the same amount of time, the number of construction workers needed for the job in India will be many times that needed in the United States. More generally, the number of hours of work needed on a building will be much greater in India than in the United States. If the economic value of the

[1]Similarly, the measure of national output should, in principle, include the monetary value of the work done around the house. However, because of the difficulty of constructing such a measure, the value of housework is excluded from official estimates of national output.

building is the same in both countries, the value of output produced per hour of work will be much lower in India than in the United States, i.e., the productivity of the Indian construction worker is a fraction of the productivity of the American construction worker.

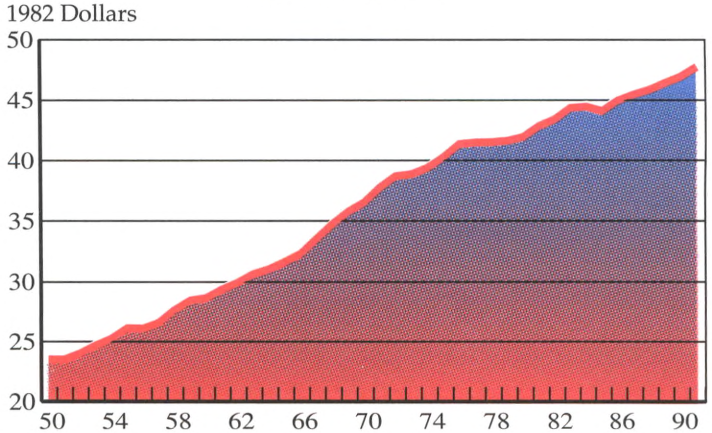
It should be clear why this is the case. An American construction worker is assisted by more capital stock than an Indian construction worker: the dollar value of pneumatic drills, jackhammers, electric screwdrivers, forklifts, cranes, and heavy earth-moving equipment used in the U.S. is many times larger than the dollar value of hammers, chisels, and shovels used in India. In other words, one important reason why productivity in the American construction industry is so much higher than in the Indian construction industry is that the capital stock used per hour of work (the capital-labor ratio, as economists call it) is much higher in the United States than in India.

What is true of India and the United States today is also true, in less extreme form, of the United States of 1950 and the United States of today. Measurements of the U.S. capital stock reveal that capital stock per hour worked has risen over the years (Figure 2). From this increase in the capital stock per hour of work, we should expect productivity to be higher in 1991 than in 1950. A natural question is: is the increase in capital stock per hour worked the sole factor, or even the most important factor, underlying the economic growth of the United States?

In the mid-1950s, several American economists attempted to answer this question for the United States. Moses Abramovitz (1956), John Kendrick (1956), and Robert Solow (1957) examined historical data on national output per hour worked and the stock of tangible physical capital per hour worked to see if increases in the capital stock per hour worked were the dominant factor driving U.S. economic growth. In perhaps the most well known of these studies, Robert Solow calculated the annual percentage change in national output per hour worked and capital stock per hour worked for the 1909-1949 period. To complete his calculations he needed to know how much a 1 percent increase in the capital stock per hour worked would increase output per hour. He observed that over this period the owners of capital received about 36 percent of national income (the rest being payment to labor), which suggested that, on average, the capital stock contributed 36 percent of the total output of the economy. With this in mind, Solow concluded that a 1 percent increase in the capital stock per hour worked would raise output per hour by 0.36 percent. Using this estimate of the contribution of capital stock to national output, Solow discovered that growth in capital stock per hour worked accounted for less than 15 percent of the increase in national output per hour worked over

FIGURE 2

Capital Stock per Hour Worked 1950 - 1991



this period. Moses Abramovitz, who used the same methodology for the period 1870 to 1953, found that increases in capital stock per hour worked accounted for only 14 percent of the increase in national output per hour worked over this period. In other words, an astonishing 85 percent of the total economic growth over the 80-odd years beginning in 1870 appeared to be caused by factors other than an increase in the capital stock per hour worked!

What else could be contributing to economic growth? Let's go back to the comparison between the American and the Indian construction worker. While it's true that the dollar value of the tools that the American worker has at his disposal is much greater than the dollar value of the tools at the disposal of the Indian worker, it's not merely the case that the American worker has *more* of the *same* tools than the Indian worker. There's a clear qualitative difference between a bulldozer and a shovel, although both are used to move soil. In other words, the higher dollar value of capital stock per hour worked in the United States is also associated with a superior construction technology. Similarly, superior technology also accompanies the larger capital stock per hour worked in 1953 relative to 1870, and output per hour worked would be higher in 1953 on this count as well. Robert Solow made the bold suggestion that technical progress explained the remaining 85 percent of economic growth.

Technical progress makes any given dollar-value increase in the capital stock per hour worked more effective in generating additional output, or, conversely, it allows any given increase in national output to be attained with less of an increase in capital stock per hour worked (which might mean less capital or less labor-time or both). This increase in output per hour worked due to technical progress is called an increase in *total factor productivity* (TFP).[2]

The quantitative importance of TFP growth in accounting for U.S. economic growth over the period 1870-1953 proved to be a phenomenon that cut across national boundaries and particular histories. Edward Denison's well-known study of the U.S. and eight European countries for 1950-1962 shows that TFP growth was by far the most important source of growth in national output (Table). For these countries, its contribution is never below 64 percent (U.S.) and reaches as high as 86 percent (Germany).[3]

We can plot the advance of TFP in the United States during 1950-1991 (Figure 3). The top line is the level of output per hour worked in the U.S and is the same line that appears in Figure 1. The lower line shows the level of output per hour worked that would have been attained if TFP had remained at its 1950 level; its rise is solely the result of the increase in capital stock per hour worked. Thus, the gap between the two lines is a measure of how much the level of output per hour worked owes to increases in TFP.

While the Table and Figure 3 display the importance of TFP growth in accounting for growth in national output, they understate its importance as a causal factor. According to Solow (1956), the reason is that capital stock per hour worked grows partly in response to an

[2]The use of the term "TFP growth" rather than "technical progress" partly reflects the controversy that followed Solow's suggestion that technical progress accounted for the unexplained growth in output. At that time, many economists were unconvinced that technical progress was the key to the unexplained growth in output per hour worked. Hence, they adopted the more neutral term, TFP growth. Since the 1950s, however, numerous authors have shown that a large portion of the unexplained growth in output can be accounted for by careful measurements of the improvements in the quality of labor-time and capital stock (see, for instance, Maddison, 1987). Since improvements in the quality of inputs is an aspect of technical progress, Solow's suggestion stands vindicated.

[3]Denison, unlike Solow, measured labor input by the number of persons employed rather than total hours worked. However, since he adjusted his figures for changes in the average number of hours worked by a person, this difference is unimportant.

TABLE

Country	Growth Rate of National Output per Person Employed (1950-62)	The percentage contribution of	
		TFP	Capital per Person Employed
United States	2.15	64	36
Belgium	2.64	75	25
Denmark	2.56	74	26
France	4.80	75	25
Germany	5.15	86	14
Holland	3.65	75	25
Italy	5.36	80	20
Norway	3.27	74	26
United Kingdom	1.63	74	26

Source: Adapted from Tables 21-2 to 21-20 in Edward F. Denison: *Why Growth Rates Differ*, Washington, D.C.: Brookings Institution, 1967.

increase in TFP: as advances in technology make labor and capital more productive, firms exploit the increase in TFP by investing in newer and better buildings, structures, and equipment. In other words, growth in capital stock per hour worked is never an entirely *independent* factor contributing to economic growth: part of the growth in a nation's capital stock per hour worked occurs to keep pace with increases in its TFP.

Is there a way to tell how much of the increase in capital stock per hour worked occurs in response to increases in TFP and how much of it occurs for other reasons? Solow pointed out that if a nation's capital stock per hour worked rose in response to factors other than an improvement in TFP, the percentage increase in national output per hour worked would be only 0.36 of the percentage increase in capital stock per hour worked, so that the *ratio* of capital stock per hour worked to output per hour worked would rise. In contrast, if capital stock per hour worked rose in response to an increase in TFP, the capital-to-output ratio need not rise because of the added effect of higher TFP on output per hour worked. For the United States the ratio of capital to output has remained roughly constant, suggesting that capital accumulation in the U.S. has largely been in response to increases in TFP.[4] Also, capital-output ratios are observed to be roughly constant as well for the European countries examined by Denison. Thus, increases in TFP have been the single most important factor driving economic growth in these countries. In short, the human ability to make more out of less is at the heart of economic progress.

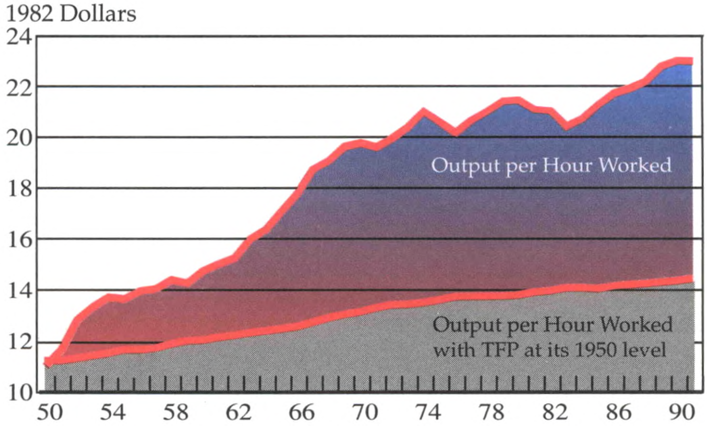
DETERMINANTS OF TFP GROWTH

Inventions. Inventing new products or new ways to make old products accounts for one of the primary sources of TFP growth in industri-

[4]To state this point differently, if capital stock per hour worked rose without an accompanying increase in TFP, diminishing returns to capital would cause the output-capital ratio to fall. For the United States, the fact that the capital-output ratio has remained constant in the face of growing capital stock per hour means that a continuous rise in TFP has offset the force of diminishing returns. In Solow's growth model, a steady increase in TFP stimulates firms to raise the capital stock per hour and, by simultaneously increasing household income, also provides the resources to finance this accumulation. Output and capital stock grow together at a rate equal to the growth rate of TFP, and thus the capital-output ratio remains constant.

FIGURE 3

Growth in TFP 1950 - 1991



alized countries. Each of these activities makes labor and capital more productive, although for somewhat different reasons. An improved process for making an existing product directly enhances productivity by economizing on labor and capital or on intermediate inputs such as raw materials and energy. A new product improves productivity indirectly by drawing away labor and capital from less valuable uses, thereby enhancing the value of goods and services produced by existing amounts of labor and capital.

In light of the international differences in TFP growth evident in the Table, an important issue is whether the pace of inventions in a country is related to the amount of resources it expends on inventive activity.

Zvi Griliches and Ariel Pakes (1984) studied the correlation between the research and development (R&D) expenditures of large U.S. firms and the number of patents issued to them each year. They found that the quantity of patents issued yearly to different firms was positively and strongly related to the level of R&D expenditures in that firm: firms that spent more on R&D generated more patents, on average.

However, all patents are not equally valuable, and not all inventions are patented. To make a more compelling case for the potency of R&D expenditures, it's necessary to examine its correlation with firm-performance measures such as the market value of common stock. Ariel Pakes (1985) documented a positive correlation between companies' expenditures on R&D and their stock value. Since many other factors besides the outcome of R&D efforts affect stock value, the correlation found is naturally not as strong as that between R&D and patents. Nevertheless, there's enough evidence to suggest that R&D expenditures generate valuable inventions for firms. Also, Saul Lach and Mark Schankerman (1989) and Saul Lach and Rafael Rob (1992) have shown that firms and industries that spend more on R&D also tend to spend more on equipment and machinery in future periods, which suggests that spending on R&D does generate profitable investment projects.

Therefore, at least for the U.S., evidence shows that the rate of TFP growth is influenced by the level of R&D expenditures. Nevertheless, this influence is imperfect because the outcome of R&D efforts involves a substantial amount of randomness. Also, as John Bound and associates (1984) document, many firms that generate patents do not report any significant R&D expenditures, perhaps because many inventions do not arise from directed R&D efforts but happen simply because individuals think up improvements in the course of doing their jobs. The R&D expenses for such accidental inventions are probably minor.

The existence of patented inventions that were apparently generated at little or no cost serves to remind us that firms do not explicitly pay for some of the most critical inputs in the R&D process. The worker who comes up with an innovative suggestion is surely using the basic skills and knowledge taught in schools and the experience acquired in previous jobs. Indeed, a portion of the worker's wage must be compensation for the valuable knowledge and experience brought to the job. Therefore, we can be reasonably confident that the average education level of the working population—education conceived broadly as years of work experience and formal schooling—also contributes to a nation's TFP growth, although this contribution may be difficult to quantify.

Recognizing that TFP growth may respond to the level of R&D expenditure and the average level of education, economists have begun to examine whether international differences in R&D expenditures and education levels help explain international differences in TFP growth. Since data on productivity are hard to compile, researchers have been content to study whether international differences in R&D expenditures and education levels help explain international differences in per capita GDP growth.

Two authors, Paul Romer (1989) and Frank Lichtenberg (1992), concluded that international differences in per capita GDP growth are influenced by international differences in R&D expenditures. However, a recent study by Nancy Birdsall and Changyong Rhee (1993) found that correcting the data set used by Romer and Lichtenberg for possible measurement errors eliminates this correlation. Therefore, the role of R&D expenditures in the explanation of international differences in TFP growth rates remains an open issue.[5] With regard to education levels, Robert Barro (1991) has documented that countries that start out with low school-enrollment rates grow at a discernably slower rate.

Economies of Scale. A second important determinant of TFP growth is economies of scale. Economies of scale exist if unit costs fall at higher levels of production. Increases in TFP that result from increases in the size or capacity of production facilities represent one type of economies of scale. For instance, in the chemical and petroleum industries the average cost of production is much lower in bigger plants if these plants operate near capacity.[6]

The increase in TFP resulting from economies of scale is distinct from that resulting from a new industrial process. A firm may be aware that a bigger plant, if operated near capacity, lowers costs but may not build or buy a bigger plant if the larger volume of output needed to make it economical cannot be profitably sold. However, with an expansion in market size, bigger plants do become practical and are acquired, thereby contributing to the increase in the productivity of labor and capital.

A second example of economies of scale is the increase in TFP that comes from specialization. Adam Smith in his *Wealth of Nations* noted that the division of labor constitutes a major source of productivity increase. A group of workers in which each worker concentrates on a limited set of tasks produces much more than an equally large group of workers in which each worker performs every task. This gain stems from several factors: less time is lost

[5]R&D expenditures may matter for growth, but differences in R&D spending across countries may not be large enough to appreciably affect productivity growth. Alternatively, a country that invents a new product or process may not be the only one to benefit from it. If inventions are patented and sold to producers in other countries, the purchasing country will also experience an increase in TFP. Indeed, Robert Evenson (1984) documents a brisk international trade in inventions. Therefore, it might be expected that differences in the rate of economic growth due to differences in R&D levels would be difficult to discern in the data.

[6]See Alfred Chandler's book *Scale and Scope: The Dynamics of Industrial Capitalism*, for an in-depth discussion of the role of increasing returns to scale in modern industries.

switching from one task to another; some workers may be better at some tasks than other workers; or simply performing a task many times over entails a gain of efficiency.

The efficiency gain from specialization of tasks within a firm extends to the specialization of production across firms as well: if production is organized so that a large number of firms produce very specialized products (which are then assembled to manufacture the final good), the productivity of labor and capital will be higher. Indeed, a proliferation of intermediate goods used in the production of goods and services always accompanies the industrial development of any country, which suggests that some of the increase in TFP must be due to gains from specialization. Again, a firm may not wish to incur the expenses involved in buying or building more specialized plants if the increased volume of its specialized output cannot be profitably sold in the market.

The economic integration of geographically dispersed markets is perhaps the most significant channel through which economies of scale contribute to the growth of TFP. When regions that did not previously trade with each other begin to do so, market size for producers in both regions expands, making it possible for more and more firms to profitably adopt bigger plants and to profitably specialize.

The integration of markets can come about for various reasons. For instance, the proliferation of railroads contributed to the higher pace of adoption of large-scale production facilities in the U.S. during the 1880s and 1890s. Similarly, the development of the U.S. interstate highway system in the 1950s and 1960s integrated markets even further. More generally, innovations in the transportation sector that reduce the costs of moving goods around have the effect of expanding the geographic size of markets. Removal of trade barriers such as tariffs promotes trade between regions and is another channel through which the geographic size of markets may increase. Edward Denison estimates that the portion of TFP growth that can be accounted for by economies of scale ranges between 23 to 36 percent.

Learning-by-Doing. Economists have identified a third source of TFP growth: learning on the job, or learning-by-doing. As individuals working together in a factory gain experience in the production of a new product or process, they learn to become more efficient, that is, they waste less time and raw materials in producing a given volume of output. Consequently, TFP increases simply as a result of experience. The existence of learning-by-doing is well documented for many manufacturing industries (see the article by Linda Argote and Dennis Epple). As one example, let's consider Alan Searle's 1945 study of the manufacture of Liberty Ships during World War II.

From December 1941 through December 1944, 14 shipyards in the U.S. produced a total of 2458 Liberty Ships, all to the same standardized design. On average, with each doubling of cumulative output, the reduction in manhours required per ship ranged from 12 to 24 percent across the 14 shipyards. Similar reductions in unit manhour requirements were also seen in the production of other ships as well. Leonard Rapping (1965) showed that after accounting for variations in labor hours and capital used in each of the shipyards, the effect of learning was to increase TFP between 11 and 29 percent over the three-year period, i.e., to increase TFP at an annual rate between 4 and 10 percent.

While learning effects can be quite substantial over a two- to three-year period, industry studies also show that TFP growth from learning ultimately stops. For instance, in the case of Liberty Ships, the maximum productivity gain had been achieved by the end of 1943, and productivity was roughly constant over the last year of production. This raises a natural question: can learning-by-doing really be a source of *sustained* increases in TFP? The answer may be yes because, as already noted, new products and new processes get added every year, so that there are fresh opportunities for learning effects to increase TFP.[7]

Indeed, some economists have recently conjectured that the effects of learning-by-doing may be an important part of the explanation of the fast economic growth of countries like Hong Kong, Korea, and Taiwan since the mid-1960s.[8] These countries have not been technological leaders, so their rapid economic growth cannot be ascribed to a rapid pace of invention. Also, although benefits from specialization exist, researchers do not consider them a major factor in the economic growth of these countries. On the other hand, these countries have rapidly added the production of more and more technologically advanced goods to their economy and have certainly been in a position to reap the TFP gains from learning-by-doing. In this they have (somewhat paradoxically) been helped by the fact that they are technological followers: they haven't had to await the outcome of costly R&D efforts to obtain "new" products.

To summarize the discussion so far, economists have identified three important determinants of TFP growth. First, TFP increases because of the invention of new products and processes. Second, it increases because of economies of scale. Third, it increases because individuals in firms learn at their jobs. The degree to which each of these sources contributes to TFP growth depends on the choices that individuals, firms, and governments make: the pace of innovation depends on the amount of resources spent on R&D and education; the importance of economies of scale depends on the speed with which the transportation network of a country develops; the extent to which a country benefits from learning-by-doing depends on how quickly a country expands the production of new goods. Thus, a country's TFP growth is shaped by the choices that its citizens make.

ECONOMIC POLICY AND TFP GROWTH

For an economist, the first and most important issue about economic policy and TFP growth is whether the government should attempt to influence TFP growth at all. Is there reason to believe that individuals and corporations acting in their own interests in this regard do not fulfill the broader interests of society? Is there a serious mismatch between private gain and social benefit in the generation of productivity improvements? This section explores some of the justifications for government intervention with regard to the three sources of TFP growth discussed in the previous section.

Economists recognized early on that the economics of technical progress had some peculiar features to it. Fundamentally, technical progress depends on our understanding of the physical universe; it draws upon the fruits of basic research in the various sciences, including medicine. However, basic research cannot be done for profit because scientists do not have property rights on the laws of nature. Once a scientific discovery is communicated to the scientific community, anyone can use it free of charge. Therefore, the government should, by and large, support basic research. Of course, this raises very thorny issues about the kinds of basic research to fund, whether the level of funding in any year is too little or too much, and exactly how to measure the benefits of past funding on basic research. However, these largely unexplored issues are beyond the scope of this article.

In contrast to basic research, the fruits of applied research are new products that can be sold for profit and new production processes that can lower costs. Therefore, applied re-

[7]The gains from learning-by-doing may be limited for a new product or process if the product or process is not produced to the same specifications each time but is customized to a significant degree.

[8]See, for instance, articles by Robert Lucas (1993), Nancy Stokey (1988), and Alwyn Young (1991).

search can, in principle, fund itself. However, even here there may be a mismatch between private return and social benefits. The mismatch arises because, in many instances, the discovery of a new product or production process stimulates the discovery of other products and processes elsewhere in the economy. If the new discoveries are closely related to the first, patent laws will allow the first discoverer to benefit monetarily from these unexpected and unintended consequences of his discovery provided the original invention was patented. However, more often than not, the discoverer cannot capitalize on these subsequent inventions.

Scholars who have attempted to measure these external benefits of research and development have generally found them to be significant and quite pervasive.[9] Thus, corporations and individuals, guided by the value of inventions to themselves alone, will undertake less R&D than is warranted by the true social benefit of their R&D efforts. Therefore, a case can be made for subsidizing applied commercial research. Again, the issue of subsidies raises difficult questions about which kinds of research should be subsidized, how the benefits from past subsidies should be evaluated, how the incentives for R&D should be structured, and how much subsidy should be provided. The U.S. government subsidizes applied commercial research through government grants and tax breaks, but economists have only an imperfect understanding of these issues.[10]

As noted earlier, the average level of education of the working population is also likely to be an important factor in promoting TFP growth. The government has an obligation to support education for reasons similar to those for supporting R&D. While most individuals value education and are willing to devote time and money to acquire it (or make sure that their family members do), they consider only their private gain. However, education confers innumerable benefits on society as a whole (faster TFP growth is one), so there is a case for subsidizing it. However, as in the case of R&D subsidies, many unresolved issues remain concerning the details of government support for education.[11]

For developing countries, the possibility of achieving TFP growth through economies of scale suggests a special rationale for government intervention. In a nutshell, the rationale stems from the fact that cost-effective large-scale production of one product typically hinges on the cost-effective large-scale production of constituent inputs, and these inputs in turn require large-scale production of other inputs and so on. If a developing country is to replicate this interlocking pattern of large-scale industrial production within a short period of time, it has to advance simultaneously across a broad industrial front. Therefore, an individual firm contemplating investment in a large-scale technology must be reasonably confident that supporting investments in large-scale technologies will occur in other industries.[12] In such a situation, the government can play a vital coordinating role by assisting firms in different sectors of the economy to commit to a common industrial plan. Virtually all developing countries have relied on such coordination of economic activ-

[9]See, for instance, the articles by Adam Jaffe (1986) and Ricardo Caballero and Adam Jaffe (1993).

[10]For an attempt to come to an understanding of this knotty issue, see Linda Cohen and Roger Noll's book *The Technology Pork Barrel.*

[11]For a discussion of the benefits of education, with explicit reference to economic growth and some of the related policy issues, see the article by T. Paul Schultz (1988).

[12]Another option for the firm is to buy the necessary inputs from abroad. However, developing countries often find the domestic price of imported inputs to be prohibitively high so that this option may not be practical.

ity, although its extent and scope have varied greatly.

Other than financing the transportation network, there are no U.S. policies designed to directly affect the advance of TFP through economies of scale. However, some policies inadvertently do so, for example, anti-trust laws and international trade rules. To reap the TFP benefits of economies of scale, firms must enlarge the scope of their operations, which, in many instances, means that a few large firms will have a major share of the market being served.[13] However, economic theory suggests that firms (or groups of firms) that are so big as to have no effective competition from their rivals tend to cut back on the supply of their product so that the artificial scarcity can generate hefty profits for the firm or the group. For this reason, the U.S. legislated anti-trust laws (the Sherman Act of 1890) that prohibit price-fixing agreements and mergers and acquisitions whose main intent is to gouge customers. However, while these laws protect consumers from errant firms, they also slow down the rate of adoption of large-scale technologies because expansions in firm size that typically accompany such adoptions need to be cleared by regulatory authorities. There are no estimates of the adverse impact of anti-trust regulation on TFP growth, but such an impact surely exists.[14] On the other hand, policies to break down barriers to international trade promote larger markets, thus allowing greater economies of scale and TFP growth.

The phenomenon of learning-by-doing *per se* does not suggest a role for government policy. While it is true that uncertainties and setbacks faced during the learning phase might mean that a firm doesn't survive to reap the benefit from learning-by-doing, it does not follow that governments should step in to help out failing firms. If private investors are not willing to risk their money in the venture, why should the government risk the taxpayers' money? Government help is justified only if surviving firms provide benefits (to the economy) for which private investors are not compensated, thereby causing private investment in new ventures to be too low. However, learning-by-doing presumably invests workers and entrepreneurs with skills that are also valuable outside of their existing firm, so that uncompensated benefits from learning-by-doing may well exist. For instance, a worker who breaks away and pioneers a valuable innovation after acquiring useful training in a firm is not obligated to share his newfound wealth with his former employers. If this kind of phenomenon is pervasive, it may be beneficial for governments to subsidize firms during the costly learning-by-doing phase.[15]

CONCLUSIONS

This article makes several important points about economic growth. First, technical progress, which economists call total factor productivity (TFP) growth, has been a major factor underlying increases in output per hour worked in the United States and Europe. In fact, it has been more important than the rising amount of capital available to each worker, although this, too, has contributed.

Second, the constancy of capital-output ratios in the United States and many European

---

[13]Indeed, as Chandler's book cited earlier documents, the adoption of large-scale technologies in the last two decades of the nineteenth century went hand in hand with the emergence of monopolies, trusts, and combines.

[14]See Crandall (1980) for a general discussion of government regulations on U.S. productivity growth. Denison (1979) estimates that 13 percent of the decline in productivity growth over the 1965-1978 period was due to increased government regulations.

[15]At least one astute observer of economic life believes that innovators are frequently breakaways. See Jane Jacobs' *The Economy of Cities.*

countries suggests that accumulation of capital stock in these countries has been mostly in response to their TFP growth. Thus, it is reasonable to think that *differences* in growth of output per hour worked across these countries are mostly the result of *differences* in TFP growth.

Third, the TFP growth experienced by a country is a result of the pace of innovations and inventions, the extent of the economies of scale experienced by the country, and the extent of productivity improvements from learning-by-doing. These sources of TFP growth in turn depend on profit calculations of individuals and firms and the wisdom of government policies. Therefore, there is a potential economic explanation for the differences in TFP growth across countries and for the same country over time. However, despite these advances in our understanding of economic growth, there's a great deal still to be learned, especially in the area of designing economic policies to promote productivity growth.

REFERENCES

Abramovitz, Moses. "Resource and Output Trends in the United States Since 1870," *American Economic Review,* Papers and Proceedings (1956), pp. 5-23.

Argote, Linda, and Dennis Epple. "Learning Curves in Manufacturing," *Science,* 247 (1990), pp. 920-24.

Barro, Robert. "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics,* 106 (1991), pp. 407-43.

Birdsall, Nancy, and Changyong Rhee. "Does R&D Contribute to Economic Growth in Developing Countries?" Rochester Center for Economic Research, Working Paper 359 (1993).

Bound, John, and others. "Who Does R&D and Who Patents?" in Zvi Griliches, ed., *R&D, Patents and Productivity.* National Bureau of Economic Research Conference Report, University of Chicago Press, 1984.

Caballero, Ricardo, and Adam Jaffe. "How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," in Olivier Blanchard and Stanley Fischer, ed., *NBER Macroeconomics Annual.* Cambridge, MA: The MIT Press, 1993.

Chandler, Alfred D. Scale and Scope: *The Dynamics of Industrial Capitalism.* Cambridge, MA: Harvard University Press, 1990.

Cohen, Linda, and Roger Noll. *The Technology Pork Barrel.* Washington D.C.: The Brookings Institution, 1991.

Crandall, Robert. "Regulation and Productivity Growth," in *The Decline in Productivity Growth.* Boston, MA: The Federal Reserve Bank of Boston, 1980.

Denison, Edward F. *Why Growth Rates Differ,* Washington, D.C.: The Brookings Institution, 1967.

Denison, Edward F. "Explanations of Declining Productivity Growth," *Survey of Current Business,* 59 (1979), pp. 1-24.

Evenson, Robert. "International Inventions: Implications for Technology Market Analysis," in Zvi Griliches, ed., *R&D, Patents and Productivity.* National Bureau of Economic Research Conference Report, University of Chicago Press, 1984.

Griliches, Zvi, and Ariel Pakes. "Patents and R&D at the Firm Level: A First Look," in Zvi Griliches, ed., *R&D, Patents and Productivity.* National Bureau of Economic Research Conference Report, University of Chicago Press, 1984.

Jacobs, Jane. *The Economy of Cities.* New York: Random House, 1969.

REFERENCES (continued)

Jaffe, Adam B. "Technological Opportunity and Spillovers of R&D: Evidence From Firms'Patent, Profits, and Market Value," *American Economic Review,* 76 (1986), pp. 984-1001.

Kendrick, John W. "Productivity Trends: Capital and Labor," National Bureau of Economic Research, Occasional Paper 53 (1956).

Lach, Saul, and Mark Schankerman. "Dynamics of R&D and Investment in the Scientific Sector," *Journal of Political Economy,* 97 (1989), pp. 880-904.

Lach, Saul, and Rafael Rob. "R&D, Investment and Industry Dynamics," National Bureau of Economic Research, Working Paper 4060 (1992).

Lichtenberg, Frank L. "R&D Investment and International Productivity Differences," National Bureau of Economic Research, Working Paper 4161 (1992).

Lucas, Robert E. Jr. "Making a Miracle," *Econometrica,* 61 (1993), pp. 251-72.

Maddison, Angus. "Growth and Slowdown in Advanced Capitalist Economies: Techniques of Quantitative Assessment," *Journal of Economic Literature,* 25 (1987), pp. 649-98.

Pakes, Ariel. "On Patents, R&D, and the Stock Market Rate of Return," *Journal of Political Economy,* 93 (1985), pp. 390-409.

Rapping, Leonard. "Learning and World War II Production Functions," *Review of Economic and Statistics,* 47 (1965), pp. 81-86.

Romer, Paul. "What Determines the Rate of Growth and Technological Change?" The World Bank, Working Paper 279 (1989).

Searle, Alan D. "Productivity Changes in Selected Wartime Shipbuilding Programs," *Monthly Labor Review,* 61 (1945), pp. 1132-47.

Schultz, T. P. "Education Investments and Returns," in Hollis Chenery and T.N. Srinivasan eds., *Handbook of Development Economics,* Vol. 1. Amsterdam: North Holland Publishers, 1988.

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

Solow, Robert M. "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics,* 39 (1957), pp. 312-20.

Stokey, Nancy L. "Learning by Doing and the Introduction of New Goods," *Journal of Political Economy,* 96 (1988), 701-17.

Young, Alwyn. "Learning by Doing and the Dynamic Effects of International Trade," *Quarterly Journal of Economics,* 106 (1991), 396-406

Is the Foreign Exchange Market Inefficient?

*Gregory P. Hopper**

Suppose you are in charge of investments for your company and you have 1 million U.S. dollars to invest for one month. You want to obtain the highest return possible for the month while assuming minimal risk, so you decide to invest in short-term government securities: Treasury bills. The rate of interest paid on U.S. Treasury bills maturing in one month is currently 4 percent annually. However, while reading the newspaper, you notice that Canadian Treasury bills maturing in one month are currently paying 5 percent annually. Why not sell the 1 million U.S. dollars for Canadian dollars at the present exchange rate, invest the proceeds in Canadian bills, and earn the 5 percent interest rate? Then, at the end of the month, convert the Canadian dollars back to U.S. dollars.

You tell your broker about your strategy, but he objects. "The foreign exchange market is efficient," he argues. "That means that investors eliminate exploitable profit opportunities. Interest rates are always different between countries. If higher interest rates in a foreign country really meant higher returns after taking expected exchange rate movements into account, in-

* Gregory Hopper is an economist in the Research Department of the Philadelphia Fed.

vestors would have recognized that years ago. They would have moved funds from one country to another to capture those higher extra returns, making interest rates converge in the process. So, this extra return cannot really be available."

"But I can earn an extra 1 percent interest in the Canadian Treasury bill market," you protest. "Why isn't that an extra 1 percent return?" "Simple," replies your broker. "You have to convert the Canadian dollars back to U.S. dollars in a month, right? Given the current level of interest rates and the exchange rate, the market must expect that during the coming month, the Canadian dollar will lose value in terms of U.S. dollars at a 1 percent annual rate. That way, you lose the extra 1 percent return when you convert the Canadian dollars back into U.S. dollars. Besides, you can't be sure what the exchange rate will be when you convert back to dollars; so you will assume a lot of exchange rate risk."

Unconvinced, you decide to see what would have happened had you followed this strategy in the past. To do this, you look at monthly U.S.-Canadian interest and exchange rate data over the period June 1973 to April 1993. For each month, you would have invested $1 million in one-month U.S. T-bills whenever their interest rate was higher than that on one-month Canadian T-bills. However, in the months when the Canadian T-bill interest rate exceeded the U.S. T-bill interest rate, you would have converted $1 million into Canadian dollars and invested the sum in Canadian T-bills; at the end of the month, you would have converted the accumulated Canadian dollars back into U.S. dollars. Over this period, there were 170 months in which you would have made a Canadian T-bill investment. Surprisingly, you find that this strategy would have made an average $1072 per month in excess returns during the months you invested in Canadian T-bills. The broker is right about the exchange rate risk, though. Because of the volatility of the U.S.-Canadian dollar exchange rate, 26 percent of the time you would have lost more than $5000 per month during the months in which you invested in the Canadian T-bill (Figure 1).[1] Sometimes, the risk would have been quite large: during some months you would have lost approximately $40,000 per month, and in one particularly bad month, you would have lost almost $60,000. Thus, the average $1072 per month in extra returns involves substantial risk. Even though the risk-return tradeoff is not very good, do the extra returns mean that the foreign exchange market is inefficient?

In this article, we will consider this question. Some economists argue that statistical problems falsely make it look like the extra returns are there. Other economists do not deny that the extra returns exist: one group claims the extra returns are available because the market's expectation of the worth of future currencies is irrational; another group maintains that the extra returns can be explained as compensation to the investor for taking on the risk of losing money. Ultimately, as we will see, economists have not yet reached agreement; thus, we may not rule out the possibility that opportunities for extra returns do exist in the foreign exchange market.

EVIDENCE AGAINST MARKET EFFICIENCY

If markets are efficient, then when the annual foreign interest rate is x percentage points above the domestic interest rate, the foreign currency is expected to decline in value at an annualized rate of x percent. If these expectations are borne out on average, over time the extra x percent interest will be offset by the currency's fall in value. But historically, these expectations are not upheld: when foreign interest rates rise above U.S. rates, the foreign

[1]Similar results arise for the short-term debt issued by the governments of other major industrial countries.

currency tends to rise in value rather than fall. Moreover, when the U.S. interest rate rises above the foreign interest rate, the U.S. dollar tends, on average, to rise rather than fall in terms of the foreign currency.[2] These results suggest a profit-making strategy for the investor: if the investor always puts his funds into the short-term government securities of the major industrial country that pays the highest interest rate, he should make extra returns over time, calling into question the efficiency of the foreign exchange market.[3]

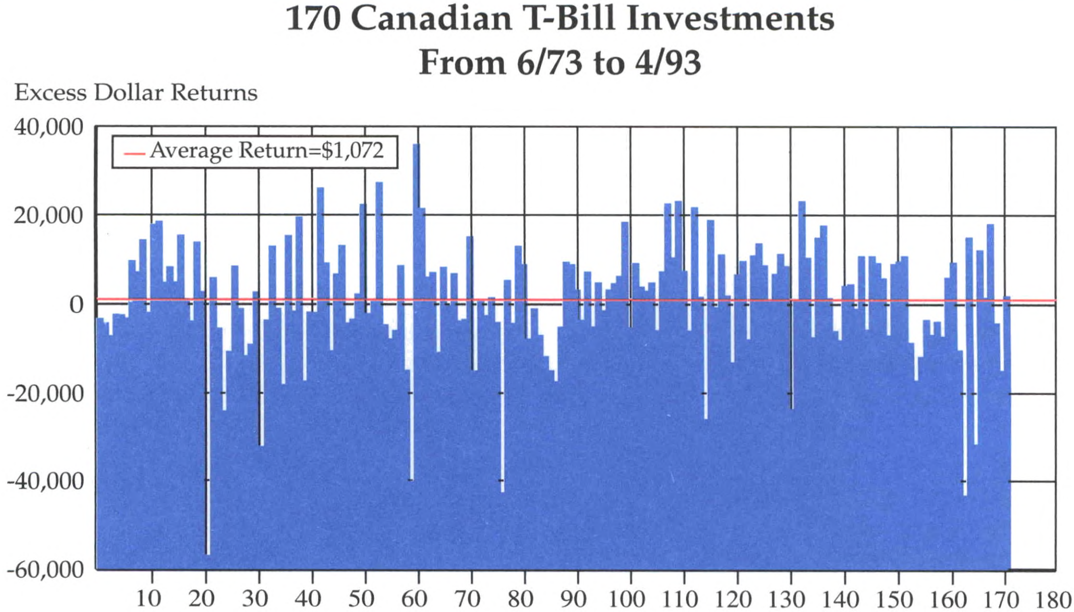
Spot and Forward Exchange Rates. The behavior of the forward exchange rate also challenges foreign exchange market efficiency. Before proceeding with this claim, however, a description of the forward and spot exchange rate markets is in order.[4] Suppose the date is September 1. If the spot exchange rate is 1 Canadian dollar per U.S. dollar, on September 1 the investor could exchange 1 Canadian dollar for 1 U.S. dollar. Similarly, on September 1 an investor can lock into an exchange rate, called the one-month forward exchange rate, for a transaction that will occur one month from that day. For example, an investor might be able to buy the Canadian dollar in the forward market at the forward exchange rate of 2 Canadian dollars per U.S. dollar on September 1. The forward exchange rate is agreed to and known on September 1, but no money changes hands. One month from that day, however, the investor is obligated to trade 1 U.S. dollar for 2 Canadian dollars.

[2]This result has been shown empirically to be generally true for the currencies of the major industrial countries. See for example Froot (1990).

[3]For a summary of the evidence against foreign exchange market efficiency, see Hodrick (1987).

[4]For exposition, the description of the spot and forward exchange rate markets has been simplified. The actual rules governing these markets are slightly more complicated. For a more detailed description, see Grabbe (1991).

FIGURE 1



In an efficient market, risk-neutral investors should set the one-month forward exchange rate equal to what they expect the spot exchange rate to be one month in the future.[5] Otherwise, the market would be allowing exploitable profit opportunities. For example, suppose on September 1 that investors expected the one-month-ahead spot exchange rate to be 1 Canadian dollar per U.S. dollar, but the market set the one-month forward exchange rate to be 2 Canadian dollars per U.S. dollar. Then, the market would be allowing an obvious profit opportunity. On September 1, an investor could enter a one-month forward contract to sell U.S. dollars in exchange for Canadian dollars. One month later, the investor could execute the forward contract by delivering 1 U.S. dollar in exchange for 2 Canadian dollars. Then, if the spot exchange rate on October 1 turned out to be 1 Canadian dollar per U.S. dollar as expected, the investor could sell his 2 Canadian dollars for 2 U.S. dollars in the spot market. The investor would then have made a $1 return, since he turned $1 into $2. However, if the market had set the one-month forward rate to be 1 Canadian dollar per U.S. dollar, and the spot exchange rate on October 1 turned out to be 1 Canadian dollar per U.S. dollar as expected, no return would have been possible.

Since expectations about a specific event usually prove incorrect, we cannot rule out extra returns in any particular month, even if the foreign exchange market is efficient. Suppose that on September 1 the one-month forward exchange rate is set to equal the expected one-month-ahead spot exchange rate of 1 Canadian dollar per U.S. dollar. But on October 1, expectations are proved wrong: the Canadian dollar exchanges for 1.25 U.S. dollars. Had an investor bought the Canadian dollar in the forward exchange market on September 1, he would have made a $0.25 return on October 1. So, in months when the Canadian dollar turned out to be worth more than expected in terms of the U.S. dollar, investors would earn extra returns. In months when the Canadian dollar turned out to be worth less than expected in terms of the U.S. dollar, investors would incur losses. As long as expectations were correct on average, over many months the positive extra returns would cancel out the negative ones and no net extra return would be earned. The market would be efficient even though extra returns appeared randomly in some months.

Biased or Unbiased Predictor? The distinct notions that the market is efficient and that expectations are correct on average can be combined in a single idea: the one-month forward exchange rate should be an unbiased predictor of the one-month-ahead spot exchange rate. In any month, the forward exchange rate in an efficient market will be the same as the market's estimation of the one-month-ahead spot exchange rate. Thus, the forward exchange rate will predict the one-month-ahead spot exchange rate. If expectations are correct on average, the forward rate prediction may not be correct in any particular month but, on average, ought to be correct. In some months, the forward exchange rate will predict a one-month-ahead spot exchange rate that is too high, and in other months, one that is too low. If the predictions are correct on average, the high predictions should cancel out the low predictions, so that the prediction will not be biased either on the high side or on the low side. Consequently, economists claim that the forward exchange rate will be an unbiased predictor of the one-month-ahead spot exchange rate when markets are efficient and expectations are correct on average.

Looking at the data on forward and spot exchange rates, however, casts some doubt on

[5]A risk-neutral investor does not need to be compensated for bearing the risk that the one-month-ahead spot exchange rate may turn out to be different from expectations. Risk will be more fully discussed in the section "Time-Varying Risk Premia."

the joint hypothesis that the market is efficient and expectations are correct on average. It turns out that the forward exchange rate is not an unbiased predictor of the one-month-ahead spot exchange rate, a fact illustrated by looking at the historical relationship between the one-month forward and one-month-ahead spot Canadian dollar-U.S. dollar exchange rates for the period June 1973 to April 1993 (Figure 2). If the forward exchange rate were an unbiased predictor of the one-month-ahead spot exchange rate, the forward exchange rate should fluctuate randomly around the one-month-ahead spot exchange rate. In that way, the forward exchange rate would overpredict the one-month-ahead spot exchange rate as often as it underpredicts it. However, the forward exchange rate for Canadian vs. U.S. dollars does not tend to fluctuate randomly around the one-month-ahead spot exchange rate, but rather tends to stay below the spot rate for extended periods when the spot rate is rising and to stay above the spot rate for extended periods when the spot rate is falling (Figure 2). The one-month forward exchange rate is therefore a biased predictor of the one-month-ahead spot exchange rate.

FIGURE 2

One-Month Forward Rate vs. One-Month-Ahead Spot Rate

(Actual)

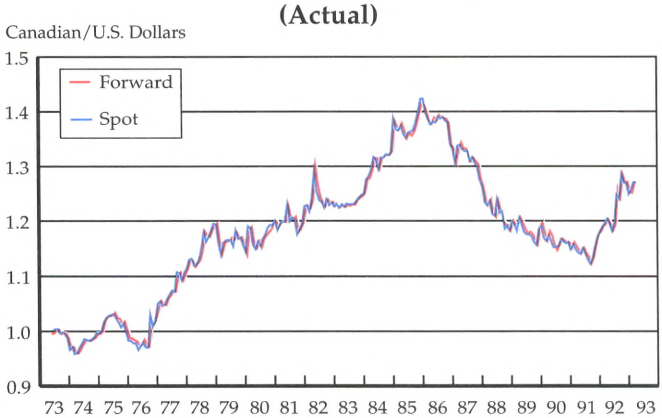
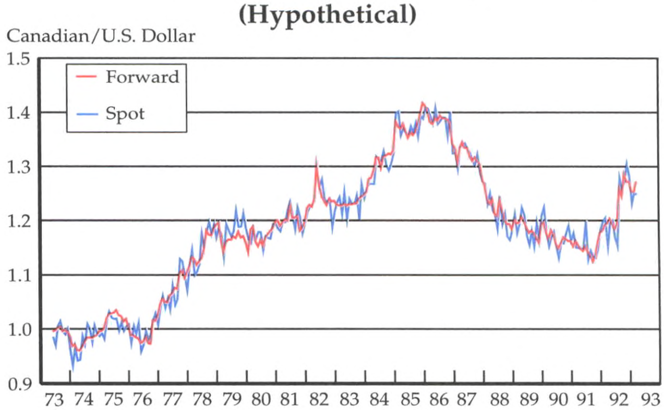


FIGURE 3

One-Month Forward Rate vs. One-Month-Ahead Spot Rate

(Hypothetical)



Contrast this behavior with an unbiased predictor I have constructed (Figure 3).[6] Notice that sometimes the forward rate underpredicts and other times it overpredicts the future spot exchange rate. But it does not systematically over- or underpredict the future exchange rate as a biased predictor would.

That the forward exchange rate is a biased predictor of the one-period-ahead future spot exchange rate suggests that the foreign exchange market may not be efficient and that it may be possible to earn extra returns. Howev-

[6]The artificial one-month-ahead spot exchange rate series was constructed by adding to each actual U.S. dollar-Canadian dollar forward exchange rate in the series the realization of an independent and identically distributed standard normal random variable.

er, economists are not convinced that forward exchange rate bias proves that the foreign exchange market is inefficient. Consequently, they have constructed explanations that allow for the bias of the forward exchange rate while at the same time maintaining market efficiency.

SOME EXPLANATIONS FOR SEEMING MARKET INEFFICIENCY

Statistical Problems. A problem that plagues the study of foreign exchange market efficiency is that the one-month forward exchange rate may be a biased predictor of the one-month-ahead spot exchange rate, even though the market is efficient. This can happen when investors expect an event that has not yet occurred to affect future exchange rates.

A real-world example of this problem concerned the behavior of the Mexican peso in the early 1970s. At that time, the Mexican government fixed the spot dollar-peso exchange rate at a constant value; however, it was widely expected that sometime in the near future the government would devalue the peso—that is, change the rate so that the peso would be worth less in terms of the dollar. Consider the situation before the government changes the fixed exchange rate. When investors form expectations of the one-month-ahead spot exchange rate, they have to take into account the chance that the government might devalue the peso. Thus, investors expect the peso to be worth less in one month than it is today, even though the spot exchange rate is currently fixed. In an efficient market, then, risk-neutral investors will set the peso in the one-month forward market to be worth less in terms of the dollar than it would be at the current fixed rate. Therefore, until the government changes the fixed exchange rate, the one-month forward exchange rate will be a biased predictor of the one-month-ahead spot exchange rate, even if the market is efficient.

It would be a mistake, then, to conclude that because the forward exchange rate is a biased predictor of the one-month-ahead spot exchange rate, the market must be inefficient. This may merely indicate that investors expect an event that has not yet occurred to affect the future exchange rate. International economists—quite naturally—call this kind of statistical problem a peso problem.

Failure of Rational Expectations. The proposition that expectations are correct on average is called rational expectations by economists. The assumption of rational expectations pervades not only international finance but also most branches of economics. Although it has been described implicitly both in the bill-investing example and in our discussion of the forward exchange rate market, it may be useful to explain it in a simple context.

Suppose you play a game in which you flip a fair coin. If heads comes up, you win \$3. If tails comes up, you lose \$1. Clearly, the average value you would win over time is \$1, since half the time you win \$3, and half the time you lose \$1 [(\$3 x 1/2) - (\$1 x 1/2) = \$1]. So, if you expect to win \$1 on average, you have rational expectations, since your estimate of the average value of the winnings is indeed its actual average value. But if you expect to win \$2 on average, you do not have rational expectations.

The assumption of rational expectations seems plausible. However, verification of rational expectations is difficult, since people's expectations are not directly observable. Typically, researchers have attacked this problem indirectly by using surveys of expectations to represent true market expectations. For example, in a 1987 study, Jeffrey Frankel and Ken Froot provided evidence indicating that investors in the foreign exchange market may not have rational expectations. Using survey data of predictions made by private exchange-rate forecasters, they found that forecasters make biased predictions of future exchange rates.[7]

---

[7]This may also indicate a peso problem.

Of course, not all economists accept their results, since they rely on survey data that may not properly reflect true market expectations.

Whether the lack of rational expectations explains the seeming failure of market efficiency is certainly controversial. Many economists find it hard to believe that people do not possess rational expectations, since this implies that investors do not learn from their mistakes but continue to make them systematically. To illustrate the point, let us return to our coin-flipping game. Imagine the following investment. A broker offers investors the chance to play a game in which the investor can win $3 or lose $1. The investor may play the game as often as he likes, but he is not permitted to know that the broker is flipping a coin to decide whether he wins $3 or loses $1. At this point, the investor cannot have rational expectations, since he does not know about the coin. However, if the investor played the game many times and saw the pattern, he would then be able to estimate the average winnings even if he never saw the coin. He may misestimate the average winnings at first, but he would not likely continue to do so.

Many economists believe that investors in the foreign exchange market develop rational expectations in the same way. Investors have a great incentive not to make systematic mistakes in estimating future exchange rates, since failure to do so can lead to large losses. The absence of rational expectations could well explain the seeming failure of foreign exchange market efficiency, but many economists are reluctant to discard the notion of rational expectations given its inherent plausibility.

Time-Varying Risk Premia. A third potential explanation for the seeming failure of efficiency in the foreign exchange market is the possibility of time-varying risk premia. Since the one-month forward exchange rate is a biased predictor of the one-month-ahead spot exchange rate, extra returns seem to be available in the foreign exchange market. But these extra returns may simply be compensation for bearing risk. In the discussion of the bill-investment example and of the forward exchange rate market, we made a crucial assumption: the investor is risk-neutral, which means he does not need to be compensated for taking risk. However, any exchange rate investment in which future exchange rates are uncertain involves exchange rate risk, risk for which a risk-averse investor must be compensated.

Before returning to our example of the Treasury-bill investment, it may be useful to explain the concepts of risk neutrality and risk aversion in the context of our coin-flipping game. Recall that if heads turns up, the investor wins $3. If tails comes up, the investor loses $1. On average, the investor wins $1 playing this game. What is the most an investor would pay each time to play this game? The answer depends on his attitude toward risk. On average the investor stands to win $1 per coin toss, but for any particular toss of the coin, he bears the risk of losing $1. Even if the investor tosses the coin 10 times, he cannot be sure of winning the $1 average return; he may have a run of bad luck. If the investor is risk neutral, he does not need to be compensated for bearing risk. In this case, he would be willing to pay up to $1 to play this game, since that is the average winnings. If he is risk-averse, he must be compensated for bearing risk. Therefore, he would pay, at most, something less than $1 to play this game. The risk premium is the amount that the investor must be compensated for bearing risk.

Suppose the risk-averse investor paid $.75 to play this game. Since the average winnings ($1) in this game exceed the cost to play ($.75), it would seem that a return of $.25 is available. But it would be wrong to conclude that the market for coin-flipping games is inefficient; rather, the $.25 is not a profit opportunity, but compensation for bearing risk.

The situation is much the same in our bill-investment example. A Treasury bill denominated in U.S. dollars represents a claim on the consumption of U.S. goods, since it is ultimately worth a certain amount of U.S. dollars upon maturity. Similarly, a bill denominated in Canadian dollars represents a claim on consumption of Canadian goods. Since Canadian dollars can be converted to U.S. dollars, and vice versa, a U.S. bill is also a claim on Canadian consumption, and a Canadian bill is a claim on U.S. consumption. But the magnitude of these claims is uncertain: when a U.S. bill is redeemed for U.S. dollars, how much U.S. or Canadian goods these U.S. dollars will buy cannot be predicted with certainty. That depends on the exchange rate, which is uncertain, and the prices of U.S. and Canadian goods when the bills are redeemed. Thus, U.S. and Canadian bills are risky assets, even though there's no risk that the governments that issued them will fail to pay investors when the bills mature.

If an investor is risk-averse, he will require a risk premium to compensate him for holding the riskier bill. If Canadian Treasury bills are judged riskier than U.S. bills, Canadian bills must pay a higher return than U.S. bills. Conversely, if U.S. bills are the riskier assets, they must pay a higher return than Canadian bills. Let's go back to our first example: suppose the annual interest rate on the Canadian bill is 1 percentage point higher than that on the U.S. bill. If risk were not compensated, on average the Canadian dollar should turn out to be worth 1 percent less on an annual basis in terms of the U.S. dollar. In that way, the return to holding U.S. bills or Canadian bills is the same. However, suppose we discover that the Canadian dollar turns out on average to be worth 1 percent more in terms of the U.S. dollar on an annual basis. According to the risk premium hypothesis, one interpretation of this situation is that investors are risk-averse. Since the Canadian dollar turns out, on average, to be worth 1 percent more in terms of the U.S. dollar on average, the investor is being paid a 2 percent premium for holding the Canadian bill: he receives a 1 percent capital gain on the Canadian dollar, and he also receives an interest rate 1 percentage point higher than that on the U.S. Treasury bill. Therefore, under the risk premium hypothesis, our interpretation is that Canadian Treasury bills are riskier than U.S. bills, and investors are being paid a 2 percent risk premium to induce them to hold Canadian bills.

The risk premium on Canadian Treasury bills can also be negative. Suppose annual interest rates on U.S. Treasury bills are 1 percentage point higher than those on Canadian bills, and we observe that the U.S. dollar turns out on average to be worth 1 percent more on an annual basis. Then the investor is giving up 2 percent in additional returns to hold the Canadian T-bill: the 1 percent in interest he would have earned on the U.S. T-bill and the 1 percent capital appreciation of the dollar. Under the risk premium hypothesis, the Canadian T-bill is seen as safer than the U.S. T-bill. Because Canadian T-bills are less risky, investors must give up 2 percent in returns in order to hold them. In this case, the risk premium on the Canadian T-bill equals -2 percent.

The risk premium on the Canadian T-bill, then, tends to be positive when Canadian T-bill interest rates exceed those on U.S. T-bills and tends to be negative when interest rates on U.S. T-bills exceed those on Canadian T-bills. Since the interest rate on U.S. T-bills frequently moves above and below the Canadian T-bill rate, the risk premium, if one exists, must frequently vary between positive and negative values. It is in this sense that economists speak of time-varying risk premia.

Risk premium explanations, although plausible, are hard to square with recent economic history. For example, during the mid-1980s, U.S. interest rates were consistently above the interest rates of many foreign countries. Over the same period, the dollar gained in value at a rapid rate against these countries' currencies. Thus, the risk premium in U.S. dollar terms seemed to be large and positive. If the risk premium explanation is true, U.S. dollar assets were seen as much riskier than foreign assets precisely at a time when common opinion held that the dollar was strong because the U.S. was a safe haven for investment. After 1985, despite the fact that U.S. interest rates remained above the interest rates of some foreign countries, the dollar lost value at a rapid rate against these countries' currencies. Thus, the risk premium seemed to become negative, meaning U.S. assets were seen as safer than foreign assets. What produced such a dramatic change? Looking back on the history of the 1980s, it's difficult to point to specific events that may account for these swings in the riskiness of U.S. or foreign bills. (See *Testing for a Time-Varying Risk Premium* for more about risk premia.)

FILTER RULE STUDIES

Another strategy for earning extra returns in the foreign exchange market recommended by some economists is a so-called filter rule. The idea is simple. Whenever a foreign currency is worth a certain percent more (like 1 percent) in U.S. dollar terms than its previous low, invest in foreign assets. Stay invested in those foreign assets until the foreign currency is worth a certain percent less than its previous high, then switch back into dollar assets.[8] In 1983, two economists, Michael Dooley and Jeffrey Shafer, showed that had an investor followed a 1 percent filter rule, he could have earned a fairly consistent return speculating in the major currencies.

Since these returns may have occurred by chance, Richard Sweeney, in 1986, tested whether filter rule profitability can reasonably be attributed to chance. He found that the returns made in filter rule strategies cannot be ascribed to luck and argues that filter rules indeed may provide excess returns, even when transactions costs are accounted for. Sweeney's tests, however, assume that the risk premium is constant. If the risk premium is time-varying, then even if the returns do not occur by chance, they may not be evidence against market efficiency. With time-varying risk premia, the filter rule

[8]The reader who is familiar with the investment literature may recognize this as a form of technical analysis.

Testing for a Time-Varying Risk Premium

Economists have tried to find statistical evidence for the existence of time-varying risk premia, but they have reached no firm conclusions. Typically, they have specified and tested an economic model of investment in the foreign exchange market. A popular economic model, much used by academics and practitioners alike, is the capital asset pricing model (CAPM). The CAPM relates the risk premium to the difference between expected return on the market portfolio of securities and the risk-free rate of interest. Mark (1988) and Hopper (1993) tested the CAPM to see if time-varying risk premia explained the fact that the one-month forward exchange rate is a biased predictor of the one-month-ahead spot exchange rate. Mark (1988) found results favorable to the risk premium hypothesis, but Hopper (1993) found evidence against it.

A related model, popular among economists, is the consumption CAPM. Unlike the standard CAPM, where investors are postulated as making only investment decisions, under a consumption CAPM, investors are also hypothesized as making consumption decisions too. Mark (1985) and Kaminsky and Peruga (1990) tested the consumption CAPM and reported results somewhat favorable to the risk premium hypothesis, but not conclusively so. Hopper (1993) tested a version of the consumption CAPM, but found evidence against the model. In general, most international financial economists believe that more evidence must be accumulated before the risk premium hypothesis can be fully accepted.

could merely be putting the investor into the foreign asset when the risk is high and taking him out when the risk is low. Thus, if there are time-varying risk premia, the returns found in filter rule studies would be compensation for bearing risk and would not be excessive on a risk-adjusted basis.

IMPLICATIONS OF THE ANALYSIS

The reader who has progressed this far might well feel somewhat disappointed with the analysis. The investor or corporate treasurer wants to know whether extra returns can be made if the foreign exchange market is inefficient. Is the foreign exchange market inefficient? On the one hand, excess returns seem to result from following simple rules. On the other hand, the returns may be explained by a peso problem, the failure of rational expectations, or time-varying risk premia. It's possible that these phenomena may be present in the foreign exchange market, but the evidence is far from conclusive. Thus, at present, the best answer economists can give to the question of market inefficiency is—maybe.

How can this help the investor or corporate treasurer? Economists may not be able to show investors how to make money in the foreign exchange market, but they do know enough to help investors avoid losing money. The insights gained by economists who study foreign exchange market efficiency can be used to assess proposed investment strategies. To see how, let us consider a final example.

Suppose once again you are responsible for managing investments for your company. An investment consulting company approaches you with a proposition. The proposition involves not a simple filter rule, but rather a complex rule involving arcane mathematics. The consulting company shows that had their rule been followed during the past 10 years, a 20 percent average annual rate of return would have resulted. Moreover, clients who have actually used the rule over the past year have continued to earn a 20 percent annual rate of return. Should you use the rule? The analysis in this article suggests some questions that can guide your decision.

First, what is the time-varying risk in following the strategy? After all, the complicated mathematics in the rule might merely be instructing the investor to invest in very risky assets at each point in time. If this were the case, it would not be surprising to find a large average return. But the return would be at the expense of assuming higher risk. Thus, the rule is valuable only if it earns returns above what it should earn when risk is accounted for. It's not enough, then, for the consulting company to claim a 20 percent average annual rate of return. The consulting company must show that the returns are above normal on a risk-adjusted basis, that is, the consulting company must have a plausible model of time-varying risk premia.

Second, could a peso problem account for the rule's profitability? As we have seen, returns may appear to be available if investors expect an event that has not yet occurred. If a peso problem exists, an investor following a rule may make money before the event occurs but would lose it after the event occurs. Thus, if the market widely expects an event that has not yet occurred (such as a country withdrawing its currency from the European Exchange Rate Mechanism), a peso problem must be suspected when a technical trading rule seems to offer excess returns.

CONCLUSION

An investor can earn extra returns by always investing in the short-term government securities of the major industrial country that pays the highest interest rate, although the risk to such a strategy will be high; moreover, the forward exchange rate is a biased predictor of the future spot exchange rate. Economists are at present undecided whether these facts should be interpreted to mean that the foreign exchange market is inefficient. A peso problem might explain why the forward exchange rate is a biased predictor of the future spot exchange rate. Alternatively, if investors do not have rational expectations or if there are time-varying risk premia, the foreign exchange market may still be efficient. Some survey evidence supports the claim that investors do not have rational expectations, but such evidence is not convincing to many economists, since surveys may not reflect true expectations of investors in the market. There is also some favorable evidence on the existence of time-varying risk premia, but the models are open to statistical dispute.

Because the foreign exchange market may well be inefficient, extra returns that cannot be explained by the assumption of greater risk might be earned in the foreign exchange markets. By employing filter rule strategies, the investor can earn above normal returns that do not appear to be attributable to chance. Whether these returns are truly evidence of market inefficiency or are merely compensation for bearing risk remains an open question.

Answering this question is important not only for economists but also for practitioners in the foreign exchange markets. Given the inconclusive nature of the literature and the possibility that the foreign exchange market is inefficient, investors cannot be sure whether they are forgoing extra returns in the foreign exchange market. How the literature on foreign exchange market efficiency will progress cannot be predicted, but investors have a practical incentive to follow its development.

REFERENCES

Dooley, Michael P., and Jeffrey Shafer. "Analysis of Short-Run Exchange Rate Behavior: March 1973 to November 1981," in D. Bigman and T. Taya, eds., *Exchange Rate and Trade Instability: Causes, Consequences, and Remedies.* Cambridge, MA: Ballinger, 1983.

Frankel, Jeffrey A., and Kenneth A. Froot. "Using Survey Data to Test Standard Propositions on Exchange Rate Expectations," *American Economic Review* (1987), pp. 133-53.

Froot, Kenneth A. "Short Rates and Expected Asset Returns," NBER Working Paper No. 3247 (January 1990).

Froot, Kenneth A., and Richard H. Thaler. "Anomalies: Foreign Exchange," *Journal of Economic Perspectives* (1990), pp. 179-92.

Grabbe, J. Orlin. *International Financial Markets.* New York, NY: Elsevier Science Publishing Co., 1991.

Hodrick, Robert J. *The Empirical Evidence on the Efficiency of the Forward and Futures Foreign Exchange Markets.* New York, NY: Harwood Academic Publishers, 1987.

Hopper, Gregory P. "Can A Time-Varying Risk Premium Explain the Failure of Uncovered Interest Parity in the Market for Foreign Exchange?" Working Paper 92-25, Federal Reserve Bank of Philadelphia (December 1992).

Hopper, Gregory P. "Time-Varying Consumption Betas and the Market for Foreign Exchange," mimeo, Federal Reserve Bank of Philadelphia (November 1993).

Kaminsky, Graciela, and Rodrigo Peruga. "Can a Time-Varying Risk Premium Explain Excess Returns in the Forward Market for Foreign Exchange?" *Journal of International Economics* (1990), pp. 47-70.

Mark, Nelson C. "On Time-Varying Risk Premia in the Foreign Exchange Market: An Econometric Analysis," *Journal of Monetary Economics* (1985), pp. 3-18.

Mark, Nelson C. "Time-Varying Betas and Risk Premia in the Pricing of Forward Foreign Exchange Contracts," *Journal of Financial Economics* (1988), pp. 335-54.

Sharpe, William F. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance* (September 1964), pp. 425-42.

Sweeney, Richard J. "Beating the Foreign Exchange Market," *Journal of Finance* (1986), pp. 163-82.

FEDERAL
RESERVE BANK OF
PHILADELPHIA

Business Review Ten Independence Mall, Philadelphia, PA 19106-1574