Is the U.S. Economy Really Growing Too Slowly? Maybe We're Measuring Growth Wrong

*Leonard Nakamura**

The central paradox of the American economy today is that we are apparently in an era of extremely rapid technological progress in which economic progress has slowed dramatically — and according to some measures stopped. In an article in the *Wall Street Journal* for June 8, 1995, G. Paschal Zachary quotes Robert M. White, head of the National Academy of Engineering: "The pace and intensity of technological advance are without historical precedent." By contrast, government data say that U.S. aggregate economic growth, after correcting for inflation, has been very slow for the past 20 years, compared with past trends. The apparent consequence is that measured economic rewards have stagnated. In particular, current measurements show real average hourly earnings are lower now than they were 20 years ago. While total real earnings per person (counting all residents of all ages) in the United States have increased, it is only because a larger proportion of us are working and because the quality of the workforce has increased: we are better educated and more experienced at our jobs. What these and other official statistics plainly assert is that changes in the tech-

*Leonard Nakamura is an economic advisor in the Research Department of the Philadelphia Fed.

nology of production no longer lead to improvements in economic well-being.

An alternative view is that the statistics are wrong: the U.S. economy has been experiencing strong growth, but our official measures fail to reflect it. In an article in the November/December 1995 issue of the *Business Review*, I argued this *mismeasurement* view — our price statistics are biased upward, thereby artificially reducing measured U.S. growth.

The policy differences implied by the two alternative economic descriptions are profound. For better or worse, we live in a society in which national economic policy is both formulated and evaluated based on national statistics. Should we stress fairness and support for the unfortunate, or should we stress efficiency and incentives for savings and investment? Should we tax all incomes equally, or should the rich bear a disproportionate burden, even though this reduces their incentives to earn? A prerequisite for answering these kinds of questions is good data. If our economy is stagnating, generosity to those currently in unfortunate circumstances may be misplaced because it may make us less able to be generous in the future. On the other hand, if the economy is growing robustly, we may be able to afford more generous policies today.

In this article, we will explore how economic progress is measured and some policy implications that arise from alternative measures of our rate of growth.

ECONOMIC PROGRESS

What Is It, and How Do We Measure It? Economic progress is best defined as the ability to better meet our needs and desires by increasing the quantity and quality of goods and services at our command. Such progress comes from making better use of existing resources as well as using more resources. Ideally, we would like a measure of economic well-being that takes into account nonmarket activities such as child rearing and home health care. In practice, however, our statistics measure goods and services sold in the marketplace. Thus we measure economic progress by the growth in output of marketed goods and services, not growth in well-being.[1] The advantage in referring only to the marketplace is that we can quantify market activities by the prices paid for marketed goods and services (and as we shall see, even that is no easy task!). Accurate quantitative measurement, in turn, provides common facts about national problems, and such information is important to the success of our policies to overcome them.

For purposes of measurement, there is an important distinction between two kinds of economic progress: progress in our ability to make *existing* goods and services versus our ability to create *new goods and services* that satisfy old needs more efficiently. Economists have long expected existing goods and services to become less important over time because, as wealth rises, we are likely to demand more variety of and higher quality goods: as wealth rises, necessity shrinks in importance and luxury gains.[2]

In all previous eras, necessity was paramount. If we turn back the clock of Western

[1]The market-based measures used in economic analysis, such as gross domestic product (GDP — the broadest measure of domestic output), thus clearly mismeasure growth in well-being. In general, market-based measures are upwardly biased measures of growth in well-being, as nonmarket activities such as child rearing and care for elderly relatives at home become market activities at daycare centers and nursing homes. In earlier times, these types of activities were provided within the household and extended family, not through paid market services, and so were not counted as part of GDP. Thus, even if the amount of these activities hasn't changed, measured GDP has grown, since these activities more frequently take place in the market and thus are counted as part of GDP. This makes the paradox of slow growth in recent years even more marked: GDP shows little growth even though it is intrinsically biased in favor of showing too much growth!

[2]See, for example, H.L. Wold, *Demand Analysis* (John Wiley, 1953).

economic development 400 years, for example, we find that in the Mediterranean world of the Renaissance, food was by far the dominant economic product, representing some four-fifths of all economic output. Of this, fully half was bread grains, primarily wheat.[3] A Spanish or Italian worker might have labored 140 days in a normal year to earn the ton of wheat that meant subsistence for his family; famine was never far away. A drought that doubled the price of grain was ruinous, and such a drought typically recurred three to six times in a worker's average lifetime.

The contrast with the situation facing the American worker is substantial. Today, a ton of wheat costs less than $150 wholesale.[4] Even at the minimum wage of $4.75 an hour this is but a week's work; at the average wage of $12, it is under two days' work. Creating the raw materials necessary for caloric subsistence used to require the preponderance of the working year; now it is a trivial part.

Americans spend more on medicine than on food, beverages, and tobacco (a category that includes restaurants). And the food purchased for home consumption includes an increasing proportion of ready-to-eat or -drink products as the boundary between supermarkets and take-out restaurants disappears. Thus, without question, over these past 400 years, there has been spectacular advance in the standard of living enjoyed by the citizens of economically advanced Western nations like the United States.

A Slower Pace of Measured Growth. In the past 20 years, however, there has been a marked slowdown in measured U.S. economic growth. According to official statistics, real gross domestic product (GDP), a measure of total marketplace economic activity that includes government, business, and consumers, grew at an annual rate of 3.8 percent from 1959 to 1974. But its growth then slowed by one-third, to an annual rate of 2.7 percent, from 1974 to 1994. Population growth slowed too, from 1.3 percent annually to 1.0 percent annually. So on a per-person basis, real GDP slowed from 2.5 percent to 1.7 percent.

This slower pace of growth has been the subject of repeated analysis along the lines that Nobel Laureate Robert Solow advanced: analyzing the sources of growth by measuring the contributions of added capital, added labor, and improved technology. Solow's original work, published in 1957, covered the period 1909 to 1949.[5] During that time, real output in the nonfarm business sector (a convenient grouping that avoids the measurement problems of the agricultural and government sectors) grew at an annual rate of 2.9 percent.[6] Of this rate, 1.1 percentage points were due to an increase in the total number of hours worked (a product, in turn, of more people working a shorter number of hours each year, with the increase in workers outweighing the shortening of the work year). Of the remaining 1.8 percentage points, Solow reckoned that one-fifth (0.4 percent annually) was due to an increase in capital per worker, that is, people having more equipment with which to do their work. The remaining four-fifths (1.4 percent annually) was due to an increase in technological progress, that is,

[3]Fernand Braudel, *The Mediterranean and the Mediterranean World in the Age of Phillip II,* Vol. 1, Sian Reynolds, transl. (Harper and Row, 1972), pp. 418-61.

[4]The cash price of wheat in December 1996 was $4 per bushel, or roughly $135 a ton.

[5]Robert M. Solow, "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics,* 39 (1957), pp. 312-20. The data on labor hours are from Solow's source, later published as John W. Kendrick, *Productivity Trends in the United States* (Princeton University Press, 1961).

[6]In agriculture, the difficulty is counting the hours of farm owners and their families. In the government sector the outputs — compulsory schooling, criminal justice — are hard to count because they are not priced in the marketplace.

having superior procedures and equipment with which to work. Subsequent work by Edward Denison of the Brookings Institution on the period 1929 to 1969 pushed up the annual contribution of technology to 1.7 percent.[7] The clearcut evidence from these and other studies was that for most of the 20th century, most of American economic growth per person was due to improvements in our technology — *how* we worked — rather than to increases in hours worked or amount of capital per worker. Exactly how technological advance of this type occurs and to what extent the improvements in technology reside in organization of the workforce (working smarter) or equipment (smarter tools) have remained unclear. Indeed, technological advance came to be known as the "black box" of economic growth.[8]

When this same approach is applied to the period of the slowdown in growth that began in the 1970s, however, a new, startling conclusion emerges.[9] The statistics show that the growth of labor and capital accounts for *all* of the increased growth beginning in the mid-1970s and that the contribution of technological advance to economic growth has disappeared! Figure 1 shows what is left over from output growth after accounting for increases in capital and labor.[10] From 1929 to 1974, our productivity advanced at an annual rate of 1.7 percent; afterwards, its growth was nearly zero.

The picture of slowdown in U.S. productivity growth in Figure 1 is at odds with the picture of intensive technological advance that appears in business and science publications. Let's take just one example: electronics. Advances in integrated chips have made electronics ubiquitous in the United States. The number of computers in use today rivals the number of cars. And the United States is at the forefront of the design, manufacture, and utilization of integrated circuits.

Similarly, U.S. universities are at the forefront of practically every discipline, from neuroscience to materials science to computer science, from comparative literature to finance to cinema. And this expertise spills over to technology and engineering, as Intel and Microsoft, Merck and Goldman Sachs, Disney and McDonald's continue to dominate world markets.

To try to reduce the dissonance between these two portraits of America, we can look at other aggregate evidence of American well-being. We have already discussed one candidate: the analysis of expenditure on necessities and luxuries.

MORE LUXURIES AND FEWER NECESSITIES: THE CHANGING COMPOSITION OF EXPENDITURES AS INCOMES RISE

A systematic way of testing for the presence of economic growth is to examine the rate at which basic economic necessities, such as food and clothing and household operations, are shrinking as a proportion of total expenditures.[11] The basic empirical principle in this regard is Engel's Law: As real income per person

[7]Edward F. Denison, *Accounting for United States Economic Growth, 1929-1969* (Brookings Institution, 1974). The 1.7 percent figure represents Denison's semiresidual, which includes both pure technological advance and economies of scale — productivity gains due to the increased scale of production. Here, I am lumping the two together. It is now generally recognized that technological advance and economies of scale are, in the long run, inseparable. Output per person grew 2.1 percent during this period.

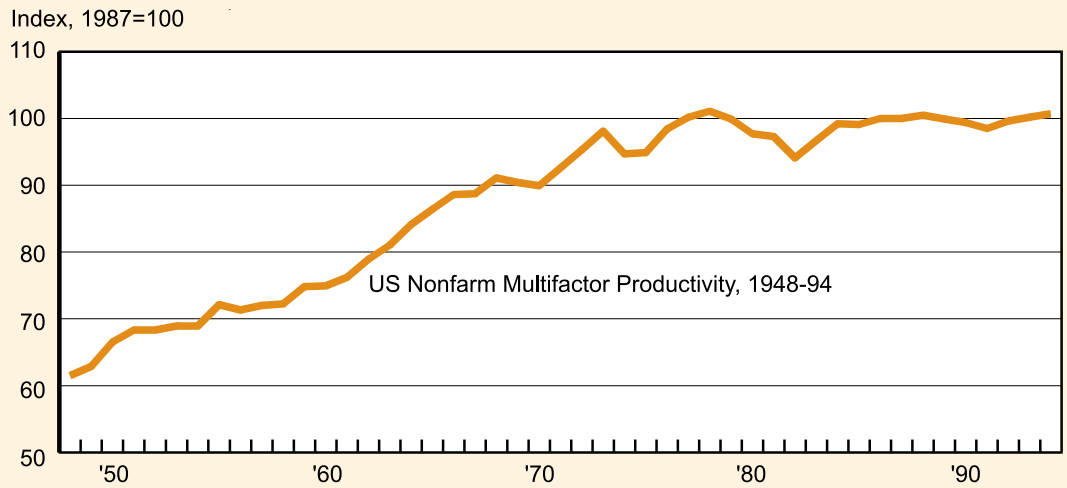
[8]See, for example, the preface to Nathan Rosenberg, *Inside the Black Box: Technology and Economics* (Cambridge University Press, 1982).

[9]Edward F. Denison, *Accounting for Slower Economic Growth: The United States in the 1970s* (Brookings Institution, 1979).

[10]Bureau of Labor Statistics, *Multifactor Productivity Trends, 1994.*

FIGURE 1

U.S. Economic Statistics Appear to Indicate No Technological Progress Since 1977



Source: U.S. Bureau of Labor Statistics, *Multifactor Productivity Trends, 1994.*

rises, the proportion spent on food declines. The eminent Harvard economist Hendrik Houthakker has said, "Of all empirical regularities observed in economic data, Engel's Law is probably the best established; indeed it holds not only in the cross-section data where it was first observed, but has often been confirmed in time-series analysis as well."[12]

Table 1 illustrates the basic idea. Suppose at time 0 real income is 1000, of which 60 percent is spent on food and other necessities, while the other 40 percent is spent on luxuries. Now suppose that real income grew 20 percent, to 1200. Demand for food doesn't increase as much as demand for luxuries, so although food purchases increase, they shrink as a percent of expenditures. Suppose that real income grows another 20 percent. Food purchases continue to rise, but less rapidly than total income and spending. The share spent on food declines over time. Moreover, equal percent increases in real income lead to equal changes in shares of nominal expenditure: in both periods, each share changes 5 percentage points.

This formulation of Engel's Law is based on

[11]This section is based on my paper, "Is U.S. Economic Performance Really That Bad?" Federal Reserve Bank of Philadelphia Working Paper 95-21/R, April 1996.

[12]Hendrik S. Houthakker, "Engel's Law," in John Eatwell, Murray Milgate, and Peter Newman, eds., *The New Palgrave: A Dictionary of Economics.* Volume 2 (Macmillan, 1987), pp.143-44.

TABLE 1

Example of Engel's Law: As Real Incomes Rise, the Share Spent on Necessities Falls

Year	Zero	Ten	Twenty	Change 0 to 10	Change 10 to 20
Spending on food and other necessities	600	825	1080		
Spending on luxuries	400	675	1080		
Nominal expenditures	1000	1500	2160		
Spending on food and other necessities	60%	55%	50%	5%	5%
Spending on luxuries	40%	45%	50%	5%	5%
Nominal income	1000	1500	2160		
Prices	100	125	150		
Real income (dollars)	1000	1200	1440	20%	20%

Note: Deaton and Muellbauer's formulation: Equal percent changes in real incomes in two periods lead to equal percentage point changes in shares of expenditures of necessities and luxuries. All data per person.

work by Angus Deaton and John Muellbauer.[13] It implies that equal percent increases in real income per person should lead to equal percentage point changes in shares of expenditure.

How do we apply their formulation to U.S. data? From 1959 to 1974, according to the official statistics, real income per person grew 45 percent. In the longer period from 1974 to 1994, real income per person grew 39 percent. If these numbers are accurate, one would expect that the share of necessities in total expenditures should have shrunk by about the same amount in the two periods (or perhaps a bit less in the second period). In fact, the proportion of the average budget spent on food fell from 27.3 percent in 1959 to 23.1 percent in 1974, or 4.2 percentage points, but fell substantially more — 7.1 percentage points — from 1974 to 1994.

The proportion of household budgets spent on other necessities, such as clothing and home heating, also almost uniformly contracted by more in the period 1974 to 1994 than in the earlier period 1959 to 1974 (Table 2). In contrast, the share spent on luxuries, such as medical care, personal business services, recreation, education, and foreign travel, generally rose more in the later period than in the earlier one.[14] This

[13]"An Almost-Ideal Demand System," *American Economic Review* 70 (June 1980), pp. 312-16. Their system formally says that holding relative prices constant, equal changes in the logarithm of real income lead to equal changes in shares in nominal expenditures. Here we discuss the system in terms of percent changes as we assume most readers are more familiar with that terminology.

[14]What is a necessity and what is a luxury is not always easy to determine. Food is the clearest example of a necessity. Goods and services whose consumption declines over long periods of time when incomes are rising are defined as necessities here; the consumption of luxuries rises over the same time periods.

TABLE 2

Nominal Spending on Each Category

	Share of Total Spending (in percent)			Change in Spending Shares	
	1959	**1974**	**1994**	**1959-74**	**1974-94**
Total	100	100	100		
Food, Beverages and Tobacco	27.3	23.1	16	-4.2	-7.1
Clothing, Upkeep, and Personal Care	11.7	10.1	7.9	-1.6	-2.2
Housing	14.1	14.3	14.8	0.2	0.5
Household Furnishings, Fuel and Operation	14.1	13.2	11.1	-0.9	-2.1
Medical Care	6.4	10.0	17.5	3.6	7.5
Personal Business Services	4.3	4.9	7.6	0.6	2.7
Transportation	12.7	12.7	11.2	0	-1.5
Recreation	5.5	6.8	7.9	1.3	1.1
Education, Welfare, and Americans' Foreign Travel	3.8	4.9	6.1	1.1	1.2
Average Absolute Change in Shares of Consumption				1.5	2.88
Real GDP per Person, Official Measures, chained 1992 $	$12,494	$18,178	$25,352		

Source: Bureau of Economic Analysis, *Survey of Current Business,* January/February and August, 1996. Population data from *Economic Report of the President*, 1996, U.S. Government Printing Office.

faster shift away from necessities as a proportion of budgets in the second period suggests that real income per person grew more in the second period than in the first, not less as the official statistics say.

How much more? To answer this question, calculate the average absolute change in shares for all consumption categories over each period, that is, take the average without considering whether each change is up or down. In this way, a decline of 2 percent for a necessity like food and a rise of 2 percent for a luxury like travel both correspond to rising real income. The nine consumption categories in Table 2 changed absolutely by 1.50 percentage points, on average, in the period 1959 to 1974, while they changed 2.88 percentage points, on average, from 1974 to 1994. If we use Deaton and Muellbauer's application of Engel's Law, the fact that the average shift in spending shares (away from necessities and toward luxuries) was almost twice as big in the second period as in the first — 2.88 to 1.50 percentage points — implies that the true rise in real income in the second period was about twice as large as that in the first (so long as prices of luxuries did not

rise at a substantially different rate than prices of necessities). If real income rose 45 percent from 1959 to 1974 as the official data show, the change in spending shares from 1974 to 1994 suggests that real income rose just over 100 percent during those 20 years, not 39 percent as reported in the official statistics.[15] Over 1974 to 1994, this represents a per-person annual growth rate of 3.7 percent, not 1.7 percent — a difference of 2.0 percentage points per year.

Now let's reexamine the productivity slowdown that began around the mid-1970s. That slowdown is reflected in the official data in Table 2 in that more real growth per person took place from 1959 to 1974 than in the longer period from 1974 to 1994. But the slowdown is *not* consistent with the changes in the consumption expenditure shares. The implication of the calculations reported above is that growth in real income per person was mismeasured by 2.0 percentage points annually from 1974 to 1994 — slightly more than the measured slowdown in productivity growth in the official statistics of 1.7 percentage points annually. That is, it is possible that the entire productivity slowdown of the past two decades revealed by the official statistics is the result of mismeasurement! Put another way, the shifts in composition of expenditures from 1959 to 1974 and from 1974 to 1994 are consistent with the view that productivity growth was the same in both periods. Households are spending in a pattern that is inconsistent with the official statistics on real output and price; that is, the average household has expanded the proportion of luxuries it buys as if its real income had doubled over the last 20 years, while the official data report that its real income rose by less than half.

INCREASING UNCERTAINTY IN OUTPUT MEASUREMENT IN THE AMERICAN ECONOMY

Is it really possible that growth could be mismeasured on this scale? Quantifying economic progress was easier in earlier periods in the industrial revolution. Mass production standardized many goods and thus made their output easier to measure. The more uniform quality of apples and wheat and cars and shoes made for standardized pricing and publication of wholesale and retail prices of these commodities. The most rapid progress took place in the production of goods whose increased quantities we were best able to measure.

Now, an increasing proportion of the economy is devoted to products whose real output we do not attempt to measure. As Zvi Griliches pointed out in his 1993 presidential address to the American Economic Association, the industrial composition of the economy has shifted to service activities that we are not well prepared to measure.[16] And an increasing proportion of goods we do measure is changing more rapidly than in the past, adding to the measurement difficulties.

The clearest example lies in two major components of current consumption expenditures, medical care and personal business services, which are predominately measured by inputs rather than by outputs. Our official data estimate the output of doctors or insurance agents by the number of hours doctors and insurance agents work, rather than the success rate of treatment or number of insurance policies written.[17,18] That is, our statistics assume produc-

[15]To see this, remember that the underlying arithmetic is being done in logs. The change in the log of real income from 1959 to 1974 is 0.375. We multiply this by the ratio between the percent changes, 0.375 x (2.88/1.50) = 0.720. The antilog of .720 is 2.05, suggesting that real per capita income in 1994 was 2.05 times real per capita income in 1974.

[16]Zvi Griliches, "Productivity, R&D, and the Data Constraint," *American Economic Review* 84 (March, 1994), pp. 1-23.

tivity growth in these areas is nonexistent. And these two categories of consumption alone have grown from about 11 percent of consumption expenses in 1959 to over 25 percent of consumption expenses in 1994, so the errors in measurement loom far larger.

IMPROVING U.S. ECONOMIC PERFORMANCE: A PROBLEM IN MEASUREMENT, DIAGNOSIS, AND PRESCRIPTION

We thus are confronted by two possibilities: One, our true economic performance has been quite good, but our measurement of that performance has been faulty. Two, our measures are right, and scientific and technological progress is not being translated into increased economic output. Is this difference important?

The data matter because political disagreements about what our problems are and how to fix them rest on statistics, as does our ability to evaluate the success or failure of our efforts to solve them. Programs that leave in their wake high inflation and low growth are clearly failures, while those that result in low inflation and high growth are successes. If the rate of increase in prices is overestimated, so that growth in output is understated, an economic policy that, in fact, has successfully generated high growth and low inflation will appear to be a failure that has generated low growth and high inflation.

Here are some examples that show why knowing the correct measures of growth is critical to our understanding of the economy. Two other examples are discussed in *Output Mismeasurement in Health and Educational Services.*

Fostering Growth. If we accept the official data at face value and productivity is stagnating, future generations may be worse off. To be generous to those future generations, we may need to decrease constraints on economic growth and increase incentives to economic efficiency, even at the expense of equity in the present. Milton Friedman has argued that government regulations increased dramatically in the 1970s under Presidents Nixon and Carter, declined under President Reagan, and then rebounded to a new high under Presidents Bush and Clinton (Figure 2). If these regulations are associated with programs that have benefited the aged (for example, Medicare and Social Security) and the unfortunate (for example, the Americans with Disabilities Act and Medicaid) at the expense of economic efficiency, perhaps this generosity was misplaced, for, according to the official statistics, our ability to produce with fixed resources is on the verge of deterioration.

Another way to foster growth is to reduce government claims on resources, permitting greater private incentives. A sharp decline in government purchases and cutbacks in the federal safety net are already in progress.

On the other hand, if real output and inflation have been mismeasured, the apparent failure of current policies may be an artifact of bad statistics. Has free trade in the United States been costly? The apparent slow growth of the U.S. economy in the 1980s and 1990s has led some critics to argue that policies too generous to our trading partners have put us on a permanently slower growth path. But are we really on a slower growth path?

[17]Technically, the Bureau of Economic Analysis deflates the nominal revenues of these service providers by a weighted average of input prices. To the extent that any increase in productivity results in higher wages, it will not be measured as increased real output.

[18]The Bureau of Economic Analysis has recently changed the method it uses for deflating medical services. It now uses the Producer Price Index (PPI) for medical services, a series first collected in late 1992, to deflate that segment of personal consumption expenditures. This series measures the costs of treating a disease, a procedure that should be substantially closer to the right measure. However, this measure does not take into consideration improvements in quality of outcomes.

Output Mismeasurement in Health and Educational Services

Measurement problems also apply to policy issues within specific industries, such as health and education. Our efforts in these areas may suffer because we confuse inflation and technical progress.

Health Care in Crisis. We are experiencing a national crisis over health care expenses. This is no wonder: medical care now counts for one-sixth of our consumption. But because we do not measure the output of doctors, nurses, or pharmacists, but rather inputs, some people have the impression that medical providers are absorbing an increasing part of American income out of pure avarice. But what economic evidence there is suggests instead that medical costs are rising entirely because of technological advance — improvements in medical practices cause us to want to buy a lot more medical care.*

As surgery has become less invasive and recoveries more rapid and as the probability of success has risen, patients and doctors have chosen surgical intervention more often. Far more heart attack victims now choose to undergo bypasses and angioplasties because the survival rates are higher and recoveries quicker. Knee surgeries have proliferated as arthroscopic procedures have shortened recovery times to days rather than months. New, expensive drug combinations now offer hope to AIDS victims, when before their cases had been considered hopeless. Thus, the improvement in medical practices has widened demand and, in Cutler's analysis, is the main force that has driven the expansion of medical care.

The Rising Cost of College. A similar issue has been raised about education. Tuition costs at private colleges and universities, for example, have risen more rapidly than the price of medical care. Again, the increased value of the education is not being measured. Unquestionably, the totality of academic information has expanded substantially, as fundamental advances have been made in every physical and social science. As colleges and universities have more to teach, the value of education has risen. This value may be hard to measure precisely, but all available evidence suggests that the rate of return to a college education is increasing rather than falling.

Both of these cases are American success stories in which the producers are at the vanguard of worldwide scientific and technological achievement. But because our statistics treat these achievements as raising prices rather than increasing output, we risk mistaking our achievements and instead seek "reform" of health care and education.

*David Cutler, "Technology, Health Costs, and the NIH," Harvard, mimeo (September 1995).

To take another example, part of the sense of crisis about the U.S. economy is a fear that the Social Security system will be unable to support a rising burden of retirees. Part of this fear may be due to the fact that once a retiree enters the Social Security pool, payments rise with the Consumer Price Index. In theory, this permits retirees to keep up with inflation. But if the Consumer Price Index is overstated, as the Boskin Commission has argued, retirees are enjoying rising real incomes. Indeed, average Social Security benefits have been rising faster than wages — virtually guaranteeing long-run instability.

Are We Saving and Investing Too Little? U.S. national income accounts are set up to measure investment in goods — plant and equipment — but not in information: research, edu-

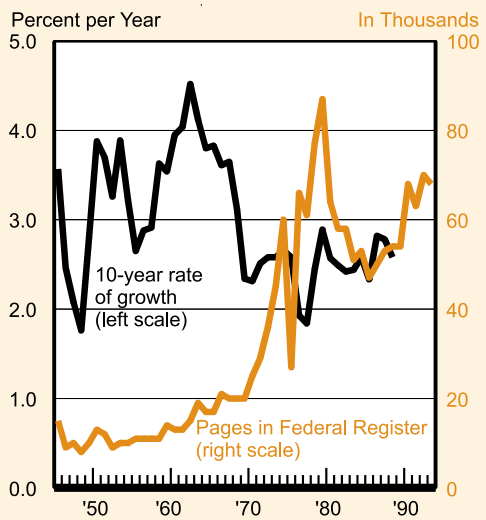
cation, computer software, data, and on-the-job experience. American investment in formal research and development efforts amounted to $169 billion in 1994, and public and private expenditures on formal education (not counting the value of the time invested by students in acquiring education) were $508 billion in 1994. These expenditures are roughly the same size as our investments in business structures ($180 billion in 1994) and equipment ($487 billion in 1994). Our gross investment rate is far larger than official data show if we consider these informational expenditures to be investments rather than costs of doing business (as our statistics currently treat research and development) or consumption expenditures (as they treat education).[19] Moreover, at an individual level, categorizing our expenditures on education as investments rather than consumption would, by lowering measured consumption, boost savings rates as well, since savings is the difference between consumption and income. Measured personal savings in 1994 was less than $200 billion, so that counting these additional investments in human capital as saving could make a substantial difference. Perhaps we are already saving and investing at unusually high rates! And fostering further saving and investment may not be so crucial after all.

This raises another paradox, however, because if we raise our estimates of capital stock, our estimates of total factor productivity would worsen. With more measured capital, the diagram in Figure 1 would show a clear long-term decline in our technological level. Only if real output growth is shown to be understated can we sensibly argue that investment is understated.

A more difficult, although related, problem is the spreading inequality of incomes. In particular, college-educated workers are increasingly better paid than the less well educated.[20] This disparity is likely related to the rate of technological advance: those who have college educations are better equipped to learn the addi-

FIGURE 2

More Rules, Less Growth?



Note: Decadal rate of real growth of U.S. national income, and annual number of pages in Federal Register, 1946-93

Source: Data courtesy of Milton Friedman. These data appeared in an article written by Dr. Friedman for the *Wall Street Journal*, August 1, 1995.

[19]In 1994, these information investments were about equal to investment in plant and equipment; in 1959, they were 78 percent as large as investment in plant and equipment and in 1974, 86 percent.

[20]Lawrence F. Katz and Kevin M. Murphy investigate the changing supply and demand of college-educated workers in "Changes in Relative Wages, 1963 -1987: Supply and Demand Factors," *Quarterly Journal of Economics* CVII February 1992, pp. 35-78. They find that demand for college-educated workers has been steadily rising, but supply has fluctuated, exceeding demand in the 1970s but falling behind in the 1980s.

tional skills necessary to keep up with a rapidly changing workplace. College graduates, for example, are far more likely than high school graduates to participate in continuing education courses in a given year (52 percent compared with 22 percent).[21] Rapid technological advance and increased economic efficiency may directly exacerbate this inequality, for example, by increasing the return to education. Thus, reducing subsidies to students or spending on education may exacerbate future problems of inequality and deprive future generations rather than help them. It may also be inefficient to shift resources away from high-return investments in education to low-return investments in physical capital.

If we believe the economy is performing poorly, we may try to reduce educational subsidies because we believe that we cannot afford them. Doing so may reduce efficiency and equity if our belief is incorrect.

THE IMPORTANCE OF STATISTICAL INFORMATION

While the fact of statistical mismeasurement may be clear, estimates of the size of statistical mismeasurement differ widely. Estimating these statistics more precisely is crucial to our nation's ability to make effective policy. They provide the yardsticks by which we measure our treatment of the poor, the rich, the elderly, the infirm, and students.

It is widely recognized that the official data underestimate growth and productivity.[22] But the size of the error is unknown. We need further studies, detailed experimental series, and new means of collecting data to obtain better estimates of economic progress. The results of these studies would progressively be incorporated in our data series as the economy evolves.

To improve statistical measurements in the short run and possibly even in the long run, our statistical agencies would need more resources; the experimental collection of information aimed at improving our official statistics is expensive both in terms of economic expertise and electronic hardware and software. While it is true that electronics can improve the efficiency of data collection, the rapid rate of change in the economy means that properly collecting the data will become more and more difficult and will require increasing the amount of intellectual analysis that goes into data collection. But if our statistics are unable to keep up with new economic realities, it will be extremely difficult for government policies to be farsighted.

[21]Data are for 1990-91 and are taken from U.S. Bureau of the Census, *Statistical Abstract of the United States, 1995* (Bernan, 1995), p. 194, which cites the U.S. National Center for Education Statistics, *Adult Education Profile for 1990-91.*

[22]For example, W. Erwin Diewert, "Comment on CPI Biases," *Business Economics* (April 1996), and Matthew D. Shapiro and David W. Wilcox, "Causes and Consequences of Imperfections in the Consumer Price Index," NBER Working Paper No. 5590, May 1996.

The Livingston Survey: Still Useful After All These Years

*Dean Croushore**

The decisions of households, firms, and government agencies depend on forecasts of the overall economy. Households face decisions such as whether to refinance a mortgage, whether to buy or lease a car, how to invest their savings, how to plan for retirement, and how to save for their children's college education, all of which depend on the future movements of macroeconomic variables like interest rates, output, the stock market, and inflation. Those same macroeconomic variables influence the decisions of business firms about what equipment to buy, how to market a product, how to invest excess funds, and how to borrow to meet liquidity needs. Government policymakers need good forecasts to calculate the costs and benefits of spending programs and to estimate tax revenues.

Households, firms, and governments can get macroeconomic forecasts in many different ways. Large firms and the federal government may have the resources to hire their own economists to provide forecasts. Some organizations may hire a consulting firm to forecast for them. But for smaller firms, households, and local governments, for which such an expense

*Dean Croushore is an assistant vice president and economist and head of the Macroeconomics section of the Philadelphia Fed's Research Department.

wouldn't be worthwhile, are there any alternatives?

One alternative is to subscribe to a survey of forecasters, such as the Livingston Survey, which, after 50 years, still provides useful forecasts of the economy.

THE SURVEY'S HISTORY

In 1946, Joseph A. Livingston, then a columnist for the *Philadelphia Record,* began asking business economists he knew to provide him with their forecasts for important economic variables. He followed through with the survey every six months, in June and December, even when he moved from the *Record* to the *Bulletin* in 1948 and to the *Philadelphia Inquirer* in 1972. Livingston wrote a lively (at least by economists' standards!) column about the results, and his survey gained a national following.

The survey had been around for a long time before economists discovered its value in the 1970s. The rise of inflation in that decade led to the formulation of a new economic theory known as rational expectations. The theory suggests that people who forecast will use all available, relevant information. Doing so implies that their forecasts won't be systematically biased (that is, forecast errors will be random and won't have any obvious pattern to them). As economists began seeking ways to test the rational expectations theory, they turned to the Livingston Survey, which was the only good collection of forecasts of macroeconomic variables.

Livingston received many requests from economists for his survey data — so many that, in 1978, he turned to the Federal Reserve Bank of Philadelphia for help. The Philadelphia Fed entered the survey data into a computer database and made the data available to researchers in an organized fashion.

In 1985, the Philadelphia Fed and Livingston, acknowledging the value of continuing the survey perpetually, agreed that if Livingston no longer wanted to continue the survey, or was unable to do so, the Fed would take over. So when Livingston died in 1989, the Philadelphia Fed took over the survey.

In a 1992 issue of this *Business Review,* Herb Taylor provided a more complete history of the survey and how Livingston wrote about it. This article will focus more on the details of the survey and its value in research.

HOW THE SURVEY WORKS

Every June and December, the Livingston Survey asks participants to forecast a set of key macroeconomic variables, including real and nominal GDP (gross domestic product, which is the best overall measure of our economy's total production), inflation (both the producer price index and the consumer price index), the unemployment rate, the three-month Treasury-bill interest rate, the interest rate on 30-year Treasury bonds, and the S&P 500 stock index. In addition, the survey covers a variety of other variables: real nonresidential fixed investment, corporate profits after taxes, industrial production, total private housing starts, average weekly earnings in manufacturing, retail trade, auto sales, and the prime interest rate.

Survey participants are asked to provide forecasts for these variables for the end of the current month, six months ahead, and 12 months ahead. They also provide forecasts for the current calendar year, the following calendar year, and (in the December survey only) the calendar year after next. In 1991, the survey added a new question, which asked for 10-year forecasts for real GDP and consumer price inflation.

The timing of the survey is crucial, since the information available to forecasters affects what their forecasts will be. Because of the importance of forecasts of consumer price inflation, the survey is mailed to participants in May and November, immediately following the government's release of the consumer price index (CPI) for the preceding month. The Fed re-

quests that the survey be returned before the next release of the CPI in June and December.

Who are the forecasters? Because of his work as a journalist, Livingston was well known in the business economics community. He polled economists with whom he had discussed issues for stories he was writing, including many economists on Wall Street, in corporations, and at forecasting firms, as well as private consultants. He also developed contacts with economists in academia who were interested in macroeconomic forecasting. The Philadelphia Fed continues the tradition of seeking a wide variety of participants. Currently, about 30 percent of the forecasters come from nonfinancial businesses (for example, a number are chief economists of corporations), 29 percent are from investment banking firms (many on Wall Street), 20 percent represent commercial banks around the country, 13 percent work in academic institutions, and the remaining 8 percent come from labor organizations, government, and insurance companies.

The Livingston Survey would not have survived the past 50 years had it not been of value to someone.

The number of participants in the survey has been fairly steady over time, averaging about 50 forecasters in each survey. Currently, about 90 participants are on the mailing list, and 55 to 65 return the survey each time.

An important issue in surveys like the Livingston is how to get people to provide their true forecasts. After all, the participants aren't paid; their only reward is that they receive the survey on a regular basis and see their name on the list of participants. However, the survey does not reveal which individual made which forecasts. If it did, some participants might shade their forecasts more toward the consensus (to avoid unfavorable publicity when wrong), while others might make unusually bold forecasts, hoping to stand out from the crowd. To prevent these publicity effects, the survey provides a list of the participants but doesn't tell who provided which forecast.

The results of the Livingston Survey are released to the press near the end of June and December each year. Staff members in the Research and Statistics Department at the Philadelphia Fed prepare a press release and tables of the survey results, which are released on a public-relations news service. Stories about the survey appear in many major newspapers and on Wall Street newswires. Staff members also report the results in on-camera or telephone interviews. The results are also made available via the Internet on the Philadelphia Fed's Web page (http://www.phil.frb.org).

IMPORTANCE OF THE LIVINGSTON SURVEY IN ECONOMIC RESEARCH

The Livingston Survey would not have survived the past 50 years had it not been of value to someone. Joseph Livingston used it as the basis for two newspaper columns each year. But its more enduring legacy is the research done by academic economists, who have used and tested the survey forecasts in many different ways to examine hypotheses about expectations. (For a synopsis of current research on expectations in economics presented at a recent conference, see *Conference on Expectations in Economics: In Honor of the 50th Anniversary of the Livingston Survey.*)

The first question one might ask is: how good

CONFERENCE ON EXPECTATIONS IN ECONOMICS: In Honor of the 50th Anniversary of the Livingston Survey

The Federal Reserve Bank of Philadelphia and the University of Pennsylvania jointly sponsored a conference in October 1996 on "Expectations in Economics: In Honor of the 50th Anniversary of the Livingston Survey." Seven new papers showed that research about expectations is still active and making new discoveries.

Robert Shiller of Yale University presented "Why Do People Dislike Inflation?" In the paper, Shiller reports the results of a survey that asks people what it is about inflation that causes problems for them, and how they rank inflation's importance compared with other problems they face. Shiller compares the results of surveys taken in the United States to those taken in Germany and Brazil.

Charles Manski of the University of Wisconsin presented his joint work with Jeff Dominitz, "Perceptions of Economic Insecurity: Evidence from the Survey of Economic Expectations." The authors report the results of a new survey that asks people about their susceptibility to the loss of health insurance, burglary, and job loss. They compare people's perceptions of the likelihood of these events to data on how often the events actually occur and find that people perceive their risk of loss due to crime to be much higher than it really is.

Garey Ramey of the University of California at San Diego discussed his joint work with George Evans, "Calculation, Adaptation, and Rational Expectations." The paper shows how intensely people calculate things based on their estimates of benefits compared with costs. It may not be worthwhile for people to spend time and effort gathering all the information possible before reaching a decision. The authors use numerical illustrations to demonstrate the theoretical work.

Stephen McNees, an economic consultant formerly at the Federal Reserve Bank of Boston, tests probability forecasts from the Survey of Professional Forecasters in his paper "Forecast Uncertainty: Can It Be Measured?" McNees examines whether the probabilities reported by the survey participants are accurate. If the participants say there's an 80 percent chance that real GDP will grow between 1 and 3 percent over the next year, does real GDP actually fall into that range 80 percent of the time? McNees finds that, on average, the probabilities are accurate, though not for all the participants in the survey.

In their paper "Measuring Predictability: Theory and Macroeconomic Applications," Lutz Kilian of the University of Michigan and Frank Diebold of the University of Pennsylvania reported on the predictability of different macroeconomic variables. The paper begins by creating a general measure of a variable's predictability. The authors show how to estimate predictability from sample time-series data and apply it to U.S. data.

Carl Bonham of the University of Hawaii presented his work (with Richard Cohen) "Heterogeneous Expectations: Aggregation Bias and the Poolability of Survey Forecasts in Tests of the Rational Expectations Hypothesis." The paper makes clear the relationship between forecasts made by individuals and the average across those forecasts in tests of rational expectations. They examine forecasts from the Survey of Professional Forecasters.

Owen Lamont from the University of Chicago asserts that older economists make bolder forecasts. In his paper "Macroeconomic Forecasts and Microeconomic Forecasters," he studies forecasts from *Business Week* and shows that economists attempt to position their forecasts to manipulate beliefs about the economists' ability. As economists become older and more established, they produce more radical forecasts, which are less accurate.

These papers show that economists are actively engaged in research on expectations. Surveys like the Livingston Survey continue to provide useful data for testing economic theories.

Copies can be obtained by writing to the authors of the papers.

are the Livingston forecasts? If a household, firm, or government based its planning on the survey, how would it make out? As we'll see, these questions have been answered in different ways as the survey has been tested in research. But it's best to begin by looking at what economists were thinking when they first began to investigate the survey.

In the 1960s and 1970s, macroeconomics was dominated by the theories of John Maynard Keynes. Keynesians said that rigid prices and wages were important aspects of the macroeconomy. Adverse shocks to the economy could lead to recessions because prices and wages would not adjust immediately. The theory implied that the government could eliminate business cycles by using appropriate fiscal and monetary policies. The Keynesian models of the era were based on adaptive expectations — they assumed that people's expectations of future inflation were just projections of past inflation.

In contrast, rational expectations theory suggested that people would use all the information available to them in forming forecasts. The idea of rational expectations originated with John F. Muth in 1961, but the theory took off in the second half of the 1970s with the publication of key papers by Robert Lucas, by Lucas and Thomas Sargent, and by Sargent and Neil Wallace.

Rational expectations theory showed how the Keynesian models were flawed by their failure to deal adequately with expectations. According to this theory, people can make better forecasts by using more information. For example, if people observe the money supply increasing more rapidly than before, they'll increase their forecasts of inflation because they know that faster money growth leads to higher inflation.

The theory of rational expectations was important because it provided a reasonable explanation of the failure of the Phillips curve in the 1970s. The Phillips curve showed a tradeoff between inflation and unemployment that Keynesian economists believed to be stable. If inflation rose, unemployment would fall, and vice versa. In the mid-1970s, a deep recession was accompanied by rising inflation — an impossibility according to the Phillips curve model. Rational expectations theory showed how inflation and unemployment could rise together when people began to expect higher rates of inflation.

The theory of rational expectations doesn't say that everybody should be a great forecaster — only people whose forecasts matter to their livelihoods. An engineer doesn't spend much time forecasting inflation because her job doesn't depend on it. But a Wall Street money manager may spend a lot of time and energy forecasting inflation because the returns on her portfolio depend a lot on what happens to inflation. So the money manager will spend a lot more time trying to understand the economy and reading through the details of economic reports than will the engineer.

How Can the Theory of Rational Expectations Be Tested? How do we know if people are forming their expectations rationally? One way is to look at people's forecasts and see if they're any good. But what constitutes a good forecast? That question has been answered in a number of different ways using the Livingston Survey.

If people have rational expectations, they shouldn't make systematic errors in producing their forecasts. If their errors had a recognizable pattern, the forecasters would modify their methods.

There are two obvious ways in which people could make systematic errors. One is that, over long periods of time, their forecast errors might not average out to zero. But if so, the forecast errors could be reduced by subtracting out the average error from the forecasts. This idea suggests some straightforward tests for the quality of forecasts. A simple test is to calculate the average forecast error and see if it's close to zero.

A more sophisticated test is to plot the actual value of the variable against the forecast. If the forecast is unbiased, the plot should be centered on a 45-degree line through the origin. That would mean that, on average, the forecasts were randomly distributed around the actual values. One can also use the statistical tool of econometrics to investigate how close those points are to the 45-degree line. Using linear regression analysis, an analyst can see if the best-fitting line through the plot of the actual values against the forecasts is significantly different from the 45-degree line. If it is, the forecasts are said to be biased; if not, the forecasts are unbiased.

Forecasters make a second type of systematic error: they fail to use all available information when forming their forecasts. If forecasters ignore some relevant piece of data, their forecast errors will be systematically related to that data. An econometric test examines whether the forecast errors are related to data that forecasters had available. If there's no such relationship, the forecasts are said to be efficient. If there is a statistically significant relationship, the forecasts are inefficient. For more technical details on these tests, see the Appendix.

Early Tests of Rational Expectations of Inflation. Inflation is the most important macroeconomic variable for testing rational expectations because Lucas and Sargent suggested that errors in forming expectations of inflation are the key to understanding the business cycle.

The first article to analyze the Livingston Survey's inflation forecasts was written in 1970 by Stephen Turnovsky. Ahead of his time, Turnovsky performed the first bias tests on the Livingston Survey data. Looking at data from 1954 to 1969, he found that inflation forecast errors were large in the late 1950s, but much smaller in the 1960s. When Turnovsky ran the econometric tests for bias (described in the Appendix), he found that, on average, inflation was forecast about 1 percentage point too low from 1954 to 1964. But he found no bias from 1962 to 1969. He suggested that something changed around 1960 to improve the inflation forecasts. Perhaps people didn't have much incentive to forecast inflation well before then, since inflation was low on average.

In a 1975 article, James Pesando evaluated the Livingston forecasts and found them to be biased in a certain way. Pesando used the data from 1959 to 1969 and showed that the 12-month-ahead forecasts weren't consistent with the six-month-ahead forecasts. When he ran bias tests as Turnovsky did, he found the six-month-ahead forecasts weren't biased, but the 12-month-ahead forecasts were. Pesando was the first person to suggest that perhaps the Livingston Survey wasn't representative of people's true forecasts. If the survey did represent true forecasts, people weren't rational, according to his statistical tests. And that's hard to believe because people would lose money in financial markets if they weren't rational.

It turned out that both Turnovsky and Pesando had used the Livingston Survey results as Livingston had published them in his newspaper column. But John Carlson of Purdue University found two serious problems with the data, as he reported in his 1977 study. First, Livingston didn't report the data quite as his respondents sent them in. In particular, if revised data came in between the time at which people made their forecast and when his column was published, Livingston modified the forecasts he reported. This was a pragmatic way for a journalist to deal with the problem of revised data, but it caused many problems for economists. So Carlson was able to get the original data from Livingston and handle it properly for statistical testing.

Carlson also pointed out that although people were treating the forecasts as covering six-month and 12-month periods, that wasn't accurate. For example, when the questionnaires were mailed out for the December survey, respondents knew only the October consumer price index (most of the time). But they were

asked to forecast the index for the following June and December, so their forecasts were really for eight months (October to June) and 14 months (October to the following December).

This methodological flaw in the survey created problems for researchers, especially since, at times, some forecasters knew only the October index when they made their forecasts, while others knew the November index. Prior to 1969, the CPI was released even later, so the forecasters may have had even less information when sending in their forecasts. This makes testing for rational expectations tricky. The flaw was finally corrected in 1992 after the Philadelphia Fed had run the survey for several years. The solution was simple: ask forecasters to forecast the current December value for the index (as well as June and the following December), so that there are definite six- and 12-month forecasts. Also, the timing of the survey was tightened to ensure that all responses to the survey were received before the CPI for May or November was released; this ensured that, for consumer price inflation at least, all forecasters had the same data.

When Carlson used corrected data and reran Pesando's tests, he found that the inflation forecasts were even worse. But Pesando's procedure had made a very strong assumption about how inflation behaved over time.[1] Carlson allowed the inflation process to change over time and compared a statistical forecast with the Livingston Survey forecasts. He found that the survey forecasts were generally better than the statistical forecasts. While this fact underlines the value of the survey, Carlson also noted the very large forecast errors in the early 1970s, a time when inflation rose to much higher levels than anticipated.

Key Rationality Tests. The next important paper in this literature came in 1978 from Don Mullineaux, then at the Philadelphia Fed. Mullineaux argued that Pesando's test (the one replicated by Carlson using the corrected data) was statistically flawed. He tried a closely related (but statistically correct) test and found no irrationality in the Livingston Survey forecasts.

Mullineaux's results in support of the survey's forecasts were almost immediately challenged by Douglas Pearce in 1979, who, using a different methodology, found that inflation forecasts were biased. Pesando, Carlson, and Mullineaux all assumed that inflation followed a particular statistical process. But a look at the data suggests that inflation follows a very different process.[2] And when Pearce generated statistical forecasts using this alternative process, they were much better than the Livingston Survey forecasts. Thus, the Livingston forecasts weren't efficient because they didn't use information on past inflation to form better forecasts. Pearce struck a strong blow against the accuracy of the Livingston Survey forecasts of inflation.

Some of the issues in the literature on rational expectations are fairly technical and thus impossible to discuss in a survey article like this one. But one technical issue turns out to be vitally important in carrying out the tests described above. Economists call it the problem of overlapping observations. Suppose in June and December 1994 you made forecasts for the next year. That is, in June 1994 you made a prediction for the consumer price index from June 1994 to June 1995, and in December 1994 you made a forecast from December 1994 to December 1995. Notice that the time periods of the two

[1]Pesando assumed that inflation was related to past inflation in a fixed pattern for the preceding five periods (each period is a half year). In technical terms, Pesando assumed that inflation was a five-period autoregressive process.

[2]The assumption was that inflation followed what is known as an autoregressive process, but Pearce's estimates suggested that a moving-average process represents the data better. The difference is important both theoretically and in practice.

forecasts overlap somewhat — both include the period from December 1994 to June 1995.

Now suppose some disaster had struck the economy during that overlapping period. For example, suppose that OPEC tripled oil prices in March 1995. Then it's likely that consumer inflation would have risen sharply in March 1995, and your forecasts for both June 1994 to June 1995 and December 1994 to December 1995 would be far too low. This is the overlapping observations problem: a shock to the variable being forecast affects the forecast errors for several forecasts, not just one. This leads to problems in interpreting both plots of the data and econometric tests for bias and efficiency.[3] Bryan Brown and Shlomo Maital, in a 1981 article, showed how to deal with this problem using a sophisticated statistical procedure. They illustrated their results with the Livingston Survey, testing forecasts for inflation and several other variables. They found the Livingston Survey forecasts to be largely free of bias. They did find, however, that the forecasts are a bit inefficient; they don't use information about monetary policy (the growth rate of the money supply) to form better forecasts.

In a 1981 article, Stephen Figlewski and Paul Wachtel suggested that rationality tests should be performed on individual forecasts, not the average across individuals, which had been used by previous researchers. Using the average across forecasters could lead to an error in evaluating whether individual forecasters have rational expectations; it's possible that the individual forecasters could have rational expectations, but the average across forecasters would be biased, or vice versa. When Figlewski and Wachtel looked at the individual forecast data from the participants in the survey, they found overwhelming evidence that the forecasts were biased and inefficient.

Many other papers found fault with the Livingston inflation forecasts. One of the more devastating notions was that simple models could do better than the Livingston Survey, as Pearce had found. In a notable 1984 paper, Eugene Fama and Michael Gibbons showed that simple models using interest rates could forecast inflation better than the Livingston Survey. They implied that because interest rates are formed in financial markets, they provide a better measure of expected inflation than the survey, which questions people who may not have a strong incentive to forecast well. As Frederic Mishkin put it in 1981, "One obvious danger with survey data is that there may be very little incentive for the respondents to answer accurately."

But recently, both the Livingston and other survey forecasts are being used more and more in the popular press and by researchers.[4] One reason is the realization that the negative results on the Livingston forecasts came during the 1970s and early 1980s, when inflation was very erratic because of oil-price shocks. For that period, the forecasts look bad and seem to be biased or inefficient. Since that time, however, inflation has become much more predictable. Rerunning the same tests for rationality (using the average across forecasters) as before leads to a much more favorable view of the Livingston forecasts, as I showed in a 1996 article.

Testing Other Forecast Variables. Though testing the inflation forecasts from the Livingston Survey has been its most important use, other variables in the survey have also been examined, including the stock market and real output.

[3]The problems arise because a shock, such as the oil-price shock in the example, that occurs in the overlapping period gives rise to forecast errors that are related to each other. But any relationship between forecast errors violates the assumptions needed to run the econometric tests for bias and efficiency described in the Appendix.

[4]See *Comparison to Other Surveys* for a discussion of some of the other forecasting surveys and how they differ from the Livingston.

Comparison to Other Surveys

The Livingston Survey isn't the only survey of macroeconomic forecasts available, though it is the granddaddy of them all. Interest in the Livingston Survey led to the formation of many other surveys.

The ASA/NBER Business Outlook Survey, now known as the Survey of Professional Forecasters, started in 1968 as a joint venture of the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER). University of Chicago professor Victor Zarnowitz was instrumental in getting the survey started in the fourth quarter of 1968. The survey was taken around the middle of each quarter and provided more detailed forecasts than the Livingston Survey. Participants were asked to provide forecasts for the current quarter, and one, two, three, and four quarters ahead, for many major macroeconomic variables. In 1981, the survey was extended to include all the components of real GNP. Unique to the survey is a set of questions about the probabilities that forecasters attach to different changes in GNP (GDP since 1992) and the price level. The Federal Reserve Bank of Philadelphia took over the survey in 1990. For further information, see my 1993 article.

In 1976, Robert Eggert discovered that interest in forecasts was so great that people wanted a more frequent survey of forecasts. His Blue Chip Economic Indicators has been a big success, and much research has been based on this survey. It has the advantage of being monthly, thus providing a lot more information than the Livingston Survey. But since it hasn't been around as long, researchers who need a longer time series can't use it. Nonetheless, many of the tests for bias and efficiency carried out on the Livingston Survey have also been done on the Blue Chip survey, with similarly mixed results.

The National Association of Business Economists (NABE) produces a survey of its members every quarter. It consists of annual forecasts for many macroeconomic variables and asks a variety of topical questions. The survey has been going on since the early 1960s and NABE maintains a database of the results.

Business publications such as the *Wall Street Journal* and *Business Week* also collect surveys of economists, although researchers are just now beginning to study their results.

The Livingston Survey's stock market forecasts have been examined in a number of articles. In 1980, Josef Lakonishok found that the forecasts (averaged across the forecasters) were biased and inefficient. In 1984, Douglas Pearce confirmed those results using the individual forecast data. But in 1989, Yoon Dokko and Robert Edelstein reached a contrary conclusion. They claimed that both earlier studies had miscalculated the expected returns to the stock market because survey participants had not provided a forecast for the current June or December value for the stock market index. For example, previous researchers assumed that in the June survey, forecasts were made on a particular day, but we know that, in fact, different participants sent in their forecasts on different days. Since day-to-day movements of the stock market may be large, each survey participant was basing his or her forecast on a different base for the stock index.

To avoid this problem, Dokko and Edelstein used just the information on expected growth between the forecast of the stock market for the end of June and the end of December, which gives an exact six-month expected return, with no question about the forecast's horizon. When Dokko and Edelstein ran tests on their data, they found no bias or inefficiency in the Livingston forecasts. These findings suggest that the results of all these statistical tests are very sensitive to researchers' assumptions.

In a recent paper I've written with Laurence Ball of Johns Hopkins University, we examined the real GDP forecasts (as well as the inflation forecasts) from the Livingston Survey and how

they are influenced by monetary policy. While our results suggested that the survey forecasts were unbiased, they also indicated that the forecasts weren't efficient with respect to monetary policy. That is, the forecasters didn't fully use the information they had about monetary policy to form better forecasts. Instead, they seem to make small but systematic errors. When monetary policy changes: (1) they don't forecast a large enough change in output; and (2) they assume that inflation will respond to the change faster than it actually does. The results suggest that the forecasters may not have rational expectations. But other explanations for the forecast errors are also reasonable, such as the fact that changes in the financial structure of the economy have led to confusion about which variable accurately measures monetary policy.

Other Uses of the Livingston Survey. All of the studies discussed above tested the quality of the Livingston Survey's forecasts, especially to see if they were consistent with rational expectations theory. However, the survey has been used to test other economic theories. The forecasts have been used to study the relationship between nominal and real interest rates, to test the response of stock market forecasts to higher expected inflation, and to test theories of consumption spending.

The Fisher effect (named after economist Irving Fisher) describes the relationship between nominal and real interest rates. The nominal interest rate consists of two components: the real interest rate plus the expected rate of inflation. According to the theory of the Fisher effect, when the expected inflation rate rises one percentage point, the nominal interest rate should rise proportionally. This theory has been tested using the Livingston series on expected inflation. In 1972, William Gibson showed that a one-percentage-point increase in expected inflation would raise interest rates on bonds with maturities of a year or less about one percentage point, confirming the Fisher effect. But Gibson's study was flawed because it was based on the data before Carlson's corrections were made, so Vito Tanzi redid Gibson's work with the corrected data. Tanzi's work did not support Gibson's results; he found that nominal interest rates don't rise one-for-one with expected inflation. Rather, both the state of the business cycle and the tax rate influenced the relationship between the nominal interest rate and the expected inflation rate.

In a 1983 study, N. Bulent Gultekin, using the Livingston Survey expectations of both inflation and stock returns, found that expected stock returns rose point-for-point with rises in expected inflation. Economic theory predicts this result, but previous empirical research, which hadn't used the Livingston forecasts, didn't support the theory.

Finally, a very influential paper in macroeconomics used the Livingston Survey to test the theory that people change the timing of their consumption expenditures depending on changes in interest rates. Robert Hall used the survey to represent the expected real returns to stocks and to short-term Treasury securities. Hall's results surprised many economists because they suggested that consumption spending didn't respond much to changes in expected real interest rates. Previous studies had reached the opposite conclusion.

SUMMARY

The Livingston Survey has proven to be a valuable tool for measuring economists' forecasts of macroeconomic variables over the past 50 years. It was instrumental in testing the theory of rational expectations in the late 1970s and early 1980s. While the survey's inflation forecasts seemed to fail the tests of rational expectations 15 to 20 years ago, more recent tests show that the forecasts are much better, though perhaps not perfect. The survey has also been used to test the rationality of forecasts of other variables and to test theories of interest rates, stock returns, and consumption spending.

REFERENCES

Ball, Laurence, and Dean Croushore. "Expectations and the Effects of Monetary Policy," Federal Reserve Bank of Philadelphia Working Paper 95-22, October 1995.

Brown, Bryan W., and Shlomo Maital. "What Do Economists Know? An Empirical Study of Experts' Expectations," *Econometrica* 49 (March 1981), pp. 491-504.

Carlson, John A. "A Study of Price Forecasts," *Annals of Economic and Social Measurem*ent 6 (Winter 1977), pp. 27-56.

Croushore, Dean. "Introducing: The Survey of Professional Forecasters," Federal Reserve Bank of Philadelphia *Business Review* (November/December 1993).

Croushore, Dean. "Inflation Forecasts: How Good Are They?" Federal Reserve Bank of Philadelphia *Business Review* (May/June 1996).

Dokko, Yoon, and Robert H. Edelstein, "How Well Do Economists Forecast Stock Market Prices? A Study of the Livingston Surveys," *American Economic Review* 79 (September 1989), pp. 865-71.

Fama, Eugene F., and Michael R. Gibbons. "A Comparison of Inflation Forecasts," *Journal of Monetary Economics* 13 (1984), pp. 327-48.

Figlewski, Stephen, and Paul Wachtel. "The Formation of Inflationary Expectations," *Review of Economics and Statistics* 63 (February 1981), pp. 1-10.

Gibson, William E. "Interest Rates and Inflationary Expectations: New Evidence," *American Economic Review* 62 (December 1972), pp. 854-65.

Gultekin, N. Bulent. "Stock Market Returns and Inflation Forecasts," *Journal of Finance* 38 (June 1983), pp. 663-73.

Hall, Robert E. "Intertemporal Substitution in Consumption," *Journal of Political Economy* 96 (April 1988), pp. 339-57.

Lakonishok, Josef. "Stock Market Return Expectations: Some General Properties," *Journal of Finance* 35 (September 1980), pp. 921-31.

Lucas, Robert E., Jr. "Econometric Policy Analysis: A Critique," Carnegie-Rochester Conferences on Public Policy 1 (1976), pp. 19-46.

Lucas, Robert E., Jr., and Thomas J. Sargent. "After Keynesian Macroeconomics," Federal Reserve Bank of Minneapolis *Quarterly Review* (Spring 1979).

Mishkin, Frederic S. "The Real Interest Rate: An Empirical Investigation," Carnegie-Rochester Conference Series on Public Policy 15 (Autumn 1981), pp. 151-200.

Mullineaux, Donald J. "On Testing for Rationality: Another Look at the Livingston Price Expectations Data," *Journal of Political Economy* 86 (April 1978), pp. 329-36.

Muth, John F. "Rational Expectations and the Theory of Price Movements," *Econometrica* 29 (July 1961), pp. 315-35.

REFERENCES (continued)

Pearce, Douglas K. "Comparing Survey and Rational Measures of Expected Inflation," *Journal of Money, Credit and Banking* 11 (November 1979), pp. 447-56.

Pearce, Douglas K. "An Empirical Analysis of Expected Stock Price Movements," *Journal of Money, Credit and Banking* 16 (August 1984), pp. 317-27.

Pesando, James E. "A Note on the Rationality of the Livingston Price Expectations," *Journal of Political Economy* 83 (August 1975), pp. 849-58.

Sargent, Thomas J., and Neil Wallace. "Rational Expectations and the Theory of Economic Policy," *Journal of Monetary Economics* 2 (April 1976), pp. 169-83.

Tanzi, Vito. "Inflationary Expectations, Economic Activity, Taxes, and Interest Rates," *American Economic Review* 70 (March 1980), pp. 12-21.

Taylor, Herb. "The Livingston Surveys: A History of Hopes and Fears," Federal Reserve Bank of Philadelphia *Business Review* (January/February 1992).

Turnovsky, Stephen J. "Empirical Evidence on the Formation of Price Expectations," *Journal of the American Statistical Association* 65 (December 1970), pp. 1441-54.

APPENDIX
Testing for Bias and Efficiency

A number of different tests have been used in the literature on bias and efficiency. The most well known are fairly simple and are described in this appendix. For more complicated tests, the reader can refer to the research studies described earlier in this article.

Given a set of forecasts, F_t, for a variable X_t, a simple bias test involves running the regression $X_t = \alpha + ß F_t + \varepsilon_t$. If the estimated value of α is 0 and the estimated value of ß is 1, the forecasts are unbiased. Formal statistical tests can be performed to see if the estimated value of α is close enough to 0 and the estimated value of ß is close enough to 1 to consider the forecasts to be unbiased.

An efficiency test requires determining if the forecast errors ($X_t - F_t$) are systematically related to some other variable that was known to the forecaster when the forecasts were made. Suppose this other variable is denoted Z_t. Then the efficiency test requires running a regression of the form $X_t - F_t = \gamma + \delta Z_t + \upsilon_t$. If the estimated value of δ is significantly different from 0, the forecasts are inefficient. It should be possible to use the existing forecasts with the information about Z_t to form a better forecast.

Bias and inefficiency are different concepts, so forecasts can be biased but efficient or they can be unbiased but inefficient.

The
LIVINGSTON
S U R V E Y

Federal Reserve Bank of Philadelphia

December 1996

Economy is holding steady

The 52 forecasters in the semi-annual Livingston Survey forecast a growth rate of 2.3 percent for the economy in the last quarter of 1996. Growth for the first half of 1997 is forecast to be 2.0 percent, with a slight increase to 2.3 percent in the second half of the year. Average growth for all of 1997 is forecast to be 2.1 percent.

Unemployment rate shows little change

Economists forecast little change in the unemployment rate over the next year: 5.3 percent in June and 5.4 percent in December 1997. Only a modest increase is forecast for 1998 when the unemployment rate is expected to average 5.6 percent.

Inflation expectations are revised slightly downward

Survey participants forecast a decline in inflation over the next year. Consumer price inflation as measured by the Consumer Price Index is expected to be 3.1 percent in the first half of 1997. Economists forecast a moderate drop in inflation, to 2.9 percent, in the second half.

The forecast for inflation over the next 10 years fell to 3.0 percent, a 0.1 percent decrease from the forecast in the last survey.

The Producer Price Index (PPI) also shows little change in inflation. Economists forecast that the PPI growth rate will be 2.4 percent in both the first and second halves of 1997.

Wage growth is expected to rise

Wages stay in line with the inflation rate—3.1 percent—for the first half of the year. But in the second half of 1997, wage growth is expected to increase to a rate of 3.3 percent, slightly higher than inflation.

Interest rates change little

Survey forecasters predict a slight increase in the U.S. Treasury bond's interest rate, from 6.6 percent at the end of 1996 to 6.7 percent by the end of June 1997. However, they forecast that rates will fall back to 6.6 by the end of 1997 and fall further, to 6.4 percent, by the end of 1998. Forecasts for the Treasury bill show the same pattern: 5.1 percent at the end of 1996, 5.2 percent at the end of 1997, and 4.9 percent at the end of 1998.

Stock market will continue to expand

Economists forecast continued expansion in the S&P 500 at a level of 738.5 for the end of 1996, 750.6 for the end of 1997, and 763.4 for the end of 1998. This translates into a growth rate of 1.6 percent for 1997 and 1.7 percent for 1998.

Economic Research Division Ten Independence Mall, Philadelphia, PA 19106-1574