Difficulties in the measurement of service outputs

Accounting for the consumer's role in the production of a service and unraveling the services within a complex bundle are two of the difficult issues that must be resolved before the output and productivity of many industries can be measured

Mark K. Sherwood

Considerable data on revenues of service industries are available to researchers interested in productivity analysis. Often, however, the data do not provide a measure of output that distinguishes changes in price over time from changes in real output. This article examines and classifies some of the generic problems that have been identified in studies that address the measurement of the real output of service industries.

In general, measuring a service industry's output involves first identifying the unit of output and then dealing with the issue of quality change. Conceptual problems can arise at either or both of these stages. Accordingly, the article considers problems falling into two broad categories corresponding to the two stages. The failure to solve these problems completely has resulted in the absence of adequate price indexes for use in the construction of output measures for many service industries. Also, the article summarizes problems encountered in measuring output when innovations occur in an industry and, finally, presents two general approaches that have been followed by researchers to measure services.

The discussion of quality change is limited to conceptual problems that make the identification of such change difficult in regard to many services. The article does not discuss the formidable problems that arise in measuring public services and that can arise even in the classification of industries and firms as producers of goods or services.

The customary approach

The usual way to measure the real output of an industry or sector when employing typical sources of data is to deflate a nominal measure of output for the industry with a price index for the industry's product. Common nominal measures of output are *value of shipments* for goods (adjusted for changes in inventories of work in progress and finished goods) and *revenues, sales,* or *receipts* for services. When constructing a price index for deflating nominal output, *it is necessary to specify first exactly what is being purchased.* Generally, this is the basic transaction unit of the product.

Once a unit of output is specified, the next step is to enumerate all the characteristics that determine its price. The variation that occurs in a given characteristic over time or among suppliers amounts to a change in quality of the product.

If the price of a product rises due to an improvement in one of the characteristics of the product, we attribute the increase to a change in the product's quality (leading to output growth), and not to an inflationary price change. Several means are employed to adjust for variations in quality in the observations.

Although not a completely clean distinction, it is useful to separate the discussion into two parts,

Mark K. Sherwood is an economist in the Office of Productivity and Technology, Bureau of Labor Statistics. The views expressed herein are the author's and do not necessarily reflect those of the Bureau of Labor Statistics or the Department of Labor. This article is based on a paper presented at the eighth annual meeting of the Voorburg Group on Service Statistics in Oslo, Norway, Sept. 27–Oct. 1, 1993.

one addressing problems that arise when defining the basic service unit and the other treating problems that arise when controlling for changes in quality in either time series or interfirm comparisons. A later section compares two general approaches to measuring services that attempt to capture both the basic output of the industry and any concomitant changes in quality. One of these methods (the "transactions approach") attempts to measure the unit of transaction of the service, while the other (the "outcomes approach") attempts to measure the outcomes of the service.

Defining the unit of service

A very general definition of a service is useful before discussing the problem of defining the *quantity* of the services performed in an industry. Based upon a concept set forth by T. P. Hill, a service is a change in the condition of a person, or of a good belonging to some economic entity, brought about as the result of the activity of some other economic entity, with the approval of the first person or economic entity.[1] For the sake of simplicity, we shall refer to the good or person undergoing the change as the *consumer* and the other economic entity causing the change as the *supplier*.

Even with this general definition, the basic unit for measuring the quantity of the services performed in an industry can be difficult to specify. Because services result in changes, in many cases the basic unit is intangible. When a service effects a change in the condition of a consumer, it may mean simply that the consumer is now allowed to do something he or she could not do before purchasing the service—for example, attend a concert.

Many services are not this abstract, however. For instance, house painting can be directly observed and the unit specified, in, say, number of coats of paint applied, square feet covered, and grade of paint used. Similarly, the output of a railroad, although intangible, can be defined as the aggregate of weighted freight ton miles and total passenger miles.[2] In both of these cases, alternative definitions would of course be possible, to match the detail of available data.

In measuring particular services, researchers have encountered several types of problems when they have tried to define the basic unit of output. Some of the major difficulties they have met with are (1) enumerating the elements of a complex bundle of services; (2) choosing among alternative representations of an industry's output; (3) accounting for the consumer's role (which manifests itself in several ways) in the generation of a service output; and (4) specifying clearly whether what is being worked with is a value-added model or a gross output model.

Enumerating the elements of a complex bundle. For some services, the transaction unit actually includes several services bundled together in a complex way. The services in the bundle may be jointly produced or interdependent. Each has its own set of characteristics, the implicit prices of which determine the implicit price of each service.

Sometimes, the problem is so complicated that it is difficult to enumerate all of the services in a bundle or keep each separate so that we can eventually match price changes for the entire bundle with changes occurring in the individual services. An example of this occurs in the retail trade industry. When a consumer goes to the supermarket to shop, he or she is buying a shopping basket full of food. Through the markup the store adds to the food prices, the consumer is also buying a complex bundle of services from the store itself. We observe only a single price, the markup, for all of the services.

Walter Y. Oi enumerates some of the elements the store may include as part of the bundle of services: consummating transactions; assembling, displaying, and providing information on goods; making the goods available at times and places convenient to customers; supplying additional services, such as delivery and credit; and packaging and processing goods into more suitable forms (for example, cold cuts at the delicatessen counter of a supermarket).[3]

The banking sector offers another example of bundled services. When a customer has a checking account, the bank provides safekeeping and accounting services, as well as facilitating payment of bills. Further, many bank services are interdependent, so that the provision of one service may include the provision of others that cannot be transacted separately.[4]

Alternative representations of output. For some industries, there is no obviously correct way (nor are there adequate data) to describe all of their outputs or functions in terms of transaction units. Banking is an example that has received much attention. Rather than describing the transaction units of this industry, researchers have made different assumptions as to the way underlying services are attached to the activities or products of a bank. The activities or products are then the output measure. In the literature, there is no agreement as to which of the different representations is best.

Banks transform liabilities (deposits) into assets (loans). They also facilitate the economy's payment system as a way to attract deposits. Specifying transaction units to describe all of the services the bank provides while performing these functions is difficult.

Outside of special accounting treatments of banking in the national income and product ac-

counts of some countries, there have been two different points of view leading to different approaches to measuring the industry's output: the production approach and the financial intermediation approach.

In the production approach, banks are viewed as firms that use capital and labor to carry out various activities associated with loans and deposit accounts. Output is then measured as the number of accounts or transactions carried out for each product. For example, the measure of bank output underlying the BLS industry productivity measure includes deposits, loans, and trust department activities as major service functions. The demand deposit function is measured in terms of the number of checks transacted and the number of electronic fund transfers, and the time and savings account functions are measured in terms of the number of transactions performed.

In the intermediation approach, the focus is on the bank as a financial firm whose goal is to choose an optimal portfolio. Of interest is the optimal dollar amount for each product the bank provides. These nominal balances are related to services because the financial services are assumed to attach to each dollar in the account. According to this approach, the values of loans and investments are the outputs.

For either of these approaches, it is difficult to classify and incorporate demand deposits as inputs or outputs in a given model. It is possible to view them as outputs because customers acquire demand deposits for safekeeping, accounting, and bill-paying purposes. On the other hand, viewed as liabilities and a source of funds for the bank, deposits should be classified as inputs. Researchers have adopted various criteria for making this determination.[5]

Triplett describes why the various criteria have not satisfactorily resolved the dilemma. In brief, the value of the services should be included in *both* a bank's output *and* a bank's input costs. The bank pays (or barters with) depositors for these inputs (funds) at least in part with outputs (services). But to include the value of deposits is problematic because some of the services are not explicitly priced. Researchers and national income accountants have had to rely on various ways to impute the prices of and control for variations in the quality of the services.[6]

Accounting for consumer involvement. The essence of this problem is that consumers supply an input—that is, they must be somehow involved—when a service takes place. In an extreme but illustrative example within Hill's framework, if a band plays to an empty stadium, there is no output because there are no consumers.[7] Further, there is no way to put this live performance into inventory and consider it an output. The situation is unlike that of a goods-producing industry, such as the auto industry, in which a car can be produced, put into inventory, and counted as output, even though a consumer does not buy it.

The customer involvement as an input forces us to confront some difficult issues when we measure certain services. First, we must distinguish between services and outcomes so that we specify clearly what is being transacted. Second, the measures need to account for the fact that output can be dependent upon the number of consumers served. Finally, we must be cognizant of the provisions that the suppliers make to accommodate situations in which it is uncertain if and when the consumer will purchase the service.

To specify a service output correctly, we need to separate the change in the consumer brought about by the supplier from other changes that might occur when the service is provided. As Hill cautions, the service cannot extend to benefits or outcomes that the provider cannot supply.[8] We might choose to measure output in terms of outcomes rather than transactions, but even if we pursue this route, in order to assign responsibility correctly to a given supplier (or sector), it will still be necessary first to enumerate carefully only those outcomes due to the supplier's efforts.

Hill presents an example from the domain of medical care. In his view, when a doctor consults with a patient in the doctor's office and recommends a course of action for an ailment, giving advice is the service. The patient's condition has been changed because he or she now possesses the knowledge the doctor has transferred. The patient's physical condition may indeed improve through following the doctor's advice, but this outcome is not the service. Even if the patient's condition worsened, the service was performed. The easiest way to view this situation is to note that the patient will be billed for an office visit, regardless of the outcome. The patient's expectations regarding the probability of success from following the recommended course of action will affect the price paid for the advice and will be discussed subsequently as a component of quality.[9]

In Hill's example, he assumes that the information has passed from the doctor to the patient. Another illustration of the complications caused by consumer participation is a situation in which a teacher is teaching a class of poor students. The teacher is trying to impart some of his or her knowledge directly to the pupils through instruction. But because of the students' inability to receive this knowledge, there is no change in their condition. Hill argues that the teacher's effort is wasted and cannot be counted as productive. Hence, there is no output.[10]

However, it is reasonable to argue that if someone is paying the teacher to be in the classroom, the teacher must be providing some service. This particular ambiguity can be resolved by specifying what the teacher agrees to provide and what the students or their parents expect to receive. The teacher agrees to appear in front of the class and to deliver lectures. Even if the students learn nothing, the teacher is enabling them to meet the legal requirement that they must attend school until a certain age. Viewed this way, there actually is an output.

Another complication due to consumer participation comes from the case of collective services. A collective service is provided when changes occur in several persons or to the goods of several economic entities as the result of the action of a single economic entity with the permission of all concerned.[11] Here, the output can be viewed as dependent upon the number of consumers.

For example, a bus company provides a collective service when several passengers ride on one of its buses. Or, in a more complicated case, the radio and television broadcasting industry provides a service to other industries that advertise on the airwaves. The broadcast industry's output is dependent upon the size of its audience. Further, the output can be dependent upon the *share* of the audience advertisers wish to reach.[12]

Collective services involve notions related to output and productivity. Hill discusses two of these. First, as the number of recipients goes up, it is necessary to account for a congestion factor.[13] For example, in education, an increase in the student-teacher ratio may lead to a reduction in the quality of the service. In an illustration of accounting for the congestion factor, Swati Mukerjee and Ann Dryden Witte explore adjusting the output of the day-care industry for a change in quality by measuring the number of staff-class hours divided by a weighted index of the number of children.[14] The weights account for the varying needs of children of different ages. (The adjustment also illustrates how the distinction between issues of defining output and issues of accounting for quality can become fuzzy.)

The second notion raised by Hill regarding collective services is that the number of recipients may decline for reasons beyond the control of the provider of the service. For example, the number of students attending school may fall due to changing demographics. The decrease will lead to reduced output and, if the number of teachers is unchanged, to lower productivity. Hill says that this is the correct result and that we should not confuse the efficiency or skill of the provider with productivity.[15]

Another example of the difficulties faced in accounting for consumer participation is when both of the previous problems occur—that is, when there are multiple consumers *and* we must separate the service from the outcome. Hill discusses measuring the output of spectator events and suggests that perhaps it would take a psychologist to measure the change in condition of the spectators.[16]

In lieu of this, however, we could examine the conditions of the "contract" between the ticket holder and the team. When a fan buys a ticket to a ball game, the team guarantees neither a win nor a nice feeling for the fans, but rather, merely that it will show up and play the game. The fans know that this is all that is guaranteed, but they buy the tickets nonetheless because they enjoy watching the game regardless of its outcome. One could let the number of tickets sold represent the output, because it represents the number of people allowed to view the game.[17]

Further, fans are likely to pay more to see a good team rather than a poor team and to see the same team if it has improved over the previous year. If a fan believes that a favorite team has a better chance of winning the game and, consequently, that he or she will experience a better psychological outcome, it is likely that the fan will be more willing to pay a higher price for a ticket. If the team improves, but ticket prices remain the same as in the previous year, more fans are apt to pay that price. From a measurement perspective, if the price of tickets increases or attendance rises, we have to make the difficult determination of whether it reflects an improvement in the team's performance. (The psychological transformation of the fans can also be viewed as part of the quality dimension, which will be discussed in the next section.)

Finally, because of consumer participation, we must account for any provisions made to accommodate the uncertain occurrence of consumer demand. This issue can be illustrated by again referencing Oi's study of retail trade. In general, to buy goods in a retail store, a customer must allocate time and resources in order to determine the right goods to buy and at what prices to buy them and in order to get the products home. The customer receives greater service if the time for the shopping transaction is reduced. For example, the store will choose its location(s) in such a way as to allow customers to reduce their travel time. This particular function performed by the store determines a part of its output, namely, convenience, which is difficult to measure.

Moreover, if and when the consumer will arrive to shop is not certain. To address this situation, a store will tolerate idle clerks so that when a customer arrives, he or she may be served without an unacceptable delay. Thus, we need to account for this service somehow in the store's output measure.[18]

Martin S. Feldstein has pointed out the need to account for uncertain demand in the context of medical care.[19] One way to accommodate the uncertain occurrence of medical problems is to design the medical sector with excess capacity. The more excess capacity, the less a patient will have to wait to have a problem treated. If this quality component is ignored in measuring and analyzing the service, greater capacity will involve greater costs without greater output, and the situation will be mistakenly viewed as one of low productivity. Feldstein argues that ignoring the excess capacity would be as wrong as measuring the fire department's productivity by the number of fires extinguished per firefighter.

Gross output versus value-added output. Two concepts of output are commonly used when an industry's output is measured. The first is *gross output*, which equals the actual goods or services produced by the industry. The second is *value-added output*, which equals gross output minus the industry's purchases of materials and services from other industries.

Because of the intangible nature of many services, it is not obvious how these concepts apply in some studies. This can lead to arguments in the literature. Two articles on the insurance industry illustrate the point.

In the more recent of the two, Andreas Hornstein and Edward C. Prescott[20] begin by defining the output of the property and casualty insurance industry in terms of the gross revenues of the industry. Thus, output is equal to the number of contracts sold, valued at base-period prices. Claims are represented by the replacement durable goods provided by the industry and are treated as intermediate inputs into the industry. We can envision the insurance industry buying replacement durable goods and passing them through to the claimants. Hornstein and Prescott then calculate value-added output by subtracting claims from gross output.

Another approach is developed by Ron Hirshhorn and Randall Geehan[21] for the life insurance industry. These researchers employ data on what are essentially administrative costs associated with the different products (policies) sold by the industry. They weight together the various policies sold in different periods with the base-year unit costs of administering them. Even though they do not label it as such, applying the base-year unit administrative costs (including the cost of purchases from other industries, but not that of claims) to weight the number of policies of a given type yields a measure of real administrative costs.

Michael Denny has criticized Hirshhorn and Geehan's measure by arguing that consumers buy insurance to avoid risks and that an output measure should reflect the quantity of risk shifted to the insurance company.[22] Denny suggests the use of base-year premiums to evaluate the services provided. This is a gross output notion because output based upon premiums includes administrative costs plus claims.

Hirshhorn and Geehan counter that Denny's suggestion will yield only a measure of funds processed by the industry.[23] The life insurance industry offers protection because it has created the facilities for pooling risks; the activities associated with maintaining its capacity to pool risks is the service provided by the industry. Hirshhorn and Geehan argue implicitly that the risk is shifted from the consumer to the pool, and not to the company, and that the pool (or claims) is an intermediate input. Their measure is effectively the value added of the industry.

Thus, viewed from within Hornstein and Prescott's framework, Denny and Hirshhorn and Geehan are actually presenting gross and value-added notions of output, rather than proposing conflicting definitions of output. There are reasons for preferring one concept of output to the other, but that is not the focus of the debate between these researchers.

Variation in quality

The quality-determining characteristics of many services may be numerous, and changes or variations in a particular characteristic may be large. Thus, measuring the quality dimension properly can be as important in the determination of real output as measuring the number of basic service units there are.

Accounting for variation in quality can be difficult. Several techniques have been employed to do the job.[24] For example, a statistical technique is available that allows us to account for variations over time or among suppliers in the characteristics of some products. This "hedonic" approach, as the method is commonly called, has proved very useful to researchers. Once we specify the transaction unit, the technique is used to derive the relationship between a product's price and its characteristics. This information can then be used to adjust a price index so as to remove the effects of variation in quality.[25]

To make quality adjustments, we would need to specify the differences in quality among products. Such differences can sometimes be easily observed. For instance, it is possible to tell that one motel has a swimming pool while another does not. This is a problem similar to one frequently encountered for goods and is not of particular interest here.

Sources of difficulty in identifying and control-

ling for variations in the quality of services include dealing with a complex bundle of services. Even for unbundled products, the units can, to a large degree, fail to be standardized with respect to quality. Further complications are due to situations wherein the supplier knows more about the probable outcome of a service than does the consumer. Finally, the consumer's role in the provision of a service must be addressed with respect to the quality of the service.

The complex bundling of services is an issue not only when trying to define the basic service unit, but also when trying to control for quality. Each of the bundled services will have characteristics that can vary and, because of their intangible nature, be hard to identify. The complexity of the product results in many elements that can change over time and among suppliers, but be reflected in only a single price.

For example, when a store hires more clerks to speed up checkouts and save the customer time, a characteristic of the "consummation of sale service" has been changed. We would need to label this an increase in quality. If the markup were raised to pay for the extra staff, it would be necessary to remove this increase from the markup before calculating the price change.

Sometimes a simple basic unit, like an hour of a lawyer's time, can be specified as an output. However, such units are not standardized.[26] One reason for this is that the quality of a given output unit is at least partially a function of the probability that the consumer will achieve the outcome he or she desires after purchasing the service. For multiple possible outcomes, there would be a probability for each outcome.[27]

Typically, clients hire defense attorneys on an hourly basis. The consumer will make a decision on which lawyer's time to purchase by considering the price of and expected benefits from the available lawyers' services. If, then, we look at the hours spent by a defense lawyer as a measure of output, a highly skilled practitioner's hours will be considerably more valuable than those of a novice because the skilled lawyer is more likely to win the case for his or her client.

The difference between the two lawyers in terms of the probability of winning a case may depend more upon their relative skills than upon the difference in hours that they spend on the case. The difference in skills can be due to differences in the two lawyers' innate talents, training, and experience. It can also change over time for each of them.

Whether the transaction unit is simple or complex, what complicates the problem further is that the quality dimension can be difficult for the consumer to determine. The supplier often has better information than the consumer does on which to make the determination. This issue is referred to as the problem of asymmetric information. A simple example is the case of a doctor, who almost always knows the chance of success of a medical procedure better than the patient does.

There are two reasons why the consumer is less informed than the practitioner regarding the probable outcome. First, there is a situation labeled "moral hazard," wherein a consumer is unable to observe the actions taken by the supplier. It is always possible that the practitioner will not give his or her best effort in a particular case because the client is unable to observe that effort.

Second is another situation, called "adverse selection," wherein the client cannot observe the qualifications of the practitioner or the conditions under which he or she operates. Thus, a patient may not be able to judge the competency of his doctor, and a consumer may not be able to judge whether all the work done by her auto mechanic was necessary.

The marketplace may address the problem of asymmetric information through the introduction of such things as contingent fees or warranties. Or it may offer the practitioner's reputation as a way to overcome the asymmetry.[28]

Asymmetric information can make the measurement and analysis of services difficult. First, we have to make sure that the relevant market corrections are contained in the service. For example, if a warranty is part of a service, any change in the warranty needs to be accounted for in a measure of that service.

Second, unless we make some adjustments in our analysis, we may assign output and productivity growth to the wrong industry. Such a misassignment would occur if the market price of the service did not accurately reflect the service being purchased from the industry being studied. One instance of this is when a correction to the asymmetry problem is made by another industry, such as the malpractice insurance or the legal industry.

Suppose, for example, that a doctor takes out a malpractice insurance policy and raises the prices he or she charges to patients. Then the doctor's patients may be better off because they are better protected from the effects of asymmetric information. In that case, there is no inflationary price increase for the total package of medical treatment plus insurance coverage. However, the resulting increase in output is due to the insurance industry, and not to the medical industry.

Finally, Griliches emphasizes the role of the consumer as a source of variation in the quality of services.[29] The earlier example of the teacher teaching poor students is a case in point. Because a poor student absorbs less information than a good student, it might be argued that the teacher in

charge of the poor students has a poorer quality of output than does a teacher with good students.

Capturing innovations[30]

When the service sector is described as being difficult to measure, one frequently mentioned reason is that it is a sector in which rapidly changing technologies have occurred or been employed. The concern is that these advances have led to innovations that produce changes in quality and productivity which are difficult to capture. The use of computers by service industries is a frequent example. Martin Neil Baily and Robert J. Gordon point out that due to the greater use of computer power, retailers can monitor and control stocks better than before and can offer customers a larger variety of goods and reduce the number of occurrences when goods are out of stock.[31]

The innovations we are interested in accounting for can be classified in one of two ways. The first variety is those that increase the effectiveness of a service and that are referred to as *product innovations*. For example, an innovation may result in a higher success rate in the diagnosis of a particular disease. Or it may allow for the diagnosis of a disease with less pain and fewer complications for the patient.

The second category of innovation is labeled *process innovation*. Here, the effectiveness of a service is not altered, but it becomes possible to provide the service with fewer resources. Griliches offers the example of the new laser-based gallbladder treatment.[32] With this technique, it is presumably possible to treat gallbladder problems with the same success rate as is achieved with other procedures, but using fewer resources.

Both types of innovation can present difficulties in measuring the output and productivity of services. Product innovations lead to changes in quality and the difficulties they can present. Process innovations lead to difficulties in matching inputs with outcomes. Suppose, for example, that a process innovation results in a new medical procedure, as in Griliches' example. If the new procedure is treated as a new service, it will be related to its own inputs and the old procedure related to its inputs. In that case, none of the productivity gains due to achieving certain outcomes (for instance, "cures") with fewer resources will be captured. (Of course, if we were to accept Hill's framework, it would preclude us anyway from using "cures" as an output measure.)

In some particularly difficult cases, both product and process innovation can be present. The advances in the personal computer industry are an illustration from the goods sector. When new models are cheaper than their inferior predecessors, the phenomenon has been referred to as "costless" quality change.

Transactions versus outcomes

To this point, we have discussed some general problems that will be encountered in trying to define the service unit and account for changes in the quality of services. There are two general approaches that researchers have followed when addressing these problems. For some services, one way is more appropriate than the other.

The first approach is to measure the service transaction unit itself and make appropriate adjustments for changes in quality. The service unit would be specified as that which is contracted for, either explicitly or implicitly. The second approach is to measure the outcomes resulting from the performance of the service in question.

Consider first the direct measurement of the transaction. Returning to an earlier example, specifying the transaction unit for railroads, while not without complications, is relatively straightforward. Even though implementing the procedure may not be simple, there are many other services that are good candidates for the transaction-based approach. In fact, the Bureau of Labor Statistics uses such an approach to measure the productivity of individual service industries representing about 42 percent of workers in the service subsector of the private business sector.

However, quantifying a doctor's advice to his or her patients would be difficult. Harking back to Hill's methodology, which we could describe as a transaction approach, it would be necessary to price the bits of information transferred from doctor to patient.[33] After that, it would be necessary to adjust the price of the bits of advice for any change in the probability of a successful cure when the advice is followed. This change in quality could be due to the doctor's increased experience,[34] or it could be due to a breakthrough in medical knowledge that makes a given unit of advice more likely to lead to a "cure"—that is, it could be a product innovation.

A quality adjustment would require us to value the change so that the price of the service before and after the change can be made comparable. In many cases, however, this may be extremely difficult to do.[35]

Measuring the service transaction unit can also miss the productivity gains from process innovation. As mentioned in the previous section, this occurs because, when a process uses fewer resources, it can be due to a new procedure, which is then treated as a new service according to the transaction approach.

The alternative to measuring the transaction units or procedures for services such as medical

consultation is to measure the outcomes of the services. There are three advantages to doing so. First, with the outcomes approach, there is no need to specify all of the elements of the basic transaction unit. Instead, it is necessary to specify some quantifiable outcome, such as "cures." Second, some difficult changes in quality would be captured, namely, those reflecting the fact that a desired outcome is more probable after the change than before it. The outcome-based output measure would increase with these changes in quality. Also, product innovations that lead to a greater number of the desirable outcomes would be captured. Third, the outcomes approach directly captures the effects of process innovations. Griliches' example of the laser-based gallbladder procedure is an illustration of this: by comparing "cures" with resources, we capture productivity gains.

Unfortunately, the outcomes approach is not without its own set of problems, chief among which is that it may be difficult to avoid blurring the distinction between service transactions and outcomes. In some cases, either too much of the outcome is credited to the supplier of the service, or some output of the supplier is not captured by the outcomes measure.

An example of this ambiguity caused by the outcomes approach is in a study by Dale W. Jorgenson and Barbara M. Fraumeni.[36] These researchers measure the output of the educational system in a given year as the difference in individuals' lifetime market and nonmarket incomes associated with the year of schooling they have completed. In such a framework, future income due to education is derived from current wages for persons who have completed different levels of education.

However, as noted by Griliches,[37] this approach requires some strong assumptions. One of them is that differences in current wages for persons of a given age and sex are due primarily to differences in education and not to other factors, for instance, differences in ability. Additionally, as argued earlier in regard to a teacher teaching poor students, it is possible to have an output in the education industry without an increase in future incomes.

Further, not all quality change and product innovation may be captured by the outcomes approach. Returning to Griliches' earlier example, assuming an unchanging success rate, counting cures would not account for product innovations that resulted in less pain for patients undergoing a new gallbladder procedure. Moreover, certain quality dimensions, such as any excess capacity in the treatment system, would be missed as well.

Finally, with the outcomes approach, it may be difficult to identify that part of the outcome which is due to consumer involvement. For example, cures may increase when a medical procedure is applied to less sick patients. This would make it dubious to assign responsibility for those increased cures (which we would construe to be productivity gains) to the medical industry. □

Footnotes

ACKNOWLEDGMENT: The author thanks Edwin R. Dean, Dennis J. Fixler, Michael J. Harper, Marilyn E. Manser, and Jack E. Triplett for comments during earlier drafts of this article.

[1] T. P. Hill, "On Goods and Services," *Review of Income and Wealth*, Vol. 123, No. 4, 1977, pp. 315–38. See especially p. 318.

[2] See John Duke, Diane Litz, and Lisa Usher, "Multifactor productivity in railroad transportation," *Monthly Labor Review*, August 1992, pp. 49–58.

[3] Walter Y. Oi, "Productivity in the Distributive Trades: The Shopper and the Economies of Massed Reserves," in Zvi Griliches, ed., *Output Measurement in the Service Sectors*, NBER Studies in Income and Wealth, Vol. 56 (Chicago, University of Chicago Press, 1992), pp. 161–91. See especially p. 168.

[4] The discussion in this paragraph and the next section draws directly upon the following articles: Jack E. Triplett, "Comment," in Griliches, *Output Measurement*, pp. 287–96; R. J. Colwell and E. P. Davis, "Output and Productivity in Banking," in Ernst Berndt, Peter Englund, and Lennart Hjalmarsson, eds., *Productivity Concepts and Measurement Problems: Welfare, Quality and Productivity in the Service Industries, Scandinavian Journal of Economics*, Vol. 94, supplement, 1992, pp. 111–29; and Dennis J. Fixler, "Measuring the Financial Service Output and Prices of Commercial Banks," *Applied Economics*, Vol. 25, 1993, pp. 983–93. The articles also contain summaries of the problems researchers have encountered in attempting to measure bank output.

[5] Triplett, "Comment," gives a summary of these criteria.

[6] See, for example, Dennis J. Fixler and Kimberly D. Zieschang, "Incorporating Ancillary Measures of Process and Quality Change into a Superlative Productivity Index," *Journal of Productivity Analysis*, Vol. 2, 1992, pp. 245–67, for a study of the use of bank branches as a proxy for the quality of the financial services delivered.

[7] Hill, "On Goods and Services," pp. 332–33.

[8] *Ibid.*, p. 337.

[9] *Ibid.*, pp. 322–23.

[10] *Ibid.*, pp. 323–24.

[11] *Ibid.*, p. 332.

[12] See Griliches, "Introduction," *Output Measurement*, p. 8.

[13] Hill, "On Goods and Services," p. 333.

[14] Swati Mukerjee and Ann Dryden Witte, "Measurement of Output and Quality Adjustment in the Day-Care Industry," in Griliches, *Output Measurement*, pp. 343–69. See especially pp. 352–53.

[15] Hill, "On Goods and Services," pp. 333–34.

[16] *Ibid.*, p. 324.

[17] *Ibid.*

[18] Oi, "Productivity in the Distributive Trades," pp. 168–69.

[19] Martin S. Feldstein, discussion to M. W. Reder, "Some Problems in the Measurement of Productivity in the Medical Care Industry," in Victor Fuchs, ed., *Production and Productivity in the Service Industries*, NBER Studies in Income and Wealth, Vol. 34 (New York, Columbia University Press, 1969), pp. 139–46. See especially p. 145.

[20] Andreas Hornstein and Edward C. Prescott, "Insurance Contracts as Commodities: A Note," *Review of Economic Studies*, Vol. 58, October 1991, pp. 917–28. Note especially p. 927, where the authors indicate that the construction of output by Ron Hirshhorn and Randall Geehan, to be discussed shortly, is in the spirit of their proposed method.

[21] Ron Hirshhorn and Randall Geehan, "Measuring the Real Output of the Life Insurance Industry," *Review of Economics and Statistics*, Vol. 59, May 1977, pp. 211–19.

[22] Michael Denny, "Measuring the Real Output of the Life Insurance Industry: A Comment," *Review of Economics and Statistics*, Vol. 62, February 1980, pp. 150–52.

[23] Ron Hirshhorn and Randall Geehan, "Measuring the Real Output of the Life Insurance Industry: A Reply," *Review of Economics and Statistics*, Vol. 62, February 1980, pp. 152–54.

[24] See, for example, BLS *Handbook of Methods*, Bulletin 2414 (Bureau of Labor Statistics, 1992), pp. 191–93, for a summary of the quality adjustment problem in the context of the BLS Consumer Price Index program.

[25] See, for example, Zvi Griliches, ed., *Price Indexes and Quality Change: Studies in New Methods of Measurement* (Cambridge, MA, Harvard University Press, 1971), for a further discussion of this technique and some examples of its use.

[26] Charles R. Hulten notes that for services provided, for example, by doctors, lawyers, and barbers, output depends as much on the quality of the product as on the hours spent in producing it. (See Charles R. Hulten, "Comment: Measurement of Output and Productivity in the Service Sector," in Robert P. Inman, ed., *Managing the Service Economy: Prospects and Problems* (New York, Cambridge University Press, 1985), pp. 127–30.)

[27] These points and the discussion that follows of services and the issues of uncertainty and information draw upon Bengt Holmstrom, "The Provision of Services in a Market Economy," in Inman, *Managing the Service Economy*, pp. 183–213.

[28] Hugues Picard of the Institut National de la Statistique et des Études Économiques, in France, discusses the issue of risk and notoriety in the measurement of services in "Calculating Service Price Indexes? It Is Possible!" a paper presented at the sixth annual meeting of the Voorburg Group on Service Statistics in Helsinki, Finland, October 1991.

[29] Griliches, "Introduction," *Output Measurement*, pp. 5–7.

[30] This section and the next draw on Hill, "On Goods and Services"; Feldstein, "Measurement of Productivity"; and Griliches, "Introduction."

[31] Martin Neil Baily and Robert J. Gordon, "The Productivity Slowdown, Measurement Issues, and the Explosion of Computer Power," *Brookings Papers on Economic Activity, 2* (Washington, The Brookings Institution, 1988), pp. 347–420. See especially pp. 391–92.

[32] Griliches, "Introduction," *Output Measurement*, pp. 12–13.

[33] Hill, "On Goods and Services," p. 322.

[34] In measuring productivity, the increased experience represents a larger input on the part of the doctor in the form of increased labor input. It is important then to reflect the increased output, or there would be a decline in productivity.

[35] For a discussion and resolution of a few of the thornier issues in valuing quality change, including "costless" quality change, see Jack E. Triplett, "Concepts of Quality in Input and Output Price Measures: A Resolution of the User-Value Resource-Cost Debate," in Murray F. Foss, ed., *The U.S. National Income and Product Accounts: Select Topics*, NBER Studies in Income and Wealth, Volume 47 (Chicago, University of Chicago Press, 1983), pp. 269–311. Manuel Trajtenberg, *Economic Analysis of Product Innovation: The Case of CT Scanners* (Cambridge, MA, Harvard University Press, 1990), provides an indepth discussion of valuing changes due to product innovations.

[36] Dale W. Jorgenson and Barbara M. Fraumeni, "The Output of the Education Sector," in Griliches, *Output Measurement*, pp. 303–38.

[37] Griliches, "Introduction," *Output Measurement*, p. 16.