ECONOMIC RESEARCH REPORTS

THREE FUNDAMENTAL PRODUCTIVITY CONCEPTS: PRINCIPLES AND MEASUREMENT

by

William J. Baumol

and

Edward N. Wolff

R.R. #85-40 December 1985

C. V. STARR CENTER FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, N.Y. 10003

November 1985

Three Fundamental Productivity Concepts: Principles and Measurement

William J. Baumol*

Edward N. Wolff**

The literature on productivity devotes considerable and deserved attention to a variety of measurement problems and to distinctions such as that between labor productivity and total factor productivity. However, some of the basic definitional issues that arise implicitly in many of the discussions do not seem to have been examined to the degree they merit. In this paper, we contrast three different basic concepts of productivity, discuss the differences in their interpretation and significance, and then demonstrate empirically that use of different notions of productivity can give rise to great differences in measurements of productivity growth.

1. Three Productivity Growth Concepts

In writings on productivity growth at least three different connotations are implicitly assigned to the term. Most often, it is interpreted as a measure of the increase in productive "capacity" attributable to technical change -- the shift in the production frontier. Sometimes it is interpreted as a measure of the increase in consumer and producer welfare produced per unit of input, regardless of the source of the improvement, whether technical change, better allocation of resources given the state of technology, or some other influence. Finally, usually with some embarrassment, statistical studies sometimes deal with changes in what we will call crude productivity -- number of units of output produced per unit of input, and with little or no attempt to take account of changes in output quality. In this paper we will refer to these, respectively, as growth in productive capacity, welfare productivity and crude productivity.

We will show that the (monotonically descending) pecking order that seems at least implicitly to be assigned to these three concepts in the literature is misleading. For as will be demonstrated, each of them has its legitimate and significant use both in analysis and in application. In particular, the crude productivity measure will be shown to be extremely important in explaining the behavior over time of the relative prices of different goods and services, in budgetary planning for various public sector activities and in planning to meet future manpower requirements. It will be seen that in some cases the productive capacity measure may deviate systematically from the welfare productivity measure, and in those circumstances the former will be inappropriate for an analysis of standards of living.

Some observations will be offered about the nature of the measurement issues raised by each of these concepts. It will be shown why only in very special cases is an unambiguous scalar index of growth in productive capacity possible. On the other hand, welfare productivity, at least in principle, turns out to be easier to measure than may be expected, despite the important role of changes in the quality of the products in question which are, of course, difficult to identify and describe and sometimes all but impossible to quantify.

2. On Productive Capacity Measures

One frequently encounters the view that a pure measure of productivity growth should confine itself to the consequences of technical change and changes in the quality of the available inputs, e.g., in the skills of the labor force, leaving out of consideration changes in output per unit of input that result, e.g., from elimination of inefficiencies.[1]

Specifically, indices of growth in productive capacity seek to measure the rate of outward shift of the production frontier hypersurface in product quantity space, holding constant the quantities of inputs utilized. That is, of course, straightforward when the shift is equiproportionate, as that from frontier AA to BB in Figure 1. One need only take the ratio of RS, the increment in length of any ray, OS, between the two frontiers and the length of ray OR to the inner frontier. But matters are not so straightforward when the outward shift is uneven, as that from AA to DD. Then there simply exists no one number which can adequately measure the enhanced productive capacity. The problem is exacerbated when a new production technique expands the ability to produce one good at the expense of another so that what may be described (somewhat barbarically) as the new technique-specific frontier (EE′ in Figure 2) crosses the old frontier, AA′. One may then assume that the new production frontier for the economy will be the envelope, EHA′, of EE′ and AA′, and in these circumstances, while no part of the economy's frontier will have shifted inward, only some portion of it will have shifted outward.[2]

This indicates why those who seek to estimate some measure of growth of productive capacity usually adopt assumptions of some degree of severity about the nature of the production set and the character of technical change. Their premises are designed to offer them an expansion path for the production frontier which can be described uniquely by a scalar measure, as in the shift from AA to BB in Figure 1. Unfortunately, reality need not follow such a simple course and then <u>any</u> scalar measure of productivity growth as an index of expansion of productive capacity per unit of input becomes at best a rough

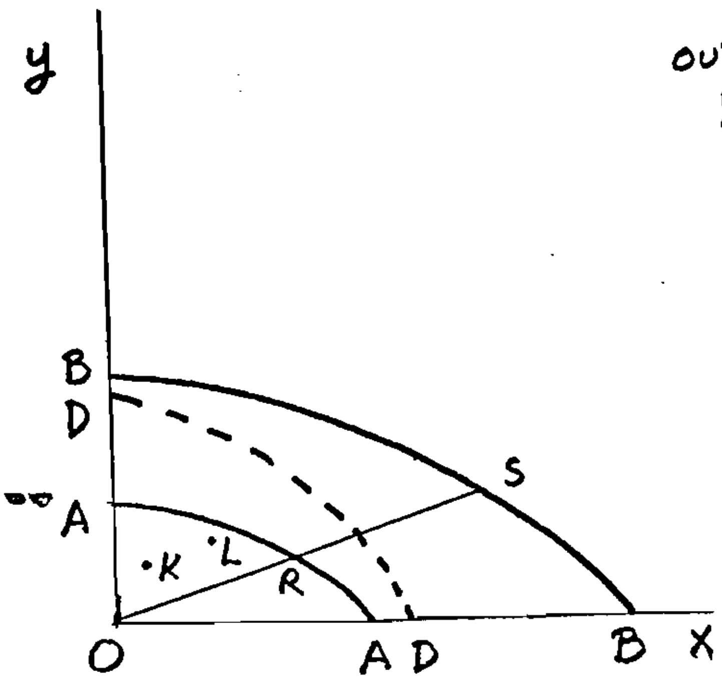

Figure 1

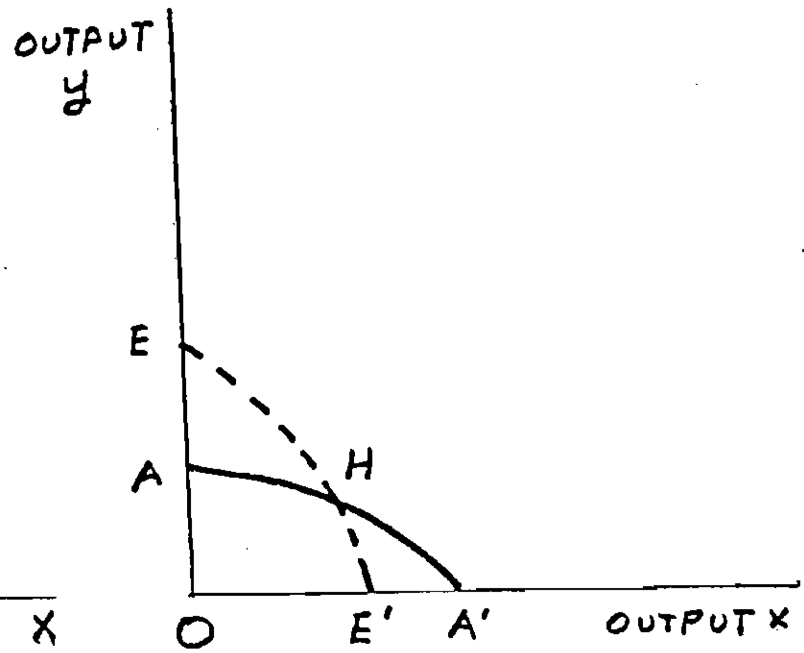

Figure 2

indicator of a development which can only be described fully by a multivariable function.

3. Imperfect Association of Productive Capacity and Welfare

However one may choose to define standard of living or level of welfare (more will be said about that presently) it is clear that there are ways other than an increase in productive capacity to enhance welfare or living standards. For one thing, allocative or X inefficiency may be reduced without any technological change. This moves the economy from a point inside the production frontier to another point closer to the frontier, but involves no shift in the frontier itself. Moreover, the product mix that constitutes output can be adapted better to consumer preferences, thereby clearly contributing to consumer welfare. Consequently, it should not be surprising that the association between growth of productive capacity and growth in welfare productivity is imperfect. In addition, disparities between the measures need not occur only haphazardly. An illustration will show how systematic differences between the two may arise.

Consider the use of a patent system to stimulate economic growth. To permit the use of a two dimensional diagram, we employ a model with a single commodity and a two period horizon, though our conclusions will be perfectly general. In period 1 the community has a choice between adoption or rejection of a patent law. The patent system will yield innovations which enhance productive capacity but it will employ as its incentive the provision of monopoly power to the innovator. The innovator will use that power to influence prices, thereby distorting the allocation of resources. Let us also assume that without patents the economy will be perfectly competitive.

Figure 3 describes the consequences of society's choice between adoption and rejection of the patent policy. The axes represent y_t and y_{t+1}, outputs in the initial and the subsequent periods. In the absence of the patent arrangement, the production frontier is AA. With a patent system, the frontier shifts outward to BA. Clearly, with the economy's resources given, patents increase productive capacity.

Welfare productivity, however, is another matter. Without patents, the economy's (competitive) equilibrium point is T, the point of tangency between production frontier, AA, and an indifference curve, II, of the social welfare function.

However, under a second period monopoly instituted via the patent system, output is distorted. Second period output, y_{t+1}, is restricted below its optimal level so that instead of the optimal point, W, the equilibrium lies to its right at a point such as R or S. If it happens to fall at R. which lies below II, the indifference curve through the patentless equilibrium, welfare productivity will clearly have fallen even though productive capacity has grown. This is no fortuitous relationship. The divergent behavior of the two productivity measures in this case may, rather, be taken as striking evidence that the patent arrangement in question is not well designed.

4. On Welfare Productivity and its Measurement

The preceding discussion may suggest, with good reason, that welfare productivity, however it may be defined, is the appropriate measure of the effectiveness with which the economy is pursuing the goals of the people who compose it. After all, we are not concerned ultimately with the size of the collection of physical objects (and services) the economy turns out, but with

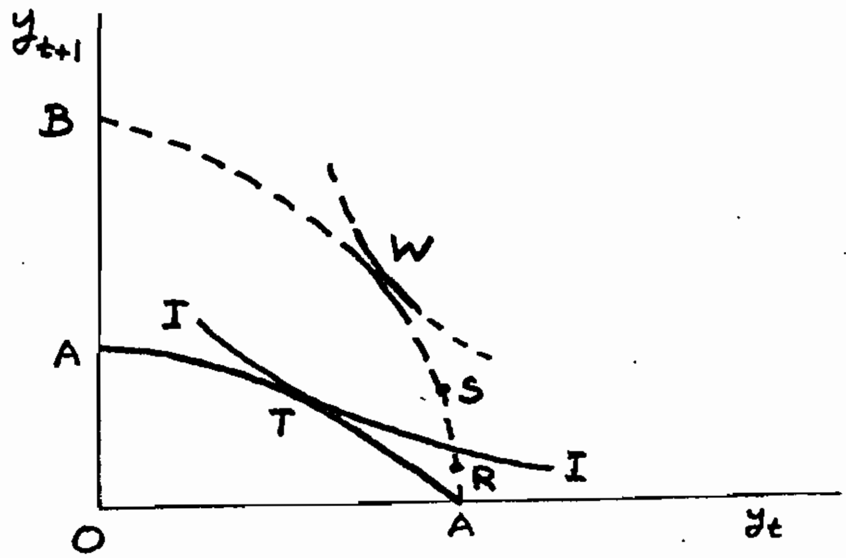

Figure 3

the degree of satisfaction they yield. That concern, for example, underlies attempts to take account of changes in product quality in the measurement of productivity growth.

Growth in welfare productivity is in some ways a more amorphous concept than that in productive capacity. Yet in one important attribute it is far more easily definable. It is more difficult to get at because welfare productivity can be increased in so many ways -- via a change in output composition or a change in product quality as well as by elimination of inefficiencies. However, it is a simpler concept because, unlike productive capacity growth, it corresponds to a move from one given point to another point in output space, not to the shift of an entire frontier. We ask by how much the move from point K to point L in Figure 1 had increased welfare, not how much of an increase in the distance of the frontier from the origin is constituted by the move from AA to DD. In mathematical terms, growth in welfare productivity is a function of the vector of outputs while that in productive capacity is a functional, i.e., a function of a nondenumerable infinity of variables: the set of points on the production frontier.

Since welfare productivity depends on product quality, this feature alone would appear to be a source of intractable difficulties. After all, it is a formidable task just to enumerate the quality changes that must be taken into account in determining what has happened, say, to the productivity of the phonograph industry or the medical profession since 1930, and all but impossible to quantify and aggregate the changes in quality of their products. This has been the bane of economists engaged in measurement of productivity in practice.

What seems at least sometimes to have been overlooked, however, is that by taking welfare productivity as a phenomenon in which we are interested, we do not thereby enmesh ourselves more deeply in the intractible complexities of measurement of quality changes. On the contrary, the welfare productivity approach, at least in principle, automatically solves those problems at a single stroke.

Here, as in so many other issues, the price mechanism once more demonstrates its amazing efficacy as a conveyor of information. In an economic sense, an improvement in product quality, after all, is not a mere concatenation of modified technological specifications -- higher quality transistors, reduced lubrication requirements and the like. Rather, it is a matter of the utility that these changes offer consumers. And on our usual premises, that contribution to utility is unambiguously measured by the behavior of a product's relative price. If a product has become better, consumers will be willing to pay more for a given quantity of the item. There is no need to go behind the change in this component to seek to determine what part of it is attributable to which element in the complex set of quality changes that the product may have undergone. One may even suspect that further inquiry into quality measurement can only be misleading and counterproductive.

Thus, price analysis is the proper way of dealing with the measurement of changes in product quality and disposes of the problem with one blow. Yet, there is a complication. We cannot for the purpose simply use the behavior of actual market prices, even in theory. True, market prices do tell us what happens to relative marginal utilities as a result of improvements in product quality. But marginal utility is not the pertinent criterion. It is easy to see that it can even be totally misleading if not interpreted with care. Thus, suppose an improvement in the quality of some product increases quantity that is sold which, in turn, elicits scale economies and a consequent reduction in price. Are we to conclude that because the product's relative price and marginal utility have fallen, its quality has deteriorated? The answer, of course, is that the fall in price has added to (and certainly not reduced) the consumers' surplus contribution of the change in product quality, which is, ultimately (together with its producers' surplus effect), the most reasonable measure of the change. That is, in measuring the welfare productivity of the particular industry, it would appear that the proper measure is the behavior of the sum of producers' and consumers' surplus per unit of input.

The need to measure changes in consumers' and producers' surpluses complicates empirical evaluation of quality changes considerably, but it certainly does not amount to total retreat to the impossible task of identification and evaluation of all of the changes in the attributes of a product that occur with the passage of time. Measurement of changes in surpluses is a far more manageable task than that. This is definitely true in principle, and to a considerable degree it is so in practice. To calculate true productivity growth in automobile production, for example, one need not undertake the (perhaps even undefinable) task of evaluating the degree of improvement in quality of a 1982 model over, say, its 1972 counterpart, including the modifications in comfort, safety, reliability, appearance, etc. In theory, for the purpose one need only estimate the relevant demand and cost relationships and use them to calculate the change in consumers' and producers' surpluses, armed with Professor Willig's [1976] assurance that the procedure is (almost perfectly) legitimate. That is, one can use the definable and conceptually observable cost and demand functions to evaluate indefinable and unmeasurable quality changes. Moreover, in practice, as a good deal of work on hedonic indices has shown, it is actually possible to obtain reasonable econometric estimates of the required magnitudes. Thus, the procedures just described transform measurement of quality changes from a mysterious and ill-defined exercise into one whose outlines, at least, are clear.

5. Crude Productivity and the Services

Anyone who has thought about the subject is apt to be uncomfortable about discussions of productivity in the services because the very concept is so elusive. Just what is the "product" of education or medical care, and how does one measure it? Is one really forced, as is often done in the services, to measure output quantities via the quantity of input used in their production so that calculated productivity remains stagnant, just because of the way it is measured?

Fortunately, for many purposes we do not have to face up to these difficulties. To explain why, we must turn to the concept of crude productivity. Crude productivity, in the calculation of which absolutely no effort is made to adjust for quality changes, is often easy to measure and is, fortunately, the correct information required for some significant types of analysis.

Crude productivity is simply a measure of number of units of observable "output" per unit of input. For example, in musical performance we can define crude productivity of labor as number of audience members (or number of concert performances) per labor hour. Similarly, in the case of higher education, the crude measure of labor productivity can be taken simply as the number of students attending colleges and universities divided by the number of faculty hours devoted to teaching, that is, the student-teacher time ratio. These are obviously easy to measure and the figures readily available. Analogous measures of multi-factor productivity in musical performance and education, taking into account other inputs besides musician or teacher time, are also easily constructed.

Why should we ever be interested in crude productivity rather than in a measure that is adjusted for quality changes? The answer is that it is the former, not the latter, that is the primary determinant of the budgets, costs and prices of the products in question. For example, ignoring other inputs, the cost of education per student is simply the wage per faculty hour multiplied by the number of faculty hours used, all divided by the number of students. But crude productivity equals number of students divided by the number of faculty hours. Therefore, the cost per student is simply the reciprocal of the index of crude productivity multiplied by the average hourly faculty salary. Immobility of labor often permits faculty salaries to lag behind wages and salaries elsewhere in the economy. But, in the long run, trends in faculty salaries seem to be determined preponderantly outside the university because of the long run mobility of labor and tend, over long periods, to be similar everywhere. This means that the only way that university administrations can affect the ratio between education cost per student and cost per unit of output in the remainder of the economy is by changing the rate of growth of crude productivity in colleges and universities. If crude productivity in education lags behind that in the rest of the economy, the cost per student must rise faster than cost per unit of output in the rest of the economy, and university fees and total budgets must follow along commensurately.

An oversimplified example will explain most clearly some important applications of productivity measurement for which it is simply wrong to take quality changes into account. In particular, we will see why this is true of certain types of budgeting decisions affected by differentials in productivity growth and for associated resource allocation decisions.

Consider an economy which produces only two outputs, call them performnce of string quartets and electronic (video) game machines. Suppose the quartets are all written for a half-hour performance, so that they always require just four musical instruments and two person hours of labor input per performance. Consequently, whatever the changes in quality of the product, crude productivity must remain absolutely fixed and immutable. Suppose also that electronic games improve in quality, in some sense, with the passage of time, and that crude total factor productivity in their manufacture grows at a rate of 7 percent per year, hence doubling every decade. Let this economy be perfectly competitive and its overall price level, P_t, stationary so that

$$P_t = P(P_{ct}, c_t, P_{gt}, g_t) = k \tag{1}$$

(where P_{ct} is the price of admission to a concert, c_t is the number of concerts performed, etc.). Suppose, finally, that both industries use similar inputs with identical input prices and that income and price elasticities of demand are such that the output proportion (i.e., the machine-concert ratio) remains absolutely constant.

Several conclusions follow. First, the price of electronic games, P_g, must fall and the price of concerts, P_c, must rise at constant percentage rates satisfying (1) and

$$(2) \qquad \left(\frac{\dot{P}_c}{P_c}\right) / \left(\frac{\dot{P}_g}{P_g}\right) = 1.07.$$

Second, with output proportions constant, the share of the economy's <u>inputs</u> devoted to concerts must increase steadily, at a rate given by the production functions and input prices for the two outputs. It should be emphasized that to calculate this change in allocation of inputs there is no need to measure the change in quality of either concert performance or electronic games.

Next, let us consider the following two applications:

(i) Schools in our imaginary economy plan how many classrooms to build for the training of musicians vis-à-vis the number they need for the training of electronic game assemblers. A moment's consideration confirms that <u>crude</u> productivity growth is the only required productivity datum. For example, if labor were the only input, and in 1970 the labor force had been divided equally between the two outputs, by 1980, since crude productivity in games doubles each decade while crude productivity in music remains constant, the fixity of output proportions requires that

$$(3) \qquad \frac{L_{c80}}{2L_{g80}} = \frac{L_{c70}}{L_{g70}} = 1 ,$$

where L_{g80} is the size of the 1980 labor force in game production, etc. Hence, we know from our calculation of crude productivity growth alone that two-thirds of the economy's 1980 labor force must be trained as musicians regardless of developments in quality.

(ii) As a second application, suppose that half the cost of each concert is obtained by public subsidy. Then budget planning by the arts support agency of the government can be carried out completely with the aid of (1), (2), and (3) which determine exactly the growth in real cost of each concert and the number of concerts on the basis of crude productivity growth data alone.

It is true, of course, that developments in quality enter the matter implicitly by determining the course of the relative demands for the two outputs, which were here subsumed in the premise that output proportions remain fixed. However, the point is that nowhere do we have to measure or even define or describe quality change to determine input training proportions or the arts subvention budget. For this we need only know crude productivity growth and changes in output proportions -- both directly observable magnitudes.

6. Some Empirical Comparisons

We shall now show, using actual data, that these concepts can yield very different measurements of both annual and average annual productivity growth. Moreover, we will show that because of ambiguities in the concept of growth in productive capacity, legitimate measures of this may yield values that vary widely.

We use data for the railroad industry, which have been compiled with care by Caves, Christensen, and Swanson [1980] and have already been used extensively in estimation work. We will only offer estimates of growth in crude productivity and in productive capacity, since welfare productivity is more difficult to measure.[3] We will show, in particular, that the estimated value of growth in productive capacity is quite sensitive to the assumptions used to impose equiproportionate movements upon the production function over time. It should be noted that although the analysis is carried out for just a single industry, the conclusions apply with equal (if not greater) force to the aggregate production function -- that is, the measurement of aggregate productivity growth must also be very sensitive to the assumptions used to impose equiproportionate growth.

We begin with what would seem to be the productivity concept whose measurement is most straightforward: crude productivity. Crude productivity growth measures the increase of output that can be produced with various combinations of inputs. In general, measurement of crude productivity growth is complicated by the fact that input quantities do not all grow in the same proportion with time, and more than one output is produced by any industry. This is certainly so in the railroad industry. As a result, to measure crude productivity growth it is necessary to define it as the difference in the rate of growth of an output <u>index</u> and a rate of growth of an input <u>index</u>. We can define an output index, $Y(t)$, and an input index, $X(t)$, at time t, respectively as

$$Y(t) = v_1 y_1(t) + v_2 y_2(t) + \ldots + v_m y_m(t), \text{ all } v_i \geq 0, \ \Sigma\, v_i = 1$$

$$X(t) = w_1x_1(t) + w_2x_2(t) + \ldots + w_nx_n(t), \text{ all } w_j \geq 0, \ \Sigma\, w_j = 1.$$

Crude productivity at time t is, then, given by

$$CP(t) = Y(t)/X(t)$$

and crude productivity growth by

$$CP(t)/CP(t-1) = \{Y(t)/Y(t-1)\} \ / \ \{X(t)/X(t-1)\}.$$

There is, unfortunately, no uniquely preferable set of weights either for the input index or the output index. We have used three sets of weights in measuring crude productivity growth in the railroad industry: (i) first period cost share weights, (ii) last period cost share weights, and (iii) an average of first period and last period cost shares. Five inputs are used in constructing the input index: (i) labor, (ii) way and structure, (iii) equipment, (iv) fuel, and (v) materials. Two outputs enter the output index: (i) freight ton miles and (ii) passenger miles; revenue shares are used as weights.

Estimates of both annual and annual average crude productivity growth are shown in Table 1 for each of the three indices. Estimates of annual rates of crude productivity growth turn out to be relatively insensitive to the choice of weights. The maximum difference in estimates resulting from the substitution of first year for last year weights is 0.53 percentage points (1957-1958) and in only one other case (1951-1952) does the difference exceed

Table 1

Estimates of Annual Rates of Crude productivity Growth in the Railroad Industry, 1951-74

	First Year Wts.	Last Year Wts.	Average Wts.
1951-52	2.405%	2.021%	2.214%
1952-53	1.079	0.995	1.038
1953-54	2.000	1.762	1.882
1954-55	7.813	8.080	7.946
1955-56	3.844	3.830	3.837
1956-57	0.026	-0.153	-0.063
1957-58	-0.653	-1.177	-0.914
1958-59	3.884	3.929	3.907
1959-60	1.932	1.840	1.886
1960-61	3.187	2.911	3.050
1961-62	5.435	5.466	5.451
1962-63	4.361	4.403	4.382
1963-64	4.481	4.433	4.457
1964-65	6.544	6.576	6.560
1965-66	4.734	4.685	4.709
1966-67	-1.325	-1.449	-1.387
1967-68	2.462	2.502	2.482
1968-69	2.352	2.315	2.334
1969-70	-2.055	-2.012	-2.034
1970-71	-2.014	-1.917	-1.966
1971-72	7.608	7.906	7.756
1973-74	5.045	4.999	5.022
Average Annual Productivity Growth, 1951-1974			
	2.927%	3.014%	2.940%

0.3 percentage points. There is only one case in which the sign is different (1956-57). Except for 1956-57, where the signs differ, there are only two cases where the percentage difference between estimates of CP growth resulting from the use of the two sets of weights differs by more than 10 percent: 1951-52 and 1957-58. The two estimates of average annual CP growth over the entire period 1951-74 differ by only 0.09 percentage points, or 3 percent.

7. Measurement of Growth in Productive Capacity[4]

The measurement of productive capacity growth differs from that of crude productivity in that we need to know at what rate the amount of output that can be produced with each combination of inputs increases over time. The basic problem, as has been noted, is that, in general, this will be different for different combinations of inputs. Indeed, the only practical way of measuring growth in productive capacity is the adoption of fairly restrictive assumptions about the nature of the techniques available to an industry and the way they change over time. In particular, it is necessary to assume that productivity grows at the same rate for each combination of inputs. But even this assumption does not yield a unique productivity growth figure, since there are various ways of imposing equiproportionate productivity growth on a production function.

We shall compare three such measurements of growth in productive capacity. The first uses a standard Divisia index and the other two are measures constructed by Caves, Christensen, and Swanson explicitly for use in the railroad industry. The reader may be quite surprised at how different the estimates are from each other and from the crude productivity measurements.

In this section we describe the three measures of growth in productive capacity that we used. Then, in the next section, the results of the statistical analysis will be reported.

a. The Divisia Measure. We will, in general, follow the derivation of the Divisia index in Hulten (1978).[5]

The assumptions are the following:

Assumption 1. The technology of a sector can be represented as a production function of the form:

$$Y = F(x_1, x_2, \ldots, x_n, t)$$

where Y is the output, F is an algebraic function, x_1 through x_n are the input quantities, and t represents time. This rather general premise already rules out discrete and abrupt changes in technology, which are common to most industries. The presence of the time parameter, t, in the production function also implies that technological change is "disembodied" (and, in particular, is Hicks-neutral).

Assumption 2. The technology represented by this production function is strictly quasi-concave and continuously differentiable.

Assumption 3. All input prices are determined in competitive markets.

Assumption 4. Producers minimize cost.

The next and critically important premise is

Assumption 5. The production function exhibits constant returns to scale (CRTS).

This assumption serves two purposes. First, it ensures that productivity growth is the same for each combination of inputs. Second, it ensures that the full value of the output is exhausted by the inputs. This last relationship is needed to derive the Divisia index.

We now define the Divisia rate of growth of productive capacity, DP*, as:

$$DP^* = \partial(\ln Y/\Sigma w_i \ln x_i)/\partial t.$$

Define the rate of growth of each input X_i as

$$X_i^* = d(\ln x_i)/dt$$

and the rate of growth of output as

$$Y^* = d(\ln Y)/dt.$$

Then, the Divisia rate of productive capacity growth is given by:

$$DP^* = Y^* - (w_1 x_1^* + w_2 x_2^* + \ldots + w_n x_n^*)$$

where q is the price of output Y, p_i is the price of input x_i, and

$$w_i = \frac{p_i x_i}{qY} .$$

In the case of multiple outputs, $y_1,\ldots,y_m$, the Divisia output index is

$$Y^* = v_1 y_1^* + \ldots + v_m y_m^*$$

where

$$v_i = q_i y_i / \Sigma\, q_i y_i$$

and q_i is the price of output i.

b. Adjustment for Returns to Scale. To measure productive capacity growth in the rail industry, Caves, Christensen, and Swanson [1980] propose a measure alternative to the Divisia index which allows explicitly for non-constant returns to scale. Their approach uses a general transformation function and its corresponding multiproduct cost function. Differentiation of the cost function leads to a productivity index which is a weighted sum of output growth less a weighted sum of input growth. The weights for output are the elasticities of total cost with respect to the output levels, and the input weights are the elasticities of total cost with respect to the corresponding input prices.

Formally, following the work of McFadden [1978] they adopt the following premises:

Assumption 1′. The technology of the railroad industry can be represented as an implicit production function of the form

$$f(y_1, y_2, \ldots, y_m;\ x_1, x_2, \ldots, x_n;\ t) = 0 \tag{4}$$

where f is an algebraic function, y_1 through y_m are various outputs, and x_1 through x_n are the inputs. This function is used because it is not, in general, possible to assign unique input quantities to each of the several outputs. As a result, the production structure must be defined implicitly as a general algebraic transformation function. As in assumption 1, discrete and abrupt changes in techniques are ruled out in this formulation and it is again assumed that technological change is disembodied.

Assumption 2′. The transformation function has a strictly convex structure.

Assumption 3′. Producers are cost minimizers.

Assumption 4′. While economies of scale may be present, their effects can be separated out.

Economies of scale mean that as the output level increases, measured productivity may increase -- without any change in technology. Roughly speaking, crude productivity -- that is, the measured ratio of outputs to inputs -- will increase either because of economies of scale or changes in technology. It is crucial to separate out these two influences when measuring productive capacity growth in an industry, since by the latter we refer only to the portion of the change in the measured ratio of outputs to inputs strictly ascribable to changes in technology. The Caves-Christensen-Swanson formulation here thus differs from the Divisia index by netting out the economies of scale factor (compare Assumption 5).

Assumption 5′. Prices need not be determined in competitive markets. In particular, as noted above, cost elasticities rather than relative prices are used to weight both outputs and inputs in the Caves-Christensen-Swanson productivity growth measure.

Technically, for this purpose the transformation function represented by equation (4) must first be reformulated as a cost function of the form

(5) $$C = g(y_1, y_2, \ldots, y_m;\ p_1, p_2, \ldots p_n;\ t)$$

where g is an algebraic function; p_1 through p_n are the prices of inputs x_1 through x_n respectively; and C is total cost given by

$$C = \sum_{i=1}^{n} p_i x_i \ .$$

Next, let the elasticity of the cost function with respect to input price i, η_i, be given by

$$\eta_i = \frac{\partial \ln g}{\partial \ln p_i}$$

and the elasticity of the cost function with respect to output y_i, e_i, be given by

$$e_i = \frac{\partial \ln g}{\partial \ln y_i} \ .$$

Then it can be shown that Assumption 4′ yields the following properties: First, the elasticity of the cost function with respect to input price is equal to the share of the input in the cost of the product:

$$\eta_i = \frac{\partial \ln g}{\partial \ln p_i} = \frac{p_i x_i}{C} = s_i$$

where s_i is the share of input i in total cost. Second, it can be shown that the rate of productive capacity growth, EP*, is now given by:

$$EP^* = \sum_{i=1}^{m} e_i y_i^* - \sum_{i=1}^{n} s_i x_i^* .$$

c. Adjustments for Returns to Scale and Capacity Utilization. In a follow-up article, Caves, Christensen, and Swanson [1981] provide two advances over their earlier paper [1980]. These are explicit measurement of economies of scale in the railroad industry and allowance for the possibility that inputs are not optimally employed.

The measurement of returns to scale uses an analytical procedure similar to the one the authors employed in the earlier article. They begin with a general transformation function describing the structure of production, which is given by

(6) $$H(\ln y_1, \ldots \ln y_m; \ln x_1, \ldots, \ln x_n; t) = 1$$

where H is an algebraic function and all other symbols are defined as before.

They argue that in the case where an industry produces only one output, productive capacity growth is defined as the rate at which output can grow over time with inputs held constant (i.e., $\partial \ln Y/\partial t$). In the case of multiple outputs, a "natural" definition of productive capacity growth is the common rate at which all outputs can grow with inputs held fixed:

$$\pi^*_y = \frac{d\ln y_i}{dt} = \frac{d\ln y_j}{dt} \qquad \text{subject to } d\ln X = 0,$$

where π^*_y is the rate of productive capacity growth from the output side. it can be shown that

$$(7) \qquad \pi^*_y = \frac{\partial F/\partial t}{\sum_{i=1}^{m} \partial F/\partial \ln y_i} .$$

It is equally "natural" to define productive capacity growth as the common rate at which all inputs can be decreased over time, with outputs held fixed:

$$\pi^*_x = - \frac{d\ln x_i}{dt} = - \frac{d\ln x_j}{dt} \qquad \text{subject to } d\ln Y = 0$$

where π^*_x is the rate of productive capacity growth from the input side. It can be shown that:

$$(8) \qquad \pi^*_x = \frac{\partial F/\partial t}{\sum_{i=1}^{n} \partial F/\partial \ln x_i} .$$

Next, it can be shown that π^*_y and π^*_x are related by the degree of returns to scale (RTS) in the transformation function. RTS is defined as the proportional increase in all outputs that result from a given proportional increase in all inputs, holding the production structure and hence time fixed. That is to say,

$$RTS = - \sum_{i=1}^{n} \partial F/\partial \ln x_i / \sum_{i=1}^{m} \partial F/\partial \ln y_i \ .$$

As a result,

(9) $$\pi^*_y = RTS\ \pi^*_x \ .$$

The returns to scale factor RTS can be estimated directly from the cost side. If there is a convex input structure and the firm minimizes cost with respect to all inputs, then the transformation function (6) associated with it has a unique cost function:

(10) $$\ln C = G(\ln y_1, \ldots, \ln y_m, \ln p_1, \ldots, \ln p_n; \ t)$$

where all symbols are as before and G is the cost function. RTS is then given by

(11) $$RTS = \left[\sum_{i=1}^{n} \partial \ln C/\partial \ln y_i \right]^{-1} .$$

The second major advance of the caves, Christensen, Swanson follow-up article [1981] is allowance for the possibility that not all inputs are employed optimally. As the authors argue, productivity growth estimates typically assume that the firm is in a position of static equilibrium -- in particular, that the firm is at a position of minimum cost with respect to all inputs. That is to say, it is normally assumed that the firm is operating at an efficient point in its production set. In reality, of course, firms often are not perfectly efficient. As the authors note, if the assumption of minimum cost is violated, "then estimates of (crude) productivity growth include the effects of...movements toward or away from equilibrium, in addition to shifts in the structure of production." (p. 994)

Their way of measuring productive capacity growth when some inputs are not in equilibrium is ingenious. Because the total cost function given by equation (10) will not be satisfied, they do not attempt to use it. Instead, they assume that the firm minimizes cost with respect to a subset of inputs (the so-called "variable" factors of production whose quantities can be changed in the short run) subject to the other input quantities remaining fixed. (These are referred to as "fixed" or "quasi-fixed" inputs.) In this way they can derive a variable cost function

$$(10) \qquad \ln CV = G^*(\ln y_1, \ldots, \ln y_m; \ln p_1, \ldots, \ln p_{n-q}; \ln z_1, \ldots, \ln z_q; t)$$

where CV, the variable cost, is given by

$$CV = \sum_{i=1}^{n-q} p_i x_i \ .$$

G* is the new cost function, and the z_i are the fixed inputs. The formulae for output productivity growth π^*_y, input productivity growth π^*_x, and returns to scale RTS must now be modified. It is shown (Caves, Christensen and Swanson [1981]) that the new equations are

$$(7') \qquad \pi^*_y = -(\partial \ln CV/\partial t)/\sum_{i=1}^{m}(\partial \ln CV/\partial \ln y_i)$$

$$(8') \qquad \pi^*_x = -(\partial \ln CV/\partial t)/\left[1-\sum_{i=1}^{q}(\partial \ln CV/\partial \ln z_i)\right]$$

and

$$RTS = \left[1-\sum_{i=1}^{q}(\partial \ln CV/\partial \ln z_i)\right]/\sum_{i=1}^{m}(\partial \ln CV/\partial \ln y_i) \ .$$

It still remains true that

$$\pi^*_y = RTS\ \pi^*_x \ .$$

This formulation of productivity growth is particularly useful when the railroad industry is not operating at full capacity.

8. Empirical Results

We shall now see that estimates of productive capacity growth are very sensitive to the choice of measure. We use the data provided in Caves, Christensen, and Swanson [1980, pp. 171, 172, and 175]. Three measures will be employed as estimates of annual productivity growth. The first, which I have called the "full Divisia index", measures productivity growth as the difference between a Divisia index of output and a Divisia index based on five inputs. The Divisia output index is a weighted sum of freight ton-miles and passenger miles, with revenue shares as weights. The Divisia input index is a weighted sum of the following inputs -- (i) labor, (ii) way and structures, (iii) equipment, (iv) fuel, and (v) materials -- with cost shares as weights. The second measure is also a Divisia index. It differs from the first in that only two inputs are used in the index -- labor and capital -- and the inputs are weighted by the share of labor and capital in the national income generated in the railroad industry. The third measure is the Caves-Christensen-Swanson index EP*, based on five inputs and two outputs.

The estimates are shown in Table 2. First, it is instructive to look at the estimates of annual average productivity growth over the period 1951-74. These estimates range from the 1.5 percent value of EP* to the 3.6 percent value of the two-input Divisia index, a 240 percent difference. Moreover, the two Divisia indices differ by 0.9 percentage points, or by 40 percent. The estimate of crude productivity growth over this period is about 3 percent per year, which also differs significantly from each of the three productive capacity growth measures.

Estimates of annual productivity growth are even more sensitive to the choice of measure. For 1952-53, estimates vary from <u>negative</u> 0.6 percent per year to positive 0.7 percent. In 1953-54, they range from -0.2 percent to 2.3 percent. In 1956-57, they range from -1.0 to 0.6; in 1957-58, from -2.6 to 1.4; in 1958-59, from 2.6 to 4.6; in 1959-60, from 1.3 to 2.7; in 1964-65, from 4.4 to 8.5; in 1966-67, from -2.3 to 1.2; in 1967-68, from 0.1 to 3.6; in 1968-69, from 0.9 to 2.9; in 1970-71, from -4.4 to 1.0; and in 1973-74, from

Table 2

Estimates of Annual Rates of Productivity Growth in the Railroad Industry Using Three Different Capacity Productivity Indices

Year	Full Divisia Index (DP*)	Divisia Index Using National Income Weights (DP*)	Caves-Christensen-Swanson Measure EP* (Bell Journal)
1951-52	-0.3%	0.5%	0.1%
1952-53	0.0	0.7	-0.6
1953-54	-0.2	2.3	0.1
1954-55	9.1	9.0	6.7
1955-56	3.8	4.3	3.1
1956-57	-0.5	0.6	-1.0
1957-58	-2.2	1.4	-2.0
1958-59	4.2	4.6	2.6
1959-60	1.7	2.7	1.3
1960-61	2.8	4.1	2.3
1961-62	5.7	5.9	4.6
1962-63	4.6	5.0	2.9
1963-64	4.4	4.3	3.4
1964-65	6.4	8.5	4.4
1965-66	4.6	5.5	3.4
1966-67	-1.5	1.2	-2.3
1967-68	2.5	3.6	0.1
1968-69	2.3	2.9	0.9
1969-70	-2.0	-1.4	-3.5
1970-71	-1.9	1.0	-4.4
1971-72	7.9	10.9	7.0
1972-73	5.5	5.2	4.6
1973-74	-0.2	-0.6	1.3
Average Annual Productivity Growth Rate, 1951-1974			
	2.5%	3.6%	1.5%

-0.6 to 1.3 percent per year. There is almost no consistency between any two of these measures.

Estimates of π^*_y and π^*_x are available only for average annual productivity growth for the period from 1955 to 1974. Comparisons of these with the three measures in Table 2 are provided in Table 3. The estimates of π^*_y and π^*_x based on the total cost function [equations (7) and (8)] range from 0.8 to 1.0 percent per year,[6] which are considerably lower than the Divisia estimate and lower than the EP* index. The estimates of π^*_y and π^*_x based on the variable cost function [equations (7′) and (8′)] fell in the range 1.8 to 2.0 percent per year, about twice the total cost function figures but still lower than the Divisia-based estimates of productive capacity growth.

9. Conclusion

In this paper we have shown that three basic concepts of productivity growth -- welfare productivity, productive capacity and crude productivity -- have very different meanings and uses and can behave very differently. All three of them are shown to be significant, even the crude productivity growth measure which makes no adjustments for quality changes and which may seem basically indefensible on first consideration. We have seen also why it may be impossible to devise any robust scalar representation of growth in productive capacity. We saw why explicit adjustments for changes in product quality may be unnecessary. That is, at least in principle, one can hope to deal with the quality change problem through reasonably accurate measurement of growth in welfare productivity. This, moreover, may in some ultimate sense, have the best claim as the true measure of productivity growth for purposes of economic analysis.

Table 3

Estimates of Annual Average Productivity Growth in the Railroad Industry, 1955-74, Using Different Capacity Productivity Indices

Measure	Productivity Growth
1. Full Divisia Index (DP*)	2.5%
2. Divisia Index Using National Income Weights (DP*)	3.7%
3. Caves-Christensen-Swanson (1980) Measure EP*	1.5%
4. Caves-Christensen-Swanson (1981) Measure π^*	
a. π^*_y (Equation (7): Total Cost Function)	0.9-1.0%
b. π^*_x (Equation (8): Total Cost Function)	0.8%
c. π^*_y (Equation (7'): Variable Cost Function)	1.8-2.0%
d. π^*_x (Equation (8'): Variable Cost Function)	1.8%

The empirical results demonstrate that estimates of both annual and annual average productivity growth over fairly long periods are highly sensitive to choice of productivity concept. Measures of crude productivity differ greatly from those of productive capacity. Moreover, estimates of productive capacity growth are very sensitive to the assumptions used to impose an equiproportionate shift on the production function over time.

Footnotes

*Princeton and New York Universities

**New York University

We would like to express our gratitude to the Division of Information Science and Technology of the National Science Foundation, the Fishman-Davidson Center for the Study of the Service Sector, and the C.V. Starr Center for Applied Economics at New York University for their support of the research reported here.

1. Where we deal with multi factor productivity the concept of a unit of input clearly involves serious aggregation problems. For the discussion here, however, that is an irrelevant complication and it will therefore be ignored. We will, in effect, proceed on the assumption that all input is homogeneous or that we are concerned with labor productivity. We will, however, return to the aggregation issues later.

2. It is even arguable that a new technique sometimes literally reduces the economy's ability to produce some outputs. For example, the use of concrete in the construction of buildings has probably reduced the opportunities for on the job training of stonemasons, whose quality of work on churches and gothic college buildings may thereby have been impeded.

3. The calculation of the consumers' surplus clearly requires estimates of the demand schedules for railroad output. We will also make no attempt to measure levels of or change in X-inefficiency over time.

4. This section describes the assumptions and the mathematical expressions used in the three measures of growth in productive capacity. The reader who is interested only in the results can proceed directly to section 8 which is comprehensible without knowledge of the formulas.

5. A similar derivation can be found in Gollop and Jorgensen [1980].

6. Two measures of π^*_Y were provided, based on alternative methods of estimation of the returns to scale (RTS).

References

Caves, Douglas, Laurits R. Christensen, and Joseph A. Swanson, "Productivity In U.S. Railroads, 1951-1974", Bell Journal of Economics, Vol.11, Spring 1980, pp. 166-81.

-------, "Productivity Growth, Scale Economics, and Capacity Utilization in U.S. Railroads, 1955-74", American Economic Review, Vol. 71, No. 5, December, 1981, pp. 994-1002.

Gollop, Frank M., and Dale W. Jorgensen, "U.S. Productivity Growth by Industry, 1947-73". in John W. Kendrick and Beatrice N. Vaccara editors, New Developments in Productivity Measurement and Analysis (Chicago: University of Chicago Press), 1980.

Hulten, Charles R., "Growth Accounting with Intermediate Inputs", Review of Economic Studies, Vol. 45 (3), No. 141, October 1978, pp. 511-18,

McFadden, Daniel, "Cost, Revenue, and Profit Functions", in M.A. Fuss and D. McFadden eds., Production Economics: A Dual Approach to Theory and Applications, (Amsterdam: North-Holland), 1978.