

The accuracy of the BLS productivity measures

Overall productivity growth may be somewhat underestimated; despite continued progress, measurement and conceptual barriers remain

Edwin R. Dean

The Bureau of Labor Statistics has made sustained efforts to improve its productivity measures. The goals of these efforts have been to enhance the reliability of the measures; to facilitate analysis of economic performance; and to provide useful information to the public.¹ The BLS clearly recognizes that, despite the beneficial results of its program, there is room for further improvement.

In the past several years, some have voiced concerns about the accuracy of the trends in the BLS productivity series, mainly to suggest that productivity growth for the business sector of the economy has been understated. The BLS has also been concerned about the accuracy of the major sector productivity trends and has devoted considerable effort to examining the accuracy of these trends.

The concerns about possible underestimation of productivity growth have been focused on data for the business sector of the economy, and especially the services components of that sector. "Services," broadly defined, include all producing activities outside the "goods" sector. The major services-producing activities are transportation, communications, utilities, retail and wholesale trade, finance and insurance, and various additional services rendered to persons and businesses.² Commentators have wondered why productivity in services has not grown nearly as rapidly as productivity in manufacturing, particularly in light of anecdotal indications of improvements in several types of services.

Concerns have also been expressed about several measurement techniques used by the BLS to

compute productivity trends. These concerns are directed to such questions as whether the BLS productivity data fully reflect changes in the quality of goods and services; whether the best techniques are used to introduce new, advanced products into the data series; and whether the BLS methods capture the full impact of new information technology on economic performance. Accordingly, several points in this paper attend briefly to these issues.

Five key issues

This article addresses five important issues related to the BLS measure of productivity in the business sector of the economy. First, the article examines whether there is in fact mismeasurement of productivity growth in the services portion of the economy. The article concludes that there are important measurement problems in some service activities and these problems may be leading to underestimation of productivity growth rates.

Second, the article addresses specific sources of mismeasurement, the sectors of the economy where these problems are found, and the possibility of determining the extent of underestimation. It concludes that there is no basis for determining the extent of the underestimation. Although existing information sheds some light on the magnitude of the problem, there is no basis for a precise determination of its extent. The available information does not indicate that the published data understate productivity growth by a large order of magnitude.

Edwin R. Dean is Associate Commissioner, Office of Productivity and Technology, Bureau of Labor Statistics.

Third, the article discusses what can be done to improve the measurement of productivity in services. It describes a number of recent improvements in the quality of Federal Government statistics that have led, in turn, to improvements in the productivity data. In addition, it discusses steps that can be taken to improve further the productivity measures, especially within the services sector.

Two other significant issues cut across discussion of these three matters. One is the impact of alleged biases in the consumer Price Index on productivity data. The other is the intrinsic difficulty of defining output for a number of service activities, activities that have been labeled “hard-to-measure” services. The discussion of all of these questions draws heavily on the articles by William Gullickson and Michael Harper (pp. 47–67), and by Lucy Eldridge (pp. 35–46) in this issue.

How the trends differ

The most effective way to begin a discussion of these measurement issues is to compare the labor productivity trends for two major sectors, the business sector and manufacturing, one of its components.³ Following are the average annual percentage rates of growth of labor productivity, or output per hour of work, for these two sectors:

	<i>Business</i>	<i>Manufacturing</i>
1949–73	3.3	2.6
1973–90	1.2	2.4
1990–98	1.4	3.7

In the business sector, the growth rate of output per hour declined from the robust rate of 3.3 percent prior to 1973 to slightly more than 1 percent in the years following 1973. This decline is often called the “productivity slowdown.” Another noteworthy result of these data is the contrast, after 1973, between continued robust growth in manufacturing productivity and the sharp deceleration of growth in overall business sector productivity.

The pattern of changes in the trends in the BLS series on multifactor productivity resembles that of the changes in labor productivity trends. Multifactor productivity is output per unit of all inputs combined, including labor, capital, and other inputs.⁴ Following are the average annual percentage rates of growth of multifactor productivity for the private business sector and manufacturing:

	<i>Private business</i>	<i>Manufacturing</i>
1949–73	2.1	1.5
1973–903	.5
1990–965	1.7

The contrasts observed in the labor productivity data are found also in the multifactor productivity data. The productivity growth rate in private business fell off markedly after 1973. The growth rate in manufacturing also fell off, but the decline was less sharp than the decline in the private business sector. After 1979, the manufacturing productivity growth rate recovered, and has even exceeded the pre-1973 rate, while the growth rate for the business sector remained quite low. Hence, for both labor and multifactor productivity, the data since 1973 show a superior productivity performance in manufacturing, compared with the larger sector of which it is a component.⁵

A wide variety of data

The methods used for the quarterly labor productivity series are fairly straightforward. Two types of information are needed to construct this series: output and labor input measured in hours worked.

For the business sector output series, BLS uses national income and product accounts (national accounts) prepared by the Bureau of Economic Analysis (BEA) of the Department of Commerce. BEA prepares its national accounts series by drawing on a wide variety of information prepared by many other Federal statistical agencies.⁶ A prime example is their computation of real expenditures on consumer goods and services. BEA prepares data on real expenditures on many goods and services by using price indexes to adjust current-dollar expenditures to remove the influence of changes in the prices at which these goods and services are sold. The use of price indexes in this way is known as deflation. The price data come primarily from the BLS. In particular, BEA draws on the BLS Consumer Price Index (CPI) series to deflate expenditures on many consumer goods and services. If there are problems underlying either the expenditure or price data for the service sector, this may affect the reliability of the national accounts information produced by the BEA and, in turn, the productivity data produced by the BLS.⁷

The BLS series on manufacturing output is constructed from an even wider variety of sources.⁸ The BLS draws on annual production data, in current dollars, prepared by the Bureau of the Census. The deflation of these data uses price information received from the BEA which, in turn, compiles its information mainly from the BLS producer price indexes. However, the BEA constructs its own quality-adjusted price series for a number of manufactured goods. In addition, the BLS manufacturing output series makes use of the BEA series on inventories as well as the BEA input-output tables.⁹

Labor input information for the BLS major sector productivity series is computed mainly from data on employment and average weekly hours paid, collected through the BLS monthly survey of employers. The data on weekly hours paid

are adjusted to hours actually worked using the BLS Hours-at-Work Survey. The labor input series also draws on employment and hours data from the BLS monthly survey of households.

Productivity trends

Two considerations establish a basis for believing that recent trends in business sector productivity growth might be underestimated. The first is the comparison of productivity growth rates for the business sector and for manufacturing for the years since 1973. The implied residual productivity growth trend for nonmanufacturing—that is, the trend in the remainder of the business sector after the manufacturing trend is removed—must be lower than the trend for business as a whole.¹⁰ As some observers have noted, the trend for this residual must then be very modest indeed, or perhaps even negative. It seems unlikely that nonmanufacturing productivity growth would be very low or negative for a time span exceeding two decades.

The second consideration is the practice of estimating trends in real expenditures for some components of gross domestic product (GDP) on the basis of input information. It should be clear that in order to produce accurate labor productivity series, it is critical for the information on output and labor input trends to be derived independently of each other. If they are not independent, the resulting trends in productivity will be incorrect. For many years, economists in the BEA and the BLS have been aware that, for some portions of GDP, trends in output were derived partly by using trends in input information. It is largely for this reason that the business sector of the economy is the largest sector for which productivity series are published. Components of GDP outside the business sector are, in several instances, computed in part on the basis of input information. This is the case for general government, nonprofit institutions, and paid employees of private households.¹¹ In addition, however, there are several relatively small, yet significant, components of business sector output that are estimated in part on the basis of input information.

These two considerations have led some analysts to believe that the BLS labor productivity trends for the business sector may be biased downward. This belief is reinforced by indications that for a number of service industries the technologies used to produce outputs, as well as the services themselves, have changed significantly in recent years. Further, in recent years there has been evidence of economic prosperity, including the rapid growth of corporate profits, that does not seem fully consistent with slow productivity growth.¹²

New light is shed on both of these considerations in the careful study by Gullickson and Harper. Their article includes

an examination of multifactor productivity trends in nonmanufacturing. The authors approach this matter by assessing the contribution of multifactor productivity growth in the manufacturing and nonmanufacturing sectors to the U.S. private business sector as a whole. The study finds that since 1979, nonmanufacturing industries have contributed very little to business sector multifactor productivity growth. In one of their analytical exercises, this contribution was zero from 1979 to 1990 and zero also from 1990 to 1996. Hence, manufacturing accounts for all of the modest growth in overall multifactor productivity that has occurred during this period. Such a result seems improbable.

Gullickson and Harper also examine multifactor productivity growth in specific nonmanufacturing industries, mostly service-producing industries. The industry data used are mostly at the two-digit level of the industry classification system—examples of two-digit industries are metal mining and water transportation—although in several cases, groups of two-digit industries are examined.

Gullickson and Harper present estimates of multifactor productivity for 34 industries and sectors. One of the sets of estimates incorporates as much national accounts data as possible; another relies more heavily on data from BLS sources. In the first case, a total of 11 of these 34 industries have negative average annual multifactor productivity growth rates for the period 1977–92. In several cases, these negative growth rates are just slightly under zero—two industries have average annual growth rates of –0.1 and –0.3 percent. In other cases, however, the rates are strongly negative, in the range of –1 to –3 percent. In the estimates which rely more heavily on data from BLS sources, 13 of the 34 industries show negative multifactor productivity growth rates. In both sets of estimates, negative productivity growth rates are found for important service-producing industries that account for a sizable portion of the private business sector output.

It would be wise to be cautious about drawing firm conclusions based on these findings. To develop estimates of multifactor productivity for these industries, it was necessary for Gullickson and Harper to work with data of limited reliability—a limitation commonly faced by researchers who work on service industries. This problem could affect their conclusions about the specific industries that contribute the most to sluggish overall productivity growth. In spite of these limitations, Gullickson and Harper used these data to develop estimates of multifactor productivity growth in industries in order to try to identify the specific industries with serious measurement problems. This will ultimately indicate which industries should be the subjects of intensive data-improvement efforts.¹³ In preparing these multifactor productivity estimates, the BLS does not mean to imply that it regards the resulting data as suitable for official publication.¹⁴ Further, it should be empha-

sized that negative productivity growth is quite possible in the real world. For example, negative productivity growth can occur over lengthy spans of time because of declining demand for an industry's products, accompanied by idle capital plant and equipment in the industry.

Negative growth possible, not probable

Even a cautious interpreter of these data, though, is strongly tempted to reach two general conclusions. First, negative multifactor productivity growth in so many industries over a period of 15 years seems implausible. There is probably something wrong with at least some of the data underlying these calculations.

Second, it appears likely that for some industries with questionable measured productivity trends, the problems originate with faulty real output series, rather than with the associated input data. One reason for this tentative conclusion is that there is no particular reason to believe that these low industry productivity trends originate with the input data. A second is that there are obvious sources of bias in the data underlying the output series for some industries. Both of these reasons deserve further discussion.

The low multifactor productivity growth rates found for some industries conceivably could be a result of overestimates of the growth rates of inputs. For example, the growth rates of capital inputs are computed using methods and data known to be imperfect. Also, problems arise from the small samples in some of the surveys used to compute employment and hours data. In particular, the hours worked by the self-employed are computed from a small portion of the Current Population Survey (CPS) sample. Also, hours paid are converted to hours worked through information from the BLS Hours-at-Work Survey, a survey using a small sample.¹⁵ In addition, the multifactor productivity data for industries are computed without any adjustment for changes in the quality of the labor force; some observers have suggested that in recent decades, labor force growth outside of manufacturing has occurred through the hiring of inexperienced workers with relatively little education.

While observations such as these have merit, it appears unlikely that they could account for most of the negative multifactor productivity growth rates outside manufacturing. The weight of capital inputs in total inputs in most industries is so small that even a substantial overestimation of capital input growth rates would not make much difference to multifactor productivity growth rates in these industries.¹⁶ While some of the surveys underlying the data on hours of labor input are subject to sampling error and other problems, there is no specific reason to expect an upward bias in the long-term growth rates of these hours. Regarding trends in the skills composition of the labor force, it is certainly possible that, in

some particular service industries, the average skills might have fallen rather than increased. If this decline did in fact take place, the computed multifactor productivity growth rates for those particular industries might be underestimates. For the total private business sector, however, the BLS series on labor composition indicates that the average skills of the labor force have increased, rather than decreased, in recent years.¹⁷ This would indicate that the trend in measured labor input, uncorrected for skills, has probably underestimated—rather than overestimated—the trend in skills-adjusted labor input. This problem could affect the multifactor productivity results for some specific industries, but is unlikely to be the explanation of most of the negative multifactor productivity trends found by Gullickson and Harper.

A second reason for believing that the productivity measurement problems for some industries originate with the real output series is that there are obvious sources of bias in the data underlying some of the output series. In particular, biases might be created by the estimation of components of output on the basis of input information. These estimation procedures are (i) deflation procedures carried out with indexes of input costs, in place of price indexes for the output categories, and (ii) extrapolation procedures for estimating output trends, where the extrapolation is based on trends in the quantity of inputs. In particular, extrapolation of output by use of labor input and deflation of current-dollar series by labor cost indexes come close to embodying an assumption of zero labor productivity growth. For example, if labor input data are used to extrapolate output trends, this will tend to yield a labor productivity trend that will be biased toward zero. This is so because the output measure as well as the input measure—in other words, the numerator as well as the denominator of the productivity ratio—will reflect the trend in labor input. BEA is quite aware of the limitations of these input-based techniques from the perspective of productivity measurement, and is working to develop the missing price indexes along the lines set forth in its formal strategic plan.¹⁸

For two industries, construction and banking, that appear to contribute substantially to negative business sector multifactor productivity growth, the methods used to construct real output trends are problematic. For the construction industry, Gullickson and Harper compute substantial portions of the output data by using input-cost indexes to deflate current-dollar output data. According to Gullickson and Harper, the contribution of negative multifactor productivity growth in construction to the multifactor productivity growth in the private business sector is about -0.1 percent.

The method Gullickson and Harper used to compute the output trend for the largest part of banking is the extrapolation of base-year output with “paid employee hours”; BEA prepares significant parts of its banking data by the same process

of extrapolation with employee hours. Under certain conditions, methods that tend to produce a bias towards a zero growth rate in labor productivity will yield a bias toward negative multifactor productivity. A negative multifactor productivity trend will appear if the aggregate of nonlabor inputs grows more rapidly than labor input. And, indeed, in the data used by Gullickson and Harper this is the case in the banking industry; if banking multifactor productivity were assumed to have a zero growth rate, in place of the implicit assumption of a labor productivity growth rate of zero, then multifactor productivity in the private business sector would have increased more rapidly in the 1977–92 period. However, the banking industry measurement problem should not be overemphasized: according to Gullickson and Harper, the contribution of negative multifactor productivity growth in banking to the multifactor productivity growth in private business is fairly small, a negative 0.09 percent.

It seems clear then, that two tentative conclusions are justified. First, negative multifactor productivity trends in a number of industries over a period of 15 years appear to support the belief that at least some of the data underlying the private business sector productivity series are faulty. And second, in a few industries the methods used to construct output trends may yield labor productivity measures that are biased toward zero.

Sources of measurement problems

The service-sector measurement problems will be better understood if their sources can be located. Further, an understanding of their sources may provide clues as to the best ways of dealing with these problems.

Gullickson and Harper help locate these problems. It was noted above that in one of their computations, those based directly on BEA's national accounts data, Gullickson and Harper found that a total of 11 out of 34 industries appeared to have negative multifactor productivity growth rates for the period 1977–92; 9 of these 11 industries are service-producing industries. In a second set of estimates, 13 of the 34 industries appeared to have negative multifactor productivity growth rates; 11 of the 13 are in the service sector. In both estimates, the nonservice industries with negative multifactor productivity were construction and oil and gas extraction.

Gullickson and Harper then engage in an exercise that sheds further light on the sources of the measurement problem. The exercise examines the relationship between their estimates of negative multifactor productivity for individual industries and the overall multifactor productivity growth rate for the private business sector. Their experiment can be described as follows: if we substitute a zero productivity growth rate for all industries that are showing negative multifactor productivity trends, what would happen to overall multifactor productivity

growth? This exercise has some interesting implications for the growth rate of multifactor productivity in the private business sector as a whole, as well as the contributions of the specific industry revisions to this revised overall multifactor productivity growth rate for private business.

This exercise raises the private business sector annual average multifactor productivity growth rate by 0.4 percent during the 1977–92 period.¹⁹ Also, this exercise indicates that five specific industries—banks, insurance carriers, utilities, health services, and construction—contributed most to the increase in multifactor productivity in the private business sector as a whole. Four of those five industries are in the service sector; only the construction industry is outside the service sector.²⁰ In Gullickson and Harper's results, the utilities and health services industries had smaller negative effects on aggregate multifactor productivity than did the other three industries.

It is useful also to look at the procedures used by BEA to estimate output for specific industries. Two procedures are the use of input quantity indicators to extrapolate output trends and the use of input-cost indexes instead of output price indexes to deflate current-dollar output data. As noted earlier, such estimation methods help explain low or negative productivity trends in banking and construction. Based on estimates using 1997 data, Lucy Eldridge concludes that such input-based estimates are used in computing approximately 14 percent of business sector output. In addition to banking and construction, these methods are also used in computing output trends for portions of insurance. Hence, input-based estimates are used in three of the five industries contributing in an important way to overall low productivity growth.

The role of price indexes

Several commentators have suggested that biases in the BLS Consumer Price Index (CPI) may also be contributing to slow measured productivity growth. The report of an Advisory Commission to Study the Consumer Price Index (Boskin Commission), appointed by the U.S. Senate Finance Committee, took the position that there is an upward bias in this index.²¹ Several commentators, reacting to the Boskin Commission report, have correctly noted that the Bureau of Economic Analysis uses price data from the CPI to convert nominal GDP data to real GDP. They have noted, further, that the BLS uses GDP data as the output measure in some of its productivity series. So these analysts have concluded that an upward bias in price data implies that there is a downward bias in the BLS business sector productivity series.

Although the BLS has questioned many of the Boskin Commission's conclusions, Eldridge examines the channels through which possible biases in the CPI might be transmitted to the major sector productivity data. On the hypothesis that

the CPI is biased upward at specific rates, the paper also explores the extent of the possible bias in the productivity data. The following are some of Eldridge's most significant conclusions:

1. Components of the CPI are used to construct approximately 57 percent of the business sector output measure used for the BLS productivity statistics. (Eldridge reaches this conclusion by studying the 1997 business sector data.) This means, of course, that about 43 percent of the output data underlying the BLS business sector productivity index are unaffected by the CPI. For this substantial proportion of business sector output, the techniques used to construct the data underlying the trends in real output do not involve the CPI. The report of the Advisory Commission concluded that there is an upward bias of 1.1 percent in the CPI. If this conclusion were to be accepted as accurate, then Eldridge concludes from studying the specific ways that the CPI is used in the preparation of the business sector data, there would be a downward bias of 0.6 percent per year in business sector output growth.
2. However, due to a lack of strong supporting evidence for important elements of the Boskin Commission's analysis, this result should be viewed skeptically. The Boskin Commission, for example, adopted a 0.6-percentage-point adjustment for unmeasured quality change in goods and services and for the effects of the introduction of new products on the CPI. However, BLS has disputed the Boskin Commission's conclusions on this point and the Boskin Commission's suggestion of a 0.6-percent quality bias should be viewed as debatable.
3. Research conducted by BLS indicates that there is evidence to suggest that there was an upward bias of 0.2 percent in the CPI, arising from the specific methods used in computing the basic component subindexes of the CPI. Eldridge concludes that this would yield a small downward bias in the BLS business sector productivity series; also of about 0.2 percent per year. Starting with the CPI data for January 1999, the BLS introduced an improved computation method for constructing the sub-indexes for most components of the CPI in order to correct this bias.

The BLS has computed revised historical indexes incorporating the new "geometric means" method, and has provided these revised indexes to BEA. In July 1998, as part of its annual historical revision of national accounts information, the BEA incorporated these revised indexes into its measures of real personal consumption expenditures for the years 1995 forward to eliminate a downward bias arising from the previous CPI methodology. BEA has plans for incorporating the

geometric means results into the national accounts for years earlier than 1995.

While there are grounds for considerable skepticism concerning the Boskin Commission's view that unmeasured quality change has led to a 0.6-percent annual bias, it is still possible that not all quality change is captured by the methods currently used. Some observers have suggested that a large share of changes in services output has come in the form of increased quality or convenience, rather than in the form of increased quantity. These observers believe that a significant share of the measured increases in the prices of services represents increases in quality rather than price inflation. BLS researchers, in fact, have suggested that difficult measurement problems, related to quality adjustment, remain in the indexes for medical care and high-tech consumer goods.²² The BEA uses indexes from the CPI to compute a small portion of real expenditures on medical care; these portions of medical care expenditures are included in the private business sector data used by BLS to measure productivity.

Component indexes of the CPI are used to compute significant proportions of real expenditures for utilities and insurance and in computing a small part of real expenditures on banking. As noted earlier, these three industries are among the industries Gullickson and Harper found to have negative multifactor productivity growth rates that have contributed in an important way to sluggish measured overall productivity growth. However, BLS has not found that the CPI components used to compute real expenditures on the outputs of these industries are likely to suffer from quality bias.

Eldridge also examines the implications of the fact that the business sector of the economy excludes the services provided by government and by nonprofit institutions. This needs to be taken into account in any assessment of the impact of measurement problems in services on possible biases in the BLS productivity series. A large part of medical services is provided by nonprofits. Nonprofit institutions and government also provide most educational services, and nonprofit institutions account for most religious and welfare activities. Quality improvements and technological change affecting these activities cannot lead either to upward or downward bias in the BLS business sector productivity data.

It would be helpful, of course, if the available analyses formed the basis for estimating the extent of the underestimation of productivity growth in services. The available studies, however, indicate that a measurement problem exists, and they point to some of the sources of the problem, but not to the numerical dimensions of the problem. In particular, a tentative conclusion that a substantial number of service industries have negative multifactor productivity growth rates over a substantial period of time, and that this is not entirely plausible, does not provide a basis for determining the correct multifactor productivity growth rates.

It is useful, however, to take note of some of the specific findings of Gullickson and Harper and of Eldridge. The following findings in the two articles do shed some light on the possible magnitude of the measurement problem:

- Gullickson and Harper show that adjusting negative industry multifactor productivity trends to zero for the years 1977–92 would raise private business multifactor productivity growth by about 0.4 percent.
- The current procedure for estimating the output of banking probably results in an understatement of multifactor productivity growth in banking.
- Eldridge indicates that the use of geometric means for computing the basic subindexes of the CPI would appropriately raise the growth rate of business sector labor productivity by about 0.2 percent.

It should be emphasized that it would be incorrect simply to add these three results together to find an estimate of total mismeasurement. The banking industry is simply one of the industries included in the exercise involving upward adjustments of multifactor productivity. Also, it would not be correct to add the whole of the 0.2-percent result concerning geometric means to the results obtained from upward adjustment of multifactor productivity growth rates. It is possible that the CPI methods in use before BLS produced geometric means subindexes were among the reasons for negative multifactor productivity growth rates for some industries. Finally, it should be recalled that BEA has already incorporated these geometric mean indexes into a portion of the national accounts data for the years 1995 to the present. Presumably, if Gullickson and Harper’s methods were used to study revised national accounts data incorporating geometric means, the results would show somewhat higher productivity growth rates. The whole mismeasurement problem could thus be larger than 0.4 percentage points; conceivably, it could actually be smaller.²³

The “hard-to-measure” sectors

The published data on the output of service-producing industries all rest on an implicit or explicit definition of that output. The task of improving output and productivity measures for service activities includes the review of these output definitions and the development of better definitions. Many economists would agree that inadequate definitions of output underly the published data for many service activities.

It is easier, however, to find fault with the output definitions

currently used than to specify the correct definitions. Indeed, for a surprisingly large number of service-producing industries there is a lack of agreement among economists on the best definition of output. Economic literature has produced no consensus definitions for banking, insurance, other financial services, medical care, a variety of business and personal services, or retail and wholesale trade. The literature that discusses the difficulties of defining output in these industries is voluminous. Zvi Griliches of Harvard University, in his Presidential Address to the American Economic Association a few years ago, referred to the “unmeasurable” sectors of the economy. He argued that as these sectors have grown in importance, the economy has “shifted into uncharted waters.”²⁴ Mark Sherwood, of the Bureau of Labor Statistics, has provided a road map of the difficulties that face researchers as they attempt to develop improved definitions of output.²⁵

The problem of output definition can be illustrated by a brief discussion of banking. Frank Wykoff has set forth succinctly the problem of defining bank output.²⁶ He notes that it is clear that the banking industry performs a variety of quite distinct services for businesses and individuals. These services include the operation of a system for making payments; safekeeping of funds, securities, and other valuables; assessment of credit-worthiness and extension of credit; and making markets in money. It is difficult to summarize these functions within a single analytical framework and to distill a concept of output from such a framework. It is so difficult, in fact, that no one has done it in a way that is generally acceptable to experts in this field of research.

Wykoff illustrates these difficulties by discussing the various ways that different researchers would treat bank deposits in a measure of banking productivity. Five different treatments of bank deposits have been recommended: deposits are treated, variously, as inputs, outputs, both inputs and outputs, either inputs or outputs and neither inputs nor outputs. Viewed in the light of this lack of agreement on the measurement of banking output, it is easy to understand why both the BEA and the BLS have opted for straightforward and simplified means of producing data on banking output. The BEA procedure extrapolates part of the bank output data by the use of input data and the BLS banking industry productivity measure includes an output measure that rests on counts of specific banking industry transactions.

Another complication in measuring services outputs arises from the possibility that some cannot be defined adequately without considering the role of the consumer. The role of the consumer of services may well be different from the role of the consumer of goods. Several studies discussed by Sherwood have argued that in the production of services, the consumer often supplies an essential input.²⁷ For example, it may not be possible to define medical output adequately with-

out considering whether the patient follows the doctor's advice or ignores it. Similar issues arise in the fields of education and entertainment. For example, the output of a jazz band may not be well-defined without considering whether the audience was 1,000 people, 10 people, or no one at all. The output might be considered to depend on whether or not the performance was recorded for the pleasure of a future audience. Further, the experience of being in the audience may depend on whether other members of the audience are enthusiastic or indifferent to the performance. Yet, in all these possible circumstances, the music actually performed might be identical. At this time, there is no widely-accepted model for incorporating the role of the consumer into the measurement of services outputs.

These problems will not prevent statisticians from improving service-sector data, because improvements are possible without providing perfect solutions to these problems. These considerations do lead us to understand that, for many services, the improvements may not come easily.

Recent progress, additional efforts

Throughout the U.S. Federal Government's statistical agencies, there is a recognized need to improve data on the service sector. Well-informed researchers—both within and outside of Federal statistical agencies—have noted the need for improvements in the service sector data produced by BLS programs measuring consumer and producer prices. Managers of the statistical agencies and outside researchers also have recognized the need for improvements in the services components of the national accounts data, including the GDP series, prepared by the BEA, and in the annual surveys and censuses of the Census Bureau that collect service industry data. In addition, many recognize that national accounts and productivity series need improvements in ways not directly related to service sector data.

The Federal Government has made efforts to meet these needs for improved data. Despite the conceptual problems discussed above, progress has been made in improving the data available for productivity measurement. Improvements have also been made in the methods used to calculate productivity measures. Some of the leading efforts:

- BLS has made important improvements in the Producer Price Index (PPI). Many new producer price indexes for the service-producing industries have been introduced. Since 1988, BLS has introduced 47 new PPIs for such industries. The Bureau of Economic Analysis is now making use of many of these new PPI components to compute its real GDP series.
- The Bureau of the Census has greatly expanded its coverage of service industries. Since 1985, it has introduced

new censuses or surveys for transportation, communications and utilities, and finance, insurance and real estate. For example, the Annual Survey of Communication Services was introduced in 1990. A census of transportation, communications, and utilities was introduced in 1987. A census of finance, insurance, and real estate began in 1992.

- As noted earlier, BLS has substantially improved the CPI in recent years and has plans for further improvements. For example, as noted above, BLS has recently introduced geometric means indexes as the index methodology for constructing the subindexes for most components of the CPI. In July 1998, using historical data made available by the BLS, the Bureau of Economic Analysis published revised GDP data that incorporated the geometric means indexes for components of consumption expenditures.
- In the years since 1985, the BEA has improved its data on annual capital investment by asset category. The BLS uses these data to prepare its series on capital services inputs, a critical component of the BLS multifactor productivity computations. These data are now available in much greater industry and asset-type detail than they were prior to 1985. For example, in 1997, BEA introduced data on investment in eight distinct categories of computers and office equipment, in place of the single aggregate category used previously. In 1997, BEA also improved its procedures for estimating economic depreciation of capital assets.
- The Bureau of Economic Analysis and the Bureau of Labor Statistics have introduced improved methods for measuring the prices of computers and other types of advanced information technology equipment. In 1985, the BEA introduced an innovative means of adjusting computer prices for quality change. Since that date, BEA and BLS have introduced improved methods for calculating quality-adjusted price indexes for a variety of other types of information technology equipment, including semiconductors. These new methods have permitted the development of substantially more accurate measures of output and productivity.

The BLS measures of productivity trends have benefited from the new censuses and surveys and the improved data on producer prices, consumer prices, and capital investment.²⁸ When the improvements have originated outside the BEA, the BEA has often incorporated these improvements directly into the national accounts data, to the subsequent benefit of the BLS productivity series.

Despite the improvements that have been made in the availability and quality of service sector data, additional improvements are undoubtedly needed. The findings of Gullickson

and Harper, along with other available information on measurement methods in the service sector, indicate that special attention should be given to output data relating to insurance, banking, construction, health services, and utilities.

For the insurance industry, it would be useful if researchers would give additional thought to the appropriate variable to measure. The output concept used for the fire and casualty insurance industry recently has been examined.²⁹ Among the questions discussed is whether the output concept should be based on total premiums received by the industry or on premiums less claims, the basis of the current national accounts data. Similar discussion of other insurance industries would be useful. The Bureau of Labor Statistics is now publishing a producer price index for fire, marine, and casualty insurance, starting with data for July 1998, and a price index for life insurance, starting with data for January 1999. The development of these new indexes has benefited from the recent work on appropriate output concepts for insurance.

For the banking industry, additional research on the appropriate concept of banking output is needed. It is not realistic to expect that lasting consensus on such a concept will be reached in the near future. There are, however, at least two short-term alternatives that could be examined. One alternative would be to adopt the output trend from the existing BLS banking industry labor productivity series. This output series is calculated from selected data on the numbers of transactions on the asset side and the liabilities side of banks' balance sheets, and so it reflects changes in banking activity. It is compiled independently of input data. This output series, when related to labor input information, shows a strong secular increase in labor productivity. A second alternative would be to adopt an output trend based on all inputs rather than on labor input alone. This second procedure would produce a historical series with a zero trend in multifactor productivity in place of a zero trend in labor productivity. The adoption of either of these procedures would result in a slightly higher historical growth rate of multifactor productivity in the private business sector. The BEA is planning to incorporate a new measure of banking activity into the national accounts in October 1999, as part of its next comprehensive revision of these accounts.

For improved data on construction output, the greatest need is for better data on real deliveries of construction projects to their final users. To develop these data, there is a need for additional price indexes that are adjusted for quality change, especially for nonresidential structures. BLS has plans to design such indexes, with a view to their use in developing trends in real construction output. The BLS is collaborating with the BEA and the Census Bureau in this effort. Data developed in this way should replace the current practice of extrapolating nonresidential construction output by the use of cost indexes, a technique that is likely to yield a zero

productivity trend. It would also be helpful to review the available hedonic price indexes that are presently used to deflate current-dollar data on single-family and multiple-family residential construction. In 1997, the Census Bureau sponsored a conference to examine problems with construction statistics and to assess the means of addressing these problems. Representatives of BEA and BLS attended this conference. Some participants expressed concerns about the accuracy of current-dollar construction expenditures, especially for improvements.

For health services, useful new ideas have been developed through the efforts of a research project funded jointly by BLS and BEA, and carried out by the National Bureau of Economic Research (NBER). Results from this work were reported at a June 1998 meeting of the Conference on Research in Income and Wealth. At this meeting, several suggestions were set forth for improved data collection and for better quality adjustments to price indexes in health services. The Bureau of Labor Statistics will follow up on these ideas by working with NBER economists to develop an experimental method for handling substitutions between medical treatments in the producer price index. Weights for competing medical treatments will be adjusted annually to ensure that price indexes are representative of current market shares. The indexes will capture the effects of substitutions that consumers make between competing medical treatments.

The most difficult problem in the improvement of service sector data will be the development of better output concepts. It would be especially fruitful if these efforts were undertaken jointly by the Federal statistical agencies and groups of academic economists and statisticians. Examples of such joint efforts are provided by two activities currently under way. One of these efforts, mentioned briefly above, is a project to improve measures of output in health services, financed by the BLS and the BEA and implemented by the National Bureau of Economic Research. The second effort is an initiative of The Brookings Institution. Brookings is assembling groups of experts from government and academia to examine the data available for measurement of output and productivity in services and to improve the underlying output concepts.

Collaboration between BLS and BEA has been strengthened in recent years. Senior managers in the two agencies meet several times a year to develop further the complementarity of their statistical programs. In addition, regular meetings are held between the managers of the Census Bureau and the BEA. Also, the Census Bureau's budget includes a special initiative to support the BEA's efforts to improve the national accounts data. Finally, the BLS and Census Bureau have a record of close collaboration. For example, their joint efforts resulted in the major redesign of the Current Population Survey implemented in January 1994.

Despite the recent improvements in productivity mea-

surement methods, and despite new efforts like the NBER and Brookings projects, difficult measurement and conceptual problems affecting service sector data remain. These problems reflect, in large part, fundamental theoretical and conceptual diffi-

culties not yet resolved, despite the efforts of some of the best minds in economic theory. It is realistic, however, to expect that steady progress can be made toward more precise measures of service sector output and productivity. □

Footnotes

ACKNOWLEDGMENT: This paper has benefited significantly from comments and suggestions provided by Katharine G. Abraham, Lucy Eldridge, Michael Harper, Kent Kunze, Steven Landefeld, Robert Parker, Mark Sherwood, and others at the Bureau of Labor Statistics and the Bureau of Economic Analysis. However, the author alone is responsible for the content of this article.

¹ The evolution of this program is described in Edwin R. Dean, Michael J. Harper, and Mark K. Sherwood, "Productivity Measurement with Changing Weight Indices of Output and Inputs," in *Industry Productivity: International Comparison and Measurement Issues* (Paris, the Organization for Economic Co-Operation and Development, 1996), pp. 183–215; and Edwin R. Dean and Michael J. Harper, "The BLS Productivity Measurement Program," paper presented at the Conference on New Directions in Productivity Analysis of the Conference on Research on Income and Wealth, Washington, D.C., March 1998.

² The "goods" and "services" sectors discussed in this article are based on an industry classification concept. GDP data are also disaggregated by type of product, which yields data for goods, services, and structures products. For data on GDP disaggregated by product, see Tables 1.3 and 1.4 in the *Survey of Current Business*, a publication of the Bureau of Economic Analysis. In these tables, the value of goods includes the producers' value as well as the transportation costs and trade margins.

³ The output of the business sector of the economy is equal to gross domestic product (GDP) less the outputs of general government, non-profit institutions, paid employees of private households, and the rental value of owner-occupied dwellings. Business sector output includes the output of government enterprises. In 1997, business sector output accounted for 77 percent of the value of GDP.

⁴ For the private business sector, inputs are confined to labor and capital. For manufacturing, the BLS takes account of inputs of labor, capital, and three categories of inputs purchased from other businesses, energy, non-energy materials, and services. The output of the private business sector is equal to the output of the business sector less the output of government enterprises, a small component of GDP.

⁵ The slowdown in measured productivity could be due either to an overestimation of productivity growth prior to 1973 or to an underestimation of growth after 1973, or to a combination of the two. In fact, the former possibility should not be ignored, because the measurement of both outputs and inputs has almost certainly improved in recent decades. However, this possibility is not explored here, because it is not essential in view of the main purposes of this article.

⁶ For more detailed information on the data used by BEA in preparing the national accounts, see "Updated Summary NIPA Methodologies," *Survey of Current Business*, Vol. 78, No. 9, September 1998

⁷ For a discussion of adjustments made by BEA that reduce the impact of these problems, see the article by Lucy Eldridge in this issue of *Monthly Labor Review*, pp.35–46.

⁸ The BLS data on manufacturing output are prepared using a "sectoral output" concept. Sectoral output, which is gross output less intrasectoral transactions, is not equivalent to the value added concept used by BEA in constructing GDP. For detailed information on the sectoral output concept, see William Gullickson, "Measurement of productivity growth in

U.S. manufacturing," *Monthly Labor Review*, July 1995, pp. 13–28. For a discussion of the problems involved in attributing multifactor productivity growth in an aggregate sector to the industries of which this sector is composed, see Gullickson and Harper in this issue of *Monthly Labor Review*, pp. 47–66.

⁹ For many years, the BLS used the BEA gross product originating data for its annual manufacturing output series. The BLS also makes use of the Federal Reserve Board's industrial production indexes. These indexes are used to develop data on quarterly trends in manufacturing output and to estimate annual output trends for the most recent years.

¹⁰ The productivity trend for nonmanufacturing cannot be correctly calculated by simply subtracting the productivity growth rate for manufacturing from the rate for the whole business sector. See the discussion of this point in the Gullickson and Harper article in this issue of *Monthly Labor Review*, pp 47–66.

¹¹ The rental value of owner-occupied dwellings is the fourth GDP component that is excluded from the business sector. This component is omitted from the business sector productivity series largely because there is little information on the labor input allocated to this activity. Discussion of this issue and related issues is presented in Jerome A. Mark, "Problems encountered in measuring single-factor and multifactor productivity," *Monthly Labor Review*, December 1986, pp. 3–11, and in the Eldridge article in this issue of *Monthly Labor Review*, pp. 35–46.

¹² See, for example, Zvi Griliches, "Productivity, R&D, and the Data Constraint," *American Economic Review* 84, March 1994, pp. 1–23; and L. Slifman and C. Corrado, "Decomposition of Productivity and Costs," Occasional Staff Studies, OSS-1 (Washington, DC, Federal Reserve Board, November 1996).

¹³ An additional cautionary comment is needed. The real GDP series is prepared mainly by adjusting current-dollar data on final expenditures (for example, expenditures data on categories of consumption and investment) for the trends in the selling prices of the goods and services sold. The Gullickson-Harper study, on the other hand, examines the available information on the outputs and inputs of industries. A measurement problem for an industry identified in the Gullickson-Harper study need not imply a measurement problem for the corresponding GDP expenditure component. In many cases, however, the national accounts and the industry data in the Gullickson-Harper study are prepared from the same sources and use the same methods.

¹⁴ The Bureau of Labor Statistics normally does not publish productivity series for most two-digit industries outside manufacturing, or for nonmanufacturing as a whole. For a number of industries, this would require the use of data that are not considered reliable.

¹⁵ For information on the Current Population Survey and the Hours-at-Work Survey, see *BLS Handbook of Methods*, Bulletin 2490 (Bureau of Labor Statistics, April 1997), chapters 1 and 10.

¹⁶ Gullickson and Harper, in this issue of *Monthly Labor Review*, examine information on inputs in 27 construction and service-producing industries. In 1992, the share of capital in total factor costs exceeded 25 percent in only 4 of these 27 industries. Among the five industries mentioned below as contributing to low multifactor productivity growth in an important way, utilities had a capital share of 39 percent; banks a share of slightly less than 25 percent, and construc-

tion, insurance carriers, and health services all had capital shares of 12 percent or less.

¹⁷ Over the period 1977–92, the period emphasized in the Gullickson-Harper study, the average annual growth rate of the BLS labor composition index in the private business sector was 0.5 percent per year. This calculation is based on data underlying the indexes published in the most recent BLS news release on major sector multifactor productivity trends. (See News Release USDL 99–36, “Multifactor Productivity Trends, 1997,” Feb. 11, 1999. The methodology for these calculations is presented in *Labor Composition and U.S. Productivity Growth, 1948–90*, Bulletin 2426 (Bureau of Labor Statistics, December 1993).

¹⁸ See Bureau of Economic Analysis, U.S. Department of Commerce, “BEA’s Mid-Decade Strategic Plan: A Progress Report,” *Survey of Current Business*, Vol. 76, No. 6, June 1996.

¹⁹ Gullickson and Harper also undertake an alternative arbitrary exercise: they adjust all negative multifactor productivity trends to a positive 1.0 percent per year. This yields an increase in private business multifactor productivity of about 0.9 percent.

²⁰ The only health services included in the Gullickson and Harper study are the “for-profit” services; health services provided by governments and nonprofit institutions are excluded from the BLS business sector and also from the Gullickson-Harper study.

²¹ “Toward a More Accurate Measure of the Cost of Living.” Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index. S. Prt. 104–72, Washington, D.C., December 1996.

²² See Commissioner of Labor Statistics Katharine G. Abraham, “Tes-

timony before the Subcommittee on Human Resources, House of Representatives Committee on Government Reform and Oversight,” April 30, 1997; and Brent R. Moulton and Karin E. Moses, “Addressing the Quality Change Issue in the Consumer Price Index.” *Brookings Papers on Economic Activity*, 1997:1.

²³ It is noted above that an alternative exercise in the Gullickson and Harper study results in an increase in multifactor productivity of 0.9 percent.

²⁴ Griliches, “Productivity, R&D, and the Data Constraints.”

²⁵ Mark K. Sherwood, “Difficulties in the measurement of service output,” *Monthly Labor Review*, March 1994, pp. 11–19.

²⁶ Frank C. Wykoff’s comment on Allen N. Berger and David B. Humphrey, “Measurement and Efficiency in Banking” in Zvi Griliches, ed., *Output Measurement in the Service Sectors* (Chicago, University of Chicago Press, 1992), pp. 279–287.

²⁷ Mark K. Sherwood, “Difficulties in the measurement of service output.”

²⁸ Brief reports on these data improvements may be found in the statistical agencies’ budget submissions and in occasional published reports by the agencies. For example, see “BEA’s Mid-Decade Strategic Plan.”

²⁹ W. Erwin Diewert, “Functional Form Problems in Modeling Insurance and Gambling,” discussion paper no. 94–31, Department of Economics, University of British Columbia, September 1994; and Mark K. Sherwood, “Output of the Property and Casualty Insurance Industry,” *Canadian Journal of Economics*, forthcoming.