Volume Title: Education as an Industry

Volume Author/Editor: Joseph T. Froomkin, Dean T. Jamison and Roy Radner, eds.

Volume Publisher: NBER

Volume ISBN: 0-88410-476-1

Volume URL: http://www.nber.org/books/jami76-1

Publication Date: 1976

Chapter Title: Productivity Trends in Higher Education

Chapter Author: June O'Neill, Dean T. Jamison, Roy Radner

Chapter URL: http://www.nber.org/chapters/c4495

Chapter pages in book: (p. 349 - 370)

8

JUNE O'NEILL
Council of Economic Advisers

Productivity Trends in Higher Education

The relation between growth in output and growth in inputs in colleges and universities is of particular interest from the standpoint of public policy. The cost of higher education relative to the costs of other goods and services is in large part determined by relative rates of productivity change.[1] Since higher education has been heavily subsidized by public funds and private philanthropy, change in its price is of more than routine interest.

The traditional organization of the higher education establishment into public or private nonprofit institutions lends additional interest to the study along with additional difficulties. Economic research on productivity has generally dealt with private firms for whom pecuniary profit maximization is presumed to be a major goal. The goals of a subsidized, nonprofit firm may be more diffuse, and it is then a question whether such firms have the same incentives to minimize the costs of a given output or to pursue new and cheaper methods of production as energetically as those motivated by profit. Thus, a differential trend in productivity between higher education and other industries could be the result of differences in incentives as well as differences in the underlying technological production possibilities.

NOTE: This paper is based on research supported by the Carnegie Commission on the Future of Higher Education and was written while the author was at the Brookings Institution. Helpful suggestions were received from Dave M. O'Neill and Barry Chiswick. The views expressed, however, should not be interpreted as reflecting the views of the Carnegie Commission, the Brookings Institution, or of those who gave comments.

This paper presents estimates of the rate of growth of output and input in higher education over the period 1930–67. The results should be viewed as preliminary because the measures are far from perfect, particularly the measurement of output. Thus, the finding that output has grown at about the same rate as inputs suggests that productivity change may have proceeded at a lower rate in higher education than elsewhere, although a better measure of output might alter the finding.

I. DEFINING AND MEASURING OUTPUT

The definition and measurement of a unit of an industry's output is often a controversial issue. The proper unit is not always apparent (consider the medical care industry) nor is the extent or manner in which the qualitative aspects of the good should be measured.

In this paper, output is confined to the educational output of colleges and universities or what roughly corresponds to the instruction students buy with their time and tuition. Other outputs are produced by institutions of higher education and it is sometimes hard to draw a line between the purely instructional and other services. Whereas education may be viewed as the transmission of knowledge, and research as the production of knowledge, there can be considerable overlap between the two. Moreover, occupational guidance, future business contacts, and the atmosphere of a private club may be among the by-products of a college education. If the inputs that produce these other outputs cannot be separated from the purely educational inputs then, at least conceptually, these other outputs may be treated as qualitative aspects of the educational service. The problem is not unique to education—television may provide more than entertainment; a meal in a restaurant, more than nourishment.

Schools typically sell courses, so a simple quantitative measure of output would show the number of students times the number of courses per student per year. This is essentially the student load carried by schools and it may be conveniently expressed by standardized credit hours.[2] In any year, students in the United States are enrolled in disparate courses of study—resident or extension, degree-credit or non-degree-credit, part-time or full-time, summer and regular academic year. By assigning different numbers of credit hours to these different types of enrollments, I first derived a simple quantitative count (akin to full-time equivalents) reflecting the student course load carried each year, from 1929–30 to 1967–68.[3] In this initial form, the measure is equivalent to counting the number of aspirins as the output of the aspirin industry, or the number of automobiles as the output of the automobile industry. Obviously the measure will be deficient for measuring productivity change, and deficient to the extent that the "quality" of a standardized credit hour has changed over time.

Measuring the "Quality" of Output

By quality change economists mean a change in the ability of a given quantity of a good or service to satisfy some consumer want or "ultimate end". One approach to measurement of quality is to try to measure some aspects of the ultimate end or want more directly—as when student test scores are used as an indicator of the quality of a school's instruction. The other major approach, and one used more by economists and compilers of price indexes, is to specify various quality dimensions of the good or service (e.g. horsepower of the car, the "grade" of the meat, etc.) and then look to the market for information on how consumers value increments in these dimensions. Indeed relative prices of variants of a good, differing in particular specifications, are taken as indications of the relative quality of the variants. Adjustments can then be made for changes in price over time which are due to changes in the quality of a good, arising from changes in the mix of its specifications.

Of course, there are problems. The investigator must identify the relevant variants of commodities before they can be priced and it is possible that the more subtle specifications are missed in compiling existing price indexes. Another serious difficulty is that the relative prices consumers are willing to pay reflect relative utility only when the market is in equilibrium—a criterion seldom strictly met. The method fails particularly when new goods, or new substitutes for old goods, are introduced. While consumers are in the process of substituting the new for the old, prices will not be in equilibrium and price differentials will underestimate quality differentials. For this reason, quality change may be underestimated; and the greater the pace of new introductions, the greater the underestimate.[4]

The absence of a market price system in higher education adds an additional difficulty in estimating change in the quality of its output. Because of the variation in subsidies from school to school, differences in tuition charges need not bear any relation to quality differentials. Admissions are not rationed solely by price but also by various other devices.

Relative costs of different kinds of schooling may be substituted for relative prices in estimating "quality" differences, although this procedure raises problems. Without a market test, higher costs could possibly reflect inefficiencies in the use of resources rather than higher quality, since different kinds of schools may maximize different kinds of functions, and it is not clear what role the strictly objective factors play in the various functions. And, of course, relative cost data pose the same difficulties as relative price data do for constructing price indexes (e.g. relative costs change over time). However, cost differences are probably the most accessible route to measuring quality change in higher education arising from change in the mix of different kinds of schooling.

Pervasive cost differences have been found among different grade levels (upper- and lower-division undergraduate and graduate instruction) and among different types of schools (public and private universities, other four-year and two-year schools). Using available, but very limited cost studies, I have made a crude adjustment in the credit hours measure for change in the grade-level mix and change in the mix of schools.

More refined adjustments for change in the output mix can be made as the information becomes available. Thus, using data on the costs of different subjects, change in quality due to change in the subject mix can be added. Having adjusted credit hours for grade level, and subject, and possibly for type of school, one would have an output measure with a quality adjustment similar in spirit to that currently made for most goods. The question is then: What aspects of quality are omitted from the measure and how serious are the omissions in education versus other industries?

Unfortunately, I do not think the question is answerable, at least not with currently available information. Most price indexes, hence real output measures, are faulty for reasons given above, and we do not have any way of knowing the magnitude of the bias. By taking account of more subtle changes in specific aspects of automobiles (horsepower, weight, and length), Griliches adjusted the automobile Consumer Price Index and estimated lower rates of price increase for automobiles than the official indexes had indicated.[5] However, the technique leaves some questions unanswered and has not been widely duplicated for other goods. More elusive qualities of the "ultimate-end" variety (such as comfort and quiet for cars) are usually just not measurable at all. Moreover, education is probably among the more complex goods and services with a larger component of elusive and unmeasurable quality characteristics. Narrowing down the education service to a specific subject, grade level, and type of school still leaves much room for quality variation.

Colleges and universities transmit knowledge. There are, however, alternative means of acquiring knowledge, such as books, television, family, and friends. Why, then, do people attend schools? Presumably, schools are more efficient through conserving the student's time and by allowing for the sharing of market inputs. Schools also provide evaluations—certifications to prospective employers and to other schools on the expertise of the pupil.

Qualitative change in school services should not be confused with change either in the quality of the knowledge itself or in the characteristics of the pupil, since these are not usually inputs controlled by the school. Thus, it would be incorrect to attribute the fruits of expenditures on research in science as a quality improvement in education, nor should quality change in students be counted as a change in output. An improvement in the writing style of authors or in the reading ability of book buyers would not be counted as a change in the output of the book-printing industry. However, knowledge has changed dramatically and students have changed too. Since different transmission processes may be required for the new knowledge or for the new types of student, evaluating quality is especially difficult. The speed and success with which schools adapt to these changes would actually be another, although difficult to measure, aspect of their quality.

Two measures—earnings and test scores—are often proposed for estimating the more elusive quality characteristics of school output. However, for the purpose of measuring quality change over time, both measures are at present inaccessible. Changes in the earnings differential between college graduates and high school graduates can hardly be attributed to changes in the quality of schools alone. Changes in the demand for different kinds of labor, changes in the physical supply of the two groups, and changes in the personal characteristics of the two groups are likely to dominate trends in earnings differentials over time, and since only physical supply changes can be readily measured, the separation of the unique effect of school quality on earnings differentials would appear to be a formidable task.

Tests of cognitive achievement could be administered to a cohort at high school graduation and then to the same cohort at different stages of college. Differential scores (college versus high school) could then be compared for different cohorts over time. The type of test would have to be such that it would be independent of changes in informational subject matter. It is well established that student characteristics dominate any explanation of differences in student achievement.[6] Therefore, for purposes of comparing results over time, it would be essential to know the relation between score improvement and starting level. Then, while a change in cohort score improvement may indicate the direction of quality change, the problem of measuring the exact value of the resulting change in output would remain. It would appear that these are difficult demands to meet.

II. DEFINING AND MEASURING INPUTS

For the purpose of estimating productive efficiency, it is necessary to identify and measure all the inputs which have been organized to produce the corresponding output. I have measured real inputs in higher education as total compensation to the factors (the sum of current operating expenditures and estimated capital costs), deflated by relevant factor price indexes.[7]

It should be noted that, due to data limitations, students' time is omitted from the measure. Since education is a highly time-intensive good, the omission would be serious if significant changes have occurred in the amount of student time spent in acquiring a credit relative to the growth of other inputs. However, no information could be obtained on this question.

With respect to the input measure actually used, two problem areas may be identified. First, the inputs used to produce instructional services may not have been properly separated from the inputs used to produce research and other jointly produced output. Second, the price deflators may be inadequate.

The decision about what portion of total expenditures to attribute to student instructional services involved some guesswork, since the financial statements of colleges (at least as reported by the Office of Education) do not provide very refined detail on the breakdown of expenditures by type of output. I have confined expenditures on student instruction to the following reported categories: instruction and departmental research, libraries, plant operation and maintenance, and general administration. The last three categories (sometimes referred to as overhead or indirect costs) include some expenses properly attributable to extension and public service and, also, to organized research. To adjust for this, I subtracted an amount equal to 5 per cent of extension expenditures and 15 per cent of organized research expenditures from the sum of the four categories.[8]

Since colleges and universities appear to have expanded their services over time, it is possible that some of the increase in the included inputs should not have been attributed to student instruction. This may be especially relevant for services such as job placement and the different types of counseling, which are not included in the measure of output. Working in the other direction is a possible downward bias in the measurement of input growth due to the possible transfer of some research activities, formerly classified as departmental research but reclassified into the organized research accounts as outside financing grew over time.

The capital input was measured as the deflated services of the capital stock and thus includes: (1) depreciation of the reproducible assets (buildings, improvements, and equipment), and (2) foregone interest on total capital assets (reproducible assets plus land), which would be equivalent to the return that the capital could earn if invested elsewhere in the economy.[9]

Depreciation was taken as 2 per cent of the gross value of buildings and 5 per cent of the gross value of equipment per year. Net replacement cost was used to approximate the market value of the capital assets and a conservative 5 per cent was used as the return to capital.

The selection of a measure of the capital input is a complex and controversial issue. However, in practice, the rate of growth in the net capital stock deflated, in the gross capital stock deflated, and in deflated capital costs, were all very close, about 4.2 per cent a year between 1930 and 1967.

Price indexes again present a problem. The deflator for physical capital is undoubtedly downward biased, due to the well-known bias in construction cost indexes. This results in a probable underestimate of the rate of growth of the real capital stock and, therefore, of deflated capital costs.

The deflator for educational operating expenditures is essentially a weighted average of separate indexes of faculty salaries, nonfaculty salaries, and supplies and services. The faculty salary index is itself a weighted average of separate salary indexes for full professors, associate professors, assistant professors, and instructors. Both the faculty and nonfaculty salary indexes have two known biases that work in opposite directions. The indexes were based on annual (or academic year) earnings, hence the well-known decline in hours worked over time suggests that hourly salaries actually increased faster than indicated. On the other hand, some of the observed increase in annual salaries is probably due to an increase in the educational level, and presumably the quality, of both faculty and nonfaculty workers. There is not enough information on the magnitudes of the two events to decide whether the opposite working biases were completely offsetting. However, whereas failure to account for hours worked biases change in the *quantity* of inputs upward, failure to adjust for educational level biases the *quality* dimension of inputs downward.

The overall input price index is necessarily crude, since there was not sufficient information to identify in detail all of the different inputs represented in operating expenditures.

III. SOME TENTATIVE ESTIMATES OF TRENDS IN PRODUCTIVITY

The quality of educational data along with the enormous conceptual problems makes measurement in this area hazardous. In reading this section, one should keep in mind all the qualifications previously mentioned.

Between 1930 and 1967, credit hours increased at an average annual rate of 4.8 per cent (Table 1, row 1). At the same time the average annual increase in total real inputs was 5.1 per cent. Therefore, taking the residual as a measure of productivity change leaves a seeming decline in productivity of −0.3 per cent a year. However, the choice of terminal years does influence the rate of change calculated this way.[10]

For some subperiods, small positive increases in the residual emerge (1954–67, 1954–60, 1960–67). An index of credit hours per unit of input recorded biennially over the same time period (annually from 1965–66 to 1966–67) suggests that there has probably been no trend, although there have been fluctuations (Table 2). These fluctuations correspond, as one would expect, to those unusual fluctuations in enrollment accompanying unusual events—depression, World War II, postwar GI boom. (See the index of credit hours, Table 2.)

Adjusting output for two kinds of quality change alters the results in the expected directions, but the magnitudes of these effects are very small. The two adjustments use cost differences by grade level as an indication of quality differences (as explained above). First an adjustment for changes in the graduate-undergraduate mix was made for the period 1930–67. Then, a further refinement was introduced by adding an adjustment for changes in the lower-division–upper-division mix of undergraduate credits (only for the period 1954–58).

The first adjustment necessarily assumes that costs, and therefore quality, are constant throughout the four years of undergraduate training, differing only between the graduate and undergraduate levels. Graduate costs were estimated to be three times as much as undergraduate costs.[11] Accordingly, the adjustment procedure gives graduate credit hours during the regular session a weight of three and undergraduate credits a weight of one. All summer credits were counted as undergraduate credits (weight of one).

The second adjustment allows for higher costs (and higher quality) in the upper division of undergraduate instruction. Upper-division costs were estimated to run 50 per cent higher than lower-division costs.[12] So the adjustment procedure gives a value of one to lower-division credits and a value of 1.5 to upper-division credits.[13] Graduate credits in four-year colleges and universities were given a value of 3.75, which retains the original 3 to 1 relationship between graduate and four-year

TABLE 1 Average Annual Percentage Rates of Change in Alternate Measures of Credit-Hour Output, Total Inputs, and of Residual, All Colleges and Universities, 1930–67 and Selected Subperiods (per cent)

	Period[a]					
	1930–67	1930–40	1940–54	1954–67	1954–60	1960–67
Total credit hours						
1. Unadjusted	4.8	3.1	3.2	8.0	6.6	9.1
2. Graduate adjusted	5.0	3.1	3.5	8.1	6.6	9.4
3. Standard lower-division units	–	–	–	8.0	6.7	9.2
Total inputs						
4.	5.1	3.1	4.2	7.7	6.3	8.9
Residual[b]						
Using as output:						
5. Unadjusted credit hours	−0.3	0.0	−1.0	0.3	0.3	0.2
6. Graduate adjusted credit hours	−0.1	0.0	−0.7	0.4	0.3	0.5
7. Standard lower-division units	–	–	–	0.3	0.4	0.3

SOURCE: O'Neill (1971).
Row 1. Based on Appendix A, Table A-10.
Row 2. Based on Appendix A, Table A-10. Undergraduate and summer credits given a weight of one, graduate credits given a weight of three.
Row 3. Based on Appendix A, Tables A-12, A-13, A-14. All credits in two-year schools are given a weight of one, all undergraduate and summer credits in four-year colleges and universities are given a weight of 1.25, and all graduate credits are given a weight of 3.75.
Row 4. Based on Appendix D, Table D-1.

[a] The dates refer to the academic year ending in the year indicated.

[b] Average annual percentage change in credit-hours output minus average annual percentage change in inputs.

TABLE 2 Indexes of Credit Hours per Unit of Input and of Credit Hours, All Colleges and Universities, 1929–30 through 1966–67

Year	Index of Credit Hours per Unit of Input (1929–30 = 100)	Index of Credit Hours (unadjusted) (1929–30 = 100)
1929–30	100	100
1931–32	94	106
1933–34	88	96
1935–36	96	110
1937–38	101	124
1939–40	100	135
1941–42	93	127
1943–44	81	109
1945–46	104	153
1947–48	123	240
1949–50	110	241
1951–52	96	206
1953–54	87	211
1955–56	94	251
1957–58	90	282
1959–60	89	310
1961–62	90	353
1963–64	87	411
1965–66	93	519
1966–67	90	571

SOURCE: O'Neill (1971); Appendix Tables A-10 and D-1.

undergraduate costs. Thus, all credits are actually expressed as lower-division equivalents.

The two adjustments are extremely crude, based as they are on very limited data on differential costs. Apart from difficulties already noted, the cost studies used may not be sufficiently representative of all schools and cost differences may have changed considerably over time.

Although graduate credit hours grew at roughly twice the rate of total credit hours, they account for only a small portion of the total (7.1 per cent by 1966–67). Thus, the adjustment for the increase in graduate credits raises the measures of output growth and of productivity change, but only by a trifle. Since lower-division credits increased relatively rapidly, a consequence of the expansion of junior colleges, the adjustment for this factor almost cancels out the effect of the graduate-undergraduate adjustment, at least for the period 1954–67.

Further adjustments can also be made by disaggregating by type of school. Table 3 gives a summary of change in an index of output per unit of input, for both groups and for public and private institutions considered separately. The patterns of change have been different in public and private schools. In public schools, growth in output appears to have lagged behind growth in inputs from the 1930s to the 1950s, while inputs increased less rapidly than output from the 1950s to the 1960s. For private schools, the pattern is reversed although the changes are small. Within public institutions, there has been a large shift towards junior colleges, which are the lowest cost institutions even after adjusting for grade-level mix. Since this shift would largely cancel out the increase in graduate instruction, productivity in public institutions is biased upwards in Table 3 relative to private institutions.

Since enrollment has grown much more rapidly in public institutions (and especially public two-year colleges) which are lower-cost, and therefore likely to be lower-quality institutions, failure to adjust for the change in mix of schools would bias the aggregate measure of productivity change upward.[14] I have calculated a crude index of output per unit of input where output is adjusted for change in the mix of graduates

TABLE 3 Indexes of Output per Unit of Input Using Alternate Measures of Output, Public and Private Institutions, Averages of 1930–34, 1954–58, 1960–64, 1966–67 (1930–34 = 100)

Type of Output Adjustment and Period[a]	All Institutions	Public Institutions	Private Institutions
Unadjusted credit hours			
1930–34	100	100	100
1954–58	97	85	108
1960–64	94	86	98
1966–67	97	90	94
Graduate adjusted credit hours			
1930–34	100	100	100
1954–58	101	89	112
1960–64	100	92	104
1966–67	104	97	100

SOURCE: O'Neill (1971); Appendix Tables A-10, A-15, D-2, D-3.
[a] The data refer to the academic year ending in the year indicated.
1930–34 = the average for 1930, 1932, 1934.
1954–58 = the average for 1954, 1956, 1958.
1960–64 = the average for 1960, 1962, 1964.
1966–67 = the average for 1966, 1967.

and upper- and lower-division undergraduates, and, in addition, where output per unit of input is calculated as a weighted average for the following types of school: private university, private four-year college, private two-year college, public university, public four-year college and public two-year college.[15] With the average of the years 1954–58 as the base of 100, the index falls to 97 for the period 1960–64 and to 98 for the period 1966–67. Thus, for the period 1954–67, the effect of the added refinements of adjusting for mix of schools as well as change in the grade mix of students, is to lower the rate of productivity increase compared to that observed when no adjustment in credit hours is made at all (as in the upper part of Table 3).

As noted above, it should be possible to extend the cost-differences approach to make additional compositional adjustments for quality change in output due to change in the mix of subjects taught. However, available information does not indicate that such an adjustment would alter the results very much. Most cost studies show that courses in the physical and biological sciences and in health professions have the highest costs. The proportion of all bachelor's and first professional degrees granted in these areas seems to have declined somewhat over time (from about 17 per cent of all degrees in the 1930s to 13 per cent in the 1960s), with a resulting small shift towards degrees in lower-cost subjects—education, business and social sciences.[16] The same tendency is evident amongst Ph.D.'s. However, it is possible that the amount of science taken has increased for all students, regardless of field of degree, which might offset the shift towards lower-cost degree majors.

Further disaggregation than the broad adjustments given here would then be one approach to future research on productivity change in higher education. However, the potentially more serious sources of bias in the measures of output and productivity change probably lie in failure to account for quality change that is unrelated to grade level, type of school, or subject. And this brings us once more to the elusive areas of quality measurement. As discussed earlier, sophisticated studies of student achievement, holding precollege endowments constant, may in the future provide a way to estimate quality change in post-high-school education.

Although it is quite possible that quality has increased, there is still a question as to whether these quality changes would be sufficient to bring the estimate of growth in output per unit of total input up to the rate of the economy as a whole. To meet this mark, the rate of growth of the quality of output in higher education would have had to average between 1 and 1.5 per cent a year over the period 1930–67 (more, if quality of output is underestimated in other industries too, which is likely).[17]

IV. THE ORGANIZATION OF HIGHER EDUCATION AND INCENTIVES FOR CHANGE

The relation between the market structure of an industry and its rate of technological change has long interested economists. However, most discussion has been confined to the effect of differences in structure within the private profit-making sector of the economy and controversy has revolved around the question of whether technological advance is more compatible with bigness and monopoly or with competition (see, for example, Mansfield, ed. 1968).

Differences in the goals and structure of public and private nonprofit firms raise issues about which economists know very little. Within the private-for-profit sector there is the presumption that, although the degree of incentive may vary, the lure of profits and the rigors of competition provide powerful incentives for efficiency. According to this view, the constant search for lower cost (higher profit) methods of production promotes technological advance, whether it take the form of more subtle changes in business organization or of more dramatic changes in machinery and combinations of inputs.

Colleges and universities operate in a strikingly different atmosphere, where profits are not expected (indeed, they are frowned upon), and where public and private contributions are an important source of revenue. Of course, it is possible for state legislatures and private contributors to withhold funds when inefficiency is discovered. And this possibility probably does serve to restrain overt extravagance. However, outside funding depends largely on other considerations, and it is unlikely that the personal incomes of university administrators are as closely tied to the discovery of cost-saving innovations and factor substitutions as are those of entrepreneurs in private enterprise.

Numerous examples of what would appear to be inefficient practices have been cited by students of the higher education industry. Perhaps the most striking is the diffusion of the management role. Faculty are quasi-administration, quasi-employees and their self-interest may well conflict with optimum cost minimizing behavior. (For a more detailed critique of various practices, see Harris 1969; Schultz 1968.)

However, one can never be certain on the basis of a priori models of motivation or on the basis of ad hoc examples. Indeed, counter-examples are cited by those who stress that cost saving is primarily impeded by underlying technological restraints in higher education (Bowen 1968; Baumol 1967). Champions of this view point out that those forces influencing resource saving which are present in much of the economy simply do not apply to education. The "production process," as it were, is constrained by the basic requirement of one teacher to so many pupils and not much can be done to alter that basic fact without affecting the quality of the instruction. Consequently, substitution of physical capital for labor, or one type of capital or labor for another, are not as viable options as they are in the industrial world. Those who follow this line of reasoning conclude that as a result of this state of affairs, productivity change in education is doomed to be lower than that in the rest of the economy.

A number of questions may be raised. Colleges and universities use many resources other than teachers to supply the educational product; faculty salaries account for less than half of total educational costs. Without even changing pupil-teacher ratios, economies may be achieved in administration or in the use of libraries and physical capital. Furthermore, the relevant question with respect to quality change and pupil-teacher ratios is whether a given change in the ratio changes costs by more or less than it changes quality. Finally, I am not aware of any overwhelming amount of experimentation with nontraditional forms of instruction which would conserve teacher time. Nor are we, as part-time armchair entrepreneurs, likely to think of such things with the same urgency as would be the case if our incomes depended on it.

Since speculation cannot resolve the question, systematic and objective empirical studies that compare cost differences and productivity-change differences among edcational institutions producing similar outputs, but in which administrators operate under different personal-incentive environments, are needed. One of the few such studies compares costs in Navy training schools with costs of similar programs in private proprietary schools (D. O'Neill 1970). The results indicate substantially lower costs in the private schools although it could not be determined whether the quality of output was the same in both kinds of schools.

Very few colleges are private profit-making schools. Indeed the case of Marjorie Webster Junior College, a school excluded from even the chance of applying for accreditation solely on the grounds of its being a profit-making institution, suggests that the fewness of such schools may be in part the result of restrictive behavior by the college community (see Koerner 1970). However, many technical proprietary schools offer courses similar to those in the public junior colleges and these can be compared. Also, colleges and universities vary considerably in their reliance on student fees, endowment, gifts and state and local funds. Detailed study of schools classified by type of funding might provide insight into differential costs and rates of change in output per unit of input.

To sum up, there is a possibility that the rate of productivity change in higher education, even if positive, has not kept pace with the rate of productivity change elsewhere in the economy. Observers of the educational scene have offered various reasons why the productivity of resources in education might lag behind the rest of the economy.

The clarification of the issue has important implications for future decisions on methods of financing higher education. If present methods of financing (which give large sums directly to institutions) are found to affect incentives adversely and, in turn, to impede efficiency and productivity change over time, this could be counted as a cost of this kind of financing package, to be balanced against the advantages. Other financing schemes, such as giving aid directly to students, may have more beneficial effects on incentives and therefore lead to cost saving in the long run.

NOTES

1. The price of a product will obviously change depending upon change in the prices of inputs (including materials) used in its production and changes in the amount of output produced per unit of input. For a more elaborate discussion of the relationship, along with empirical results (see Kendrick 1961).
2. Woodhall and Blaug used a measure similar in spirit to measure the output of British universities: degrees awarded, weighted by length of course required for the degree, and including an allowance for "wastage"—those who did not complete the degree (Woodhall and Blaug 1965). One difficulty with the degree measure is that one may not be able to match the student load accurately with inputs for the appropriate year. Also, in the United States, where a very large percentage of students do not obtain degrees, and where the number of years taken to obtain a degree is extremely variable, the measure would be exceedingly cumbersome to put into standardized form.
3. The methods and data sources used to convert enrollment data into credit hours are described in detail in (O'Neill 1971). Note that the credit hour measure may be viewed as an alternate—and possibly more convenient—way of estimating comprehensive full-time equivalent enrollment.
4. For an interesting exchange on problems of measuring quality and constructing welfare-oriented price indexes see (Griliches 1964) and the comments following by (Jaszi 1964) and (Denison 1964). For some evidence showing that the Consumer Price Index may sometimes overstate quality improvements see (Triplett 1971).
5. The Griliches method is particularly useful for estimating quality changes that result in changes in the mix of characteristics which are so inherently a part of the basic commodity that they are not sold separately and therefore would not be counted as specifications in the making of official price indexes. Griliches uses multivariate techniques to impute shadow prices to these characteristics (Griliches 1964).
6. For a review of the literature referring specifically to higher education see (Berls 1969).
7. See (O'Neill 1971) for a detailed description of the construction of the index.
8. The data on expenditures are taken from publications of the Office of Education and have been adjusted and reported by O'Neill (1971). Note that current operating expenditures on student instruction accounted for only 50.7 per cent of total current expenditures in 1966–67.

9. The capital stock was estimated by cumulating increments of investment adjusted for price changes and depreciation (the perpetual inventory method). (See O'Neill, 1971 Chapter III for full details.)
10. Woodhall and Blaug concluded that there had been a *decline* in output per unit of input in British universities. However, their finding could be due to their choice of terminal years if British universities followed cyclical patterns similar to American colleges and universities. For the years that they selected (1938–52; 1952–62; 1938–62) output per unit of input declined in the United States too.
11. The adjustment factors used here are based on unit-cost differences found in a study of public institutions in Michigan. (See O'Neill 1971 for further details.)
12. Ibid.
13. All credits taken in two-year schools were counted as lower-division credits. Undergraduates and summer credits in four-year schools were assumed to be divided equally between upper and lower division.
14. The fact that private schools continue to attract even a shrinking portion of all students is a bit of circumstantial evidence of higher quality in private schools compared to public ones—especially when one considers that the ratio of tuition in the two sectors is much greater than the ratio of costs. Tuition and fees per credit hour in all private schools were 4.2 times as much as in public schools in 1966–67. Costs per lower-division credit were only 1.2 times greater in private than in public schools in the same year.
15. Index derived from Table 28 (O'Neill 1971). Fixed weight average is a weighted average across the six types of schools using the 1954 distribution of credit hours (adjusted for mix of credits by grade level).
16. From the beginning of the century to the 1930s the decline in health professions particularly (and also law) and the rise of education and business was still more dramatic. For the distribution by field of B.A.'s and first professionals (and also Ph.D.'s) by five-year periods from 1911 to 1953 see (Wolfle 1954, Appendix B). For more recent data see the Office of Education series on *Earned Degrees Conferred.*
17. Denison (1962, p. 224) gives estimates of growth in output per unit of input in the economy (for the period 1929–57) which range from .82 to 1.15 per cent a year depending upon the assumptions used to obtain measures of growth rates in output for those sectors of the economy where no direct output measures exist (e.g. government and nonprofit industries). While no comparable estimates exist for the whole period 1930–67, it seems likely that productivity growth during the sixties may have exceeded the trend for the 1929–57 period. It should be noted that Denison's estimates are based on a measure which relates changes in real output to changes in total factor input, and which counts changes in the quality of a factor as growth in the amount of that factor.

REFERENCES

Baumol, William J. "Microeconomics of Unbalanced Growth: The Anatomy of Urban Crisis," *American Economic Review* 57 (June 1967): 415–426.

Berls, Robert H. "An Exploration of the Determinants of Effectiveness in Higher Education." In *The Economics and Financing of Higher Education in the United States.* A compendium of papers submitted to the Joint Economic Committee, 91st Congress, 1st Session, Washington, D.C. 1969.

Bowen, William G. *The Economics of the Major Private Universities*. Berkeley, Calif.: Carnegie Commission on Higher Education, 1968.
Denison, Edward F. "Comment" to Griliches paper (Griliches 1964).
———. *The Sources of Economic Growth in the United States and the Alternatives Before Us*. Supplementary Paper No. 13. New York: Committee for Economic Development, 1962.
Griliches, Zvi. "Notes on the Measurement of Price and Quality Changes." In *Models of Income Determination*. Studies in Income and Wealth, Volume 28. New York: NBER, 1964.
Harris, Seymour E. "Financing Higher Education: An Overview." *The Economics and Financing of Higher Education in the United States*. A compendium of papers submitted to the Joint Economic Committee, 91st Congress, 1st Session, Washington, D.C. 1969.
Jaszi, George. "Comment" on Griliches paper (Griliches 1964).
Johnson, Harry G. "Correspondence on Productivity Trends in British University Education." *Minerva* 4 (Autumn 1965): 95–98.
Kendrick, John W. *Productivity Trends in the United States*. New York: NBER, 1961.
Koerner, James D. "The Case of Marjorie Webster." *Public Interest* 20 (Summer 1970): 40–64.
Mansfield, Edwin, ed. *Monopoly Power and Economic Performance*. New York: W.W. Norton and Co., 1965.
O'Neill, David M. *Meeting the Navy's Needs for Technically-Trained Personnel: Alternative Procurement Strategies*. CNA Research Contribution No. 151. Arlington, Va.: Institute of Naval Studies, Center for Naval Analyses, 1970.
O'Neill, June. *Resource Use in Higher Education: Trends in Output and Inputs 1930–1967*. Berkeley, Calif.: Carnegie Commission on Higher Education, 1971.
Schultz, Theodore W. "Resources for Higher Education: An Economist's View." *Journal of Political Economy* 76 (June 1968): 227–247.
Triplett, Jack E. "Determining the Effects of Quality Change on the CPI." *Monthly Labor Review*. Washington, D.C.: U.S. Bureau of Labor Statistics, May 1971.
U.S. Office of Education. *Enrollment for Advanced Degrees*. Washington, D.C. Published annually since 1959.
Wolfle, Dael. *America's Resources of Specialized Talent*. New York: Harper and Brothers, 1954.
Woodhall, Maureen, and Blaug, Mark. "Productivity Trends in British University Education, 1938–62" *Minerva* 3 (Summer 1965): 483–498.

8 COMMENTS

George B. Weathersby
University of California

In general, this is a provocative and well-written report summarizing a massive study of educational costs and outputs over the past forty years. Essentially, the reported results of essentially no productivity gains (or even slight productivity losses) over this period seem to confirm Radner's observation that the relationship between educational inputs and outputs has changed very little from the days of Socrates to our own. While this paper is well girded with caveats, qualifications, and disclaimers, there remain several areas of potential difficulty which are discussed below. These are not criticisms of O'Neill's paper alone but are more generally applicable to basic definitions, data, and their interpretation. Finally, I conclude with some general comments on policy analysis and the educational industry.

COMMENTS SPECIFIC TO THIS PAPER

O'Neill begins by declaring a public policy focus for her paper and then concentrates almost exclusively on institutional efficiency. In other words, all of the costs and outputs are viewed from the institutional perspective. Therefore, only direct operating and capital costs were considered, with no adjustment for either direct social costs or social opportunity costs, and the only outputs considered were full-time equivalent student years, which were derived from full-time and part-time head count enrollments by an assumed credit-hour equivalency. Research, public service, extension, regional economic development, and all other noninstructional outputs were ignored. To the extent that these noninstructional outputs have constituted an increasing proportion of the total outputs of higher education over the past forty years—casual empiricism suggests that they have—then the results reported in this paper underestimate the actual productivity increases of higher education.

The use of student credit hours (SCH) through the full-time student year concept as an output measure is subject to challenge. The credit value of a course is often determined arbitrarily by faculty members, in a manner unlinked to its resource use or to the production process. Some schools have assigned the same unit value to all courses regardless of their frequency or duration of meeting. Independent study is a growing component of many schools' instructional program with some schools reporting over one-third of their total SCH's in independent study. Typically, the credit value of a student's independent efforts is personally negotiated between each student and faculty member and is again independent of the resource use or relative productivity.

Focusing on the SCH ignores attrition. While the SCH is a process measure of sorts, it contains little information about the progress of students through an institution or the number of students who successfully complete their degree program. Roughly one-half of the entering freshmen do not receive a bachelor's degree. This may be irrelevant to an institution because degrees are free, or nearly so. (The recent decision by many law schools to provide former LL.B. recipients with a J.D. for only twenty-five dollars established a market price for doctorates at major institutions for the first time.) However, many public agencies and private firms have been persuaded that certification is valuable. Occupational placement and mobility depend far more on degrees than on SCH's. Manpower planning depends far more on degrees than on SCH's. I believe that many aspects of public policy depend far more on degrees than on SCH. This is not to argue that SCH's are meaningless, but that their public policy significance beyond the institution is somewhat limited.

The joint input/joint output problems are either handled arbitrarily (libraries, administration, and extension) or are ignored completely. The efficacy of this is again a matter of perspective. Most institutions of higher education are small, with enrollments of less than two thousand students, with undergraduate instruction as the institutions' main mission. Many of these schools are private and church-related. On the other hand, most college students and most resources are found in large institutions offering a wide range of instructional, research, and public service activities, and which are predominately public. Joint output activities abound in such schools. Once again, ignoring these joint outputs would bias productivity estimates downward.

As pointed out in the paper, using cost as a proxy for quality is problematical at best because it obscures the very relationship one is seeking—the relationship between efficiency and quality. It is all too easy to label any relative increase in resource use as an increase in quality with no supporting evidence. Alternative measures of quality that one might use are the ratio of the number of applicants to the size of the entering class, the student admit/accept ratio, the number of faculty applications/vacancy, or the number of job offers/junior faculty members. An index of institutional quality constructed by David Brown some years ago may be somewhat anachronistic but may still be useful.

In another vein, the dollar value of the capital facilities recorded on institutional ledgers may seriously overestimate both their value and their cost to the institution. Since 1963, the federal government has supported the construction of new educational facilities under the auspices of the Higher Education Facilities Act. This act provided that the federal government would bear the bulk of the costs of approved facilities. Therefore, while the total cost reported on the books of the institution might accurately reflect the market value of the facility, they could also overstate the cost to the institution by a factor of four to ten times.

Another interesting problem associated with physical facilities is the extensive use of restricted funds for capital outlay. Often the marginal utility to the institution of restricted funds is much less than their face value. A tragicomic example of this is one public institution whose largest bequest was a $10 million gift to be used exclusively for a canon law library. Although medieval church law is neither a major nor a rapidly growing field in a modern, socially conscious law school, this institution did accept the gift. However, when questioned, some faculty and administrators would have eagerly preferred $1 million or even $200,000 in an unrestricted gift. Because there were no other alternative uses of the money available, $10 million for libraries was added to the books even though neither the direct cost to the institution nor the value to the institution even closely resembled $10 million.

These last two points are only two obvious examples of one of the major pitfalls in meaningfully comparing institutional cost data—the changing mix of institutional revenues. In many respects, an educational institution is a little like a hungry family; it essentially spends all that it makes. In this sense, unit costs are more accurately described as unit revenues. Meanwhile, the O'Neill study used institutionally reported expenditure levels without adjusting for the different and changing mix of the sources of funds; the inclusion or exclusion of self-supporting enterprises, such as dormitories, cafeterias, hotels, hospitals, and even airlines further bias the results. In some institutions, the ratio of total expenditures per student is as much as twice the ratio of instructionally related expenditures per student, whereas in other institutions the two measures are nearly identical. If the appropriate data were available, only the instructionally related expenditures should be compared with the instructional outputs (ignoring joint outputs again). Furthermore, if the proportion of non-instruction-related expenditures has increased in the last forty years, then the reported productivity figures would be biased downward once more.

Finally, as the paper points out, in the recent past the acquisition of resources for education has not depended upon demonstrated efficiency. Therefore, one should not be surprised when there is no consistent pattern of efficiency demonstrated.

BROADER COMMENTS ON EDUCATION AS AN INDUSTRY

One conclusion that seems apparent, particularly from the papers by Levin and O'Neill, is that the theory of the firm is an incomplete and possibly inappropriate paradigm for *policy* analysis in education—particularly higher education. It is unnecessary to repeat here the many ways in which education falls short of the competitive market assumptions, or to speculate on what minimum behavioral assumptions are necessary for a rational analysis of public sector resource allocation.

However, we are presented with no evidence that either the variables chosen or the educational processes modeled in these papers are appropriate in the judgment of those who must use the analysis or that decisions based on these analyses would be significantly different from those made without these results. We are caught in a strange paradox of demanding a fistful of computer-calculated test statistics to indicate the statistical significance of a variable in a relationship, without presenting a shred of empirical evidence that any of these variables are relevant for policy purposes.

The quantitative analysis of cause and effect relationships undoubtedly has an important role to play in policy analysis but it is not yet in the proper context. Basically, we have no reference for the relevance of any particular variable or relationship without specifying the set of decisions and the level of decision making one is trying to affect. This seems to be the essential first question for meaningful policy analysis.