The Higher Education Production Function: Theoretical Foundations and Empirical Findings

DAVID S. P. HOPKINS

This paper reviews the literature of educational economics for evidence concerning the higher education production function. The results from a large number of sources are synthesized and presented in a systematic way so as to reveal what is known and what is not known about the character and form of the production function. From this synthesis, we derive a set of important unanswered questions and suggestions for future research.

The Production Function in General Form

The production function is intended to represent the process by means of which an institution—here, a college or university—transforms inputs (typically labor and capital) into outputs. In order to specify the function at all precisely, we must be able to: a) identify and quantify all relevant inputs and outputs, and b) describe the relationship between inputs and outputs in mathematical terms.

The general mathematical formulation is given in Hopkins and Massy (1981). Let

$\_ \quad {}_{1}, Y_2, \ldots, Y_m)$ be a vector of outputs, and

$\underline{X} = (X_1, X_2, \ldots, X_n)$ be a vector of inputs.

The production process is described by one or more functions of the type

$$F^k\,(\underline{Y}, \underline{X}) = 0. \tag{1}$$

To be a true production function, F^k (. , $\underline{X}$) should represent the *maximum* output $\underline{Y}$ obtainable from the inputs $\underline{X}$. The output measures must relate to the three primary missions of the higher education institution, namely, the transmission of knowledge (instruction), the creation of new knowledge (research), and so-called public service. One aspect of the diversity of the higher education industry in the U.S. is that different institutions place differing degrees of emphasis on these three missions; yet the research mission is clearly not limited to the major research university just as the public service mission is not limited to the publicly funded institution. Thus, most institutions have more than one major mission and, hence, a plurality of important outputs.

As Hopkins and Massy have pointed out (1981), the intangible features of both the inputs and outputs of the higher education production process are every bit as important as (many would say much more important than) the more tangible, easily quantifiable ones. These authors go on to identify a set of inputs and outputs similar to those shown in Table 1. They do not, however, carry the analysis any further in terms of specifying the exact functional form in (1) that relates these specific inputs to these specific outputs.

At this point, it would be well to observe that no researcher to date has successfully characterized the production function in terms as precise as the set of input and output variables listed in Table 1, and it is doubtful whether anyone ever will. The reasons for this are many, but they all boil down to the fact that the technologies of instruction, research, and public service are poorly understood, and the tools for estimating the requisite functional forms and coefficients are woefully inadequate to the task. To be more specific, not only are we lacking appropriate measures of quality, but the very nature of the interactions between, for example, teaching and research is difficult to express in mathematical terms. Definition of these intangibles can at least be approached by capturing these more subjective variables in a utility function that is optimized, as proposed by Hopkins and Massy (1981) and Garvin (1980), but that still does not solve the problem arising from the joint production of teaching and research. Another related problem for the economist is that, in the absence of any uniform, exogenously provided set of prices for the inputs and outputs of higher education, there is simply no way to escape the multidimensional character of the production function and all of the specification and estimation problems that it entails. Finally, we note that the concept of a true production function is based on an optimal technology, one that achieves maximum levels of output for a given set of inputs. Yet, as Levin (1976),

TABLE 1
Identification of Inputs and Outputs of Higher Education

	Tangible	*Intangible*
Inputs	New students matriculating	Quality and diversity of matriculating students
	Faculty time and effort	Quality of effort put forth by faculty
	Student time and effort	Quality of effort put forth by students
	Staff time and effort	Quality of effort put forth by staff
	Buildings & equipment	Quality, age, and style of buildings; age and quality of equipment
	Library holdings and acquisitions	Quality of library holdings and acquisitions
	Endowment assets	
Outputs	Student enrollment in courses	Quality of education obtained
	Degrees awarded	Quality of education obtained
	Research awards, articles, and citations	Quality of research performed (also quantity)
	Services rendered to the general public	Quality of services rendered
		Goodwill
		Reputation

*Adapted from Table 3.1 in David J.P. Hopkins and William F. Massy *Planning Models for Colleges and Universities*, Standard University Press. Copyright 1981.

among others, has pointed out, there is no reason to believe that the educational enterprise has been operating on the efficient frontier of production possibilities; and there are many reasons to believe that it has not. This means that, even if we were able to specify the true and complete functional form, we would still be unable to estimate the true coefficients of the model from any existing set of data.

It is apparent, therefore, that all efforts to date directed at specifying and estimating the higher education production function have provided only partial results. For example, quality measures have often been omitted, and most empirical studies have focused on the instructional production function alone without regard to the research objective, either on its own or in interaction with the instructional objective. Yet there are a great many such efforts reflected in the literature, and we shall summarize the important results below. First, however, it is useful to describe research concerning the nature and measurement of inputs and outputs.

Identification and Measurement of Inputs and Outputs

For purposes of estimating the production function for instruction, it is imperative to separate *student* input characteristics (e.g., numbers of students, aptitudes, family backgrounds, etc.) from *institutional* characteristics, such as faculty size and quality. Solmon (1973) has proposed specific measures of quality that can be used to characterize both student and institutional inputs. On the output side, nearly all researchers take some form of standardized test scores as a proxy for the amount of knowledge gained by students through the process of instruction. Astin (1973) has argued that attempts to develop a single overall measure of educational output are unrealistic; what is needed instead is a battery of measures that are sufficiently broad to capture the major outcomes of the educational process. He states, for example, that it is probably not sufficient to measure just the change in *mean* test scores resulting from students' college experience, since other features of the distribution of test scores, such as the variance or spread, may be just as important.

Perhaps the most ambitious effort to date to catalog the myriad outputs of higher education is represented by Western Interstate Commission for Higher Education (WICHE) (1970). The complete list of proposed output measures is shown in Table 2. Unfortunately, by concentrating solely on quantifiable measures, these authors have practically neglected the all-important quality dimension in their proposed output variables for research and public service. As Attiyeh has pointed out, "The number of pages of published research reports or the number of patents applied for or any other simple measure does not tell anything about the quality of research done in a university" (Lumsden, 1974, p. 6).

A somewhat more complete set of output measures for graduate education and research may be found in the National Research Council's (NRC) 1982 *Assessment of Research-Doctorate Programs in the United States*. This document gauges doctoral programs at American universities according to eighteen separate variables which are arrayed along the following six dimensions: program size, characteristics of program graduates, reputational survey results, university library size, research support, and publication records. Several direct measures of quality are included, especially in the reputational survey results. It is significant, however, that the NRC made no attempt to combine these variables into a single, composite indicator of university output.

TABLE 2
An Accounting Structure for the Outputs of Higher Education:
One Proposal

Instructional Outputs

Variables	*Source of Measures*
Cognitive Attributes of Students:	
Level of General Knowledge	Test Scores
Level of Knowledge in Chosen Field	Test Scores
Basic Language Arts Skills	Test Scores
Critical Thinking and Reasoning	Test Scores
General Intelligence	Test Scores
Affective Attributes of Students:	
Self-concept	Questionnaire Responses
Satisfaction with Educational Experience	Questionnaire Responses
Citizenship	Questionnaire Responses
Values	Questionnaire Responses
Achievement Motivation	Questionnaire Responses
Tangible Attributes of Students:	
Earning Power	Placement and Employment Data
Awards	Number and Stature of Awards
Affiliations	Number and Kind of Affiliations
Avocations	Number and Kind of Hobbies
G.P.A.	Academic Record Data
Level of Educational Attainment	Academic Record Data
Flexibility of Employment	Placement and Employment Data
Areas of Career Interest	Questionnaire Responses

Institutional Environment Outputs

Variables	*Source of Measures*
Academic Environment Attributes:	
Rate of Student Success	Dropout Data
Mean Time to Reach Degree	Student Record Data
Faculty Turnover	Faculty Record Data
Faculty Availability to Students	Student Questionnaire
Academic Resources Available	Library Data
Quality of Instruction	Faculty & Student Questionnaire
Academic Aptitude Mix	Entering Student SAT Scores
Student Stress	Student Questionnaire
Faculty Stress	Faculty Questionnaire

TABLE 2 *(continued)*

Institutional Environment Outputs (continued)

Variables	*Source of Measures*
Social Environment Attributes:	
Degree of Social Activity on Campus	Activity Records and Questionnaire
Racial Mix	Student & Faculty Records
Socio-Economic Mix	Student Records
Family Attitude Characteristics	Questionnaire
Social Involvement of Student Body	Questionnaire
Per cent Resident (on campus) Students	Housing and Student Records
Rate of Marriage Among Students	Student Records
Physical Environment	Physical Plant Data and Questionnaire

Research Outputs

Variables	*Source of Measures*
Reorganization of Knowledge	Number of new books, textbooks, etc.
New Inventions and Developments (Applied Research Products)	Number of patents, adopted procedures, etc.
New Ideas and Concepts (Pure Research Outputs)	Number of articles, papers, awards, citations, etc.
Personal Involvement of Students and Others (instructional spinoff)	Number of hours involvement on projects by students, industry, personnel, etc.

Public Service Outputs

Variables	*Source of Measures*
Student Involvement in Community	Hours of time, type of project, questionnaire
Faculty Involvement in Community	Hours of time, type of project, questionnaire
Cultural Activities Available	Number, type, duration, attendance, participation

TABLE 2 *(continued)*

Public Service Outputs (continued)

Variables	*Source of Measures*
Recreation Activities Available	Number, type, duration, attendance, participation
Continuing Education Activities	Number, type, duration, enrollment, quality, and satisfaction, questionnaire
Social Criticism	Amount, frequency, intensity, effects of confrontation —Students and Community —Faculty and Community
Personal Services	Number of health care patients, counseling patients, psychological testing, legal advice requests, etc. (dollar value of such services)
Indirect Community Benefits	Students available as employees, drawing power of the community as a place of residence for professional and skilled persons
Community Psychic Income	Public pride, awareness that expertise is available if needed
Product Testing	Number and types of products and materials tested for government and industry

Source: WICHE (1970).

The Relationship Between Inputs and Outputs

Production Functions for Instruction

We begin this summary of the empirical findings concerning the production function of higher education by limiting ourselves to the instructional objective alone, unrelated to any other objectives, since most of the empirical work to date has been carried out in this domain. These studies may further be categorized according to the predominant level of analysis, that is, whether the model is intended to represent the production process of an entire institution or of a single academic department, or the learning process of an individual student. Of course, the models and results in this section are directly applicable to those institu-

TABLE 3

Measures Compiled on Individual Research-Doctorate Programs in the Social and Behavioral Sciences

Program Size[1]

01 Reported number of faculty members in the program, December 1980.

02 Reported number of program graduates in last five years (July 1975 through June 1980).

03 Reported total number of full-time and part-time graduate students enrolled in the program who intend to earn doctorates, December 1980.

Characteristics of Graduates[2]

04 Fraction of FY1975-79 program graduates who had received some national fellowship of training grant support during their graduate education.

05 Median number of years from first enrollment in graduate school to receipt of the doctorate—FY1975-79 program graduates.[3]

06 Fraction of FY1975-79 program graduates who at the time they completed requirements for the doctorate reported that they had made definite commitments for postgraduation employment.

07 Fraction of FY1975-79 program graduates who at the time they completed requirements for the doctorate reported that they had made definite commitments for postgraduation employment in Ph.D.-granting universities.

Reputational Survey Results[4]

08 Mean rating of the scholarly quality of program faculty.

09 Mean rating of the effectiveness of the program in educating research scholars/scientists.

10 Mean rating of the improvement in program quality in the last five years.

11 Mean rating of the evaluators' familiarity with the work of the program's faculty.

University Library Size[5]

12 Composite index describing the library size in the university in which the program is located, 1979-80.

Research Support

13 Fraction of program faculty members holding research grants from the Alcohol, Drug Abuse, and Mental Health Administration, the National Institutes of Health, or the National Science Foundation at any time during the FY1978-80 period.[6]

14 Total expenditures (in thousands of dollars) reported by the university for research and development activities in a specified field, FY1979.[7]

Publication Records[8]

17 Number of published articles attributed to the program faculty members, 1978-80.

18 Fraction of program faculty members with one or more published articles, 1978-80.

[1]Based on information provided to the committee by the participating universities.

[2]Based on data compiled in the NRC's Survey of Earned Doctorates.

[3]In reporting standardized scores and correlations with other variables, a shorter time-to-Ph.D. is assigned a higher score.

[4]Based on responses to the committee's survey conducted in April 1981.

[5]Based on data compiled by the Association of Research Libraries.

[6]Based on matching faculty names provided by institutional coordinators with the names of research grant awardees from the three federal agencies.

[7]Based on data provided to the National Science Foundation by universities.

[8]Based on data compiled by the Institute for Scientific Information.

Source: *An Assessment of Research-Doctorate Programs in the United States: Social and Behavioral Sciences,* © Copyright 1982, by the National Academy of Sciences.

tions classified as "predominantly teaching institutions," whereas the interactions of instruction with research must be considered in the case of more research-oriented universities.

At the *institutional* level, the grossest form of production function can be represented by unit-cost ratios, such as dollar expenditures for instruction per student credit hour. Such ratios are often used as crude productivity indices (Wallhaus, 1975). In these instances, the sole measure of input is cost and the sole measure of output is credit hours. The implicit production function assumes a single-efficient-point technology with constant returns to scale; that is, it is of the form

$$y = ax, \tag{2}$$

where x = total expenditures on instruction and y = number of student credit hours produced.

A somewhat more sophisticated form of the instructional production function is expressed in Gulko and Hussein (1971). There, output is measured in terms of student enrollments, inputs are faculty and staff full-time equivalents and other university resources, and the level of aggregation is a cluster of academic departments representing a "discipline." The model takes the simple linear form:

$$\underline{x} = A \cdot \underline{y}, \tag{3}$$

where, now, $\underline{x} = (x_1, x_2, \ldots, x_n)$ is a vector of resource requirements, $\underline{y} = (y_1, y_2, \ldots, y_m)$ is a vector of student enrollments by level and major discipline, and $\underline{A} = [a_{ij}]$ is a matrix of input-output coefficients, such as the ratio of full professors in economics to upper-division undergraduate history majors. It is implicit in (3) that for any given set of outputs, there is a unique set of inputs, although the reverse is not true.

As I have pointed out in my article "On the Use of Large-Scale Simulation Models for University Planning" (Hopkins, 1971), this formulation of the university production function is simply not credible. Not only does it fail to take account of the university's research objective, but the implicit assumption of a single-efficient-point technology with constant returns to scale is patently unrealistic at this level of disaggregation. Hence, the aptness of such a model in describing the higher education production function is limited to institutions in which faculty are paid only to teach quite rigidly prescribed classes and students are severely restricted in their choices of which classes to take.

A somewhat different approach to representing the instructional production function for a university is given by Oliver and Hopkins (1976).

By introducing a time dimension, these authors portray the production process as a network of cohort flows in which students enter the system at various levels (freshman, undergraduate transfer, graduate student, etc.), remain for a certain period of time (cumulatively, that is—attendance is not assumed to occur in consecutive time periods) and then either graduate or drop out. A simplified representation of the network is shown in Figure 1. Outputs are represented in terms of degree-earners and dropouts, not just student-years of enrollment at the institution, while inputs are new matriculants in various degree programs plus the resources (faculty, staff, facilities, etc.) provided by the university. One particularly interesting feature of the model is that it explicitly accounts for the feedback effect whereby undergraduate enrollments create demand for graduate research assistants which, in turn, leads to requirements for new graduate admissions. Once again, the technology of

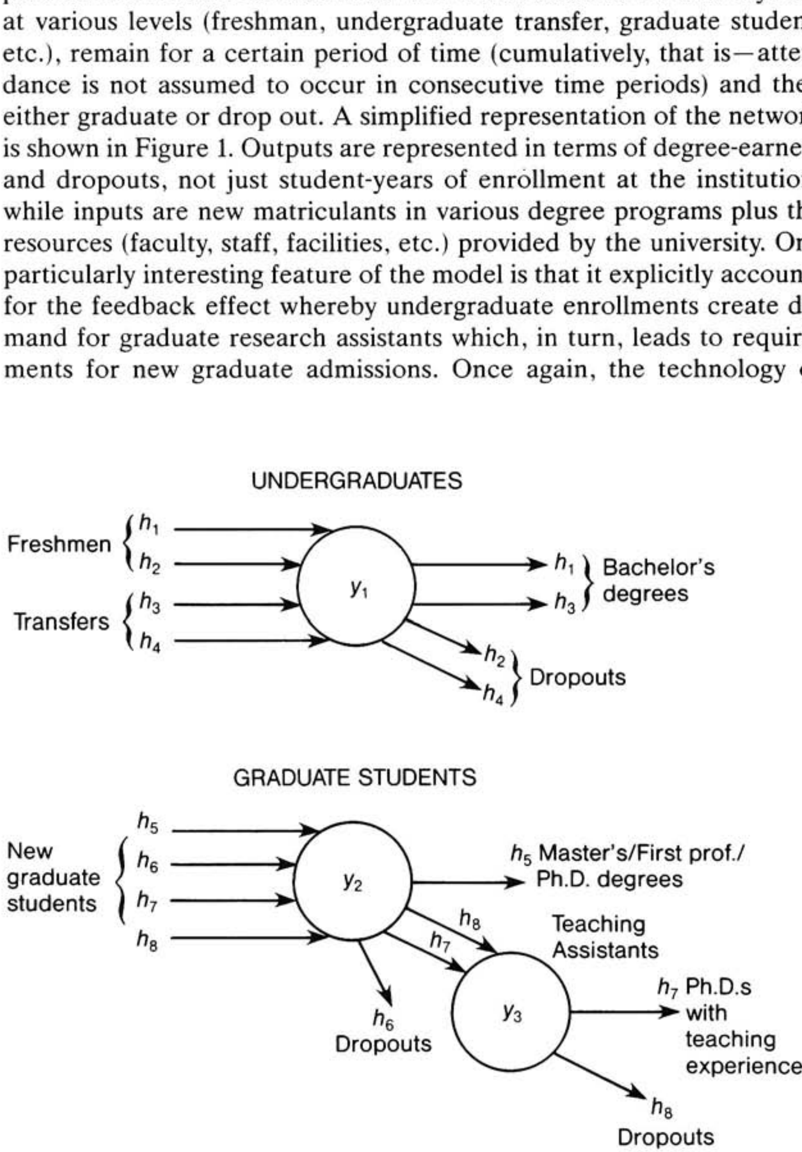

FIGURE 1. Network representation of the eight-cohort model. Reprinted from Figure 5.3 in David S.P. Hopkins and William F. Massy, *Planning Models for Colleges and Universities*, Stanford University Press. Copyright 1981.

instruction is represented by a matrix of faculty-student ratios, but in this case, the small dimensionality of the matrix (3 x 4) makes it feasible to vary its coefficients so that, in this sense, the model is neither linear nor does it assume a single efficient point.

One limitation of this model is that the flow network is assumed to be in equilibrium. That is, the flow rates of student cohorts, along with all behavioral and technological constraints, are fixed from one year to another. A time-varying version is given by Radner and Miller (1975, Chapter 10). While their model was formulated to represent the production function of the entire system of higher education, in which teachers are produced in one time period to instruct students in another, it is applicable to a single institution if one allows for an inflow of teachers from outside that particular institution.

An interesting application of production theory at the individual departmental level is given by Breneman (1976). His work is aimed at explaining the Ph.D. degree-granting behavior of academic departments in a prestigious university. A simple input-output structure of the following type is used to describe the Ph.D. production process (which we assume to be in equilibrium):

let

x = number of new matriculants per time period (input)
y_1 = number of successful degree-earners per time period (output)
y_2 = number of degree program dropouts per time period (output)
s = number of students enrolled in the degree program
f = success rate of matriculants (ratio of degree earners to new matriculants)
l_1 = average length of time to the degree
l_2 = average length of time to dropout.

From these definitions, we can write:

(4) $$y_1 = f \cdot x$$

and

(5) $$s = l_1 y_1 + l_2 y_2.$$

Breneman studies variations in the parameters f and l, which he assumed to be easily manipulated by the faculty, to differences in departmental prestige, demand for the Ph.D. output, and the amount of resources provided to the department. His chief hypothesis was that a faculty's concern over the prestige of their department would limit the production of Ph.D.'s in situations of limited demand for the product, regardless of the enrollment level, since prestige is directly related to the placement of graduates in other top schools. Although Breneman cites a variety of evidence in support of his theory, the evidence is certainly not conclusive.

Next, we turn to a brief review of production models which use the individual student as the unit of analysis. In all such models, educational output is measured in terms of level of student achievement in one or more categories or, better, in terms of the *change* in level of student achievement that results from the schooling process. Here, *student achievement* is typically measured in terms of standardized test scores. Within this category of models, we must distinguish between those that relate student achievement to a variety of school-related and non-school-related inputs, and those relating student achievement to investment to the student's time. Models of the former variety are represented generically by the following set of simultaneous equations:

$$A^k_{it} = f(\underline{b}^t_i, \underline{e}^t_i, \underline{s}^t_i), \qquad (6)$$

where

A^t_{it} = is the achievement level (or change in achievement level) in area k of student i at time t,

$\underline{b}^t_i$ = $(b^t_{i1}, b^t_{i2}, \ldots, b^t_{il})$ is a vector of background characteristics for the i^{th} student at time t (ideally containing some measure of innate ability),

$\underline{e}^t_i$ = $(e^t_{i1}, e^t_{i2}, \ldots, e^t_{im})$ is a vector of environmental influences affecting the i^{th} student at time $\underline{t}$ and

$\underline{s}^t_i$ = $(s^t_{i1}, s^t_{i2}, \ldots, s^t_{in})$ is a vector of school-related variables for student i at time t.

Obviously, the production function described in (6) can only be derived empirically if the functional form can be specified and its coefficients estimated from available data. Generally, the function is assumed to be linear (or log-linear) in all the independent variables, and the coefficients are estimated from large data bases using (simultaneous) least-squares regression.

A great deal of work was carried out during the 1960s and 1970s using the model formulation and estimation methodology described above for the purpose of estimating the 'school effect' on achievement in the primary and secondary schools. (One of the first such efforts led to the widely referenced and highly controversial Coleman Report, 1966.) This work has been well summarized and subjected to extensive critique by Cohn (1979), Hanushek (1979), Heim and Perl (1974), Lau (1979), and Levin (1976), among others. In spite of numerous attempts that were made over a period of some twenty years to relate various achievement measures to various characteristics of the student, his or her environment, and the schools, the results are often contradictory and largely inconclusive. It is not difficult to imagine the reasons for this, as the methodology can do nothing to reject, or even to indicate, the misspecification of variables and functional relationships. Yet our understanding of the true learning process is extremely limited and, hence, does not lead us to a unique specification of the appropriate production function.

The above model and approach have been applied directly to higher education in at least two instances. The often-referenced work by Astin (1968) attempted to relate social backgrounds and ability levels of college students and several measures of the quality of the college which they attended to their achievement scores on the Graduate Record Examination. The results were somewhat disturbing in that, once student ability and background had been taken into account, no differential effect from attending a highly selective (and presumably expensive) institution could be discovered. Manahan (1983) performed a similar analysis on a much more microscopic level. Using data obtained from a class in economics taught at Illinois State University in the fall of 1978, he found some positive association of change in standardized test scores with "quality of instruction," as measured by attendance and class participation. These studies are obviously subject to the same form of criticism as those performed using data from the primary and secondary schools. Suffice it to say that no reliable estimates of the true production function for individual student learning have been derived to date.

A different tack is taken in Becker (1983) and in Polachek, *et al.* (1978). These authors postulate a relationship between a gain in knowledge or achievement (the output variable) and the amount of time that a student invests in the learning process. Becker's production function is of the Cobb-Douglas variety and uses a measure of precourse aptitude along with time allocated to the course as inputs. Polachek, *et al.*, use a more general constant partial elasticity of substitution form with three input variables (single measure of precourse aptitude plus separate measures of time allocated to classroom instruction and studying outside of class). The latter authors were able to fit their model to data obtained from a special survey of students enrolled in a first-year economics course at the University of North Carolina at Chapel Hill. The results yielded some interesting figures on the marginal product of one hour's worth of class attendance versus the marginal product of the same amount of time spent studying. These figures are probably not generalizable to other situations, however, nor does the model account for any differences in output that relate to institution-specific variables such as method of instruction, quality of teacher, etc.

Some more recent studies analyze the substitution possibilities of new technologies in the individual instruction process. For example, Lewis, *et al.* (1985) report some evidence on the substitutability of computer-assisted instruction (CAI) for independent study time in terms of a student's gaining mastery of a fixed set of course material. These authors go on to report, however, that practically no current data exist on the cost-effectiveness of this instructional method compared with that of any others.

The Joint Production Function for Instruction and Research

In our rather extensive review of the literature, we were unable to locate any evidence (either theoretical or empirically based) concerning the production function for university research alone. We did, however, find several works dealing with the *joint* production of instruction and research. Models that incorporate the major interactions between instructional and research activities of faculty and students obviously are necessary if we are accurately to describe the production function of the research university. Yet the current state of understanding of these interactions—at least in any quantitative sense—is quite rudimentary.

A simple theoretical framework is provided by Nerlove (1972). This author examines the joint production of undergraduate education and graduate education coupled with research. He postulates that the production possibility curve must be the shape displayed in Figure 2. This curve shows a region close to each axis, in which the two outputs are postulated to be complementary to one another (more output of both is feasible under a fixed resource constraint), and a wider region in the middle in which the two function as substitutes. It is important to note that the outputs depicted along the axes of Figure 2 are intended to be measured in "quality-adjusted" units so that increases in output occur whenever quantity *or* quality is increased.

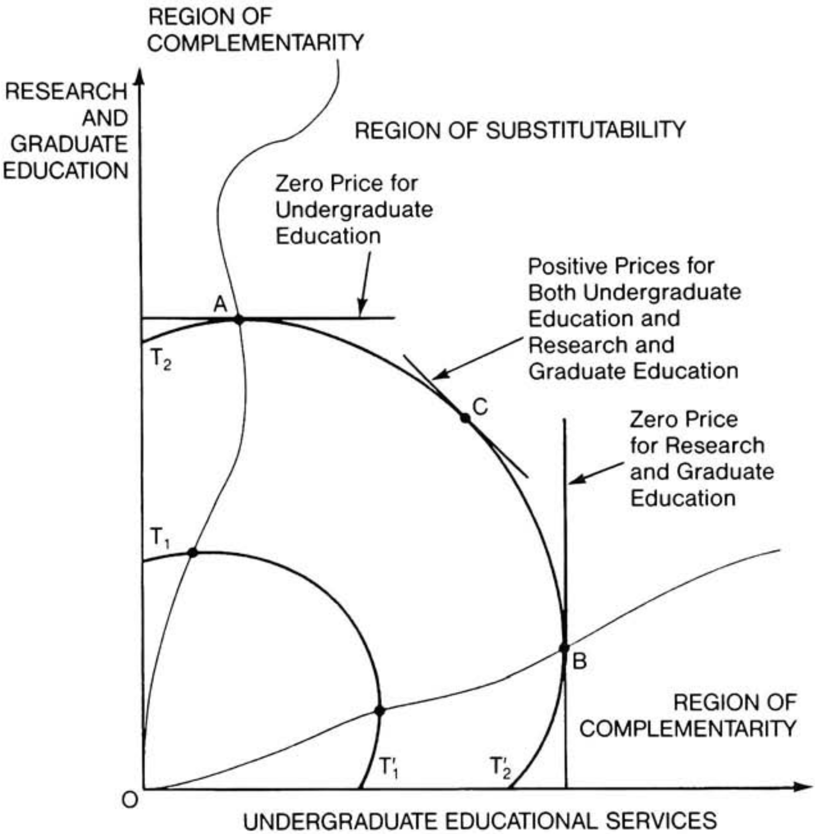

FIGURE 2. Combinations of undergraduate education and of research and graduate education showing the regions of complementarity and the region of substitutability

Source: Reprinted from Figure 1 in Nerlove (1972). ©1972 by The University of Chicago.

One inference drawn by Nerlove from his model is that it is more efficient to produce learning and new knowledge in the same institution than in entirely separate ones. This conclusion follows from the complementary supposedly exhibited between teaching and research when either activity is largely subordinate to the other one. Presumably, this concept of complementarity lies at the very heart of the existence of the research university.

Empirical models of the joint production function are extremely rare in the literature. Two particular attempts at direct estimation are reported by Southwick (1969) and Sengupta (1975). These two authors tried to fit numerous models to a statistical data base covering a variety of input and output measures for sixty-eight land grant colleges and universities over a six-year time period. Unfortunately, no model that was specified would fit the data with any reasonable degree of statistical significance, and so one must conclude that the variance in output measures (student enrollments and research expenditures) was largely explained by factors other than the input variables that were included in these various models.

Of course, even if we were successful in fitting such a model to real data, there would be no guarantee that the result would represent the efficient frontier of production possibilities. In fact, following the reasoning advanced by Cootner (1974) and the evidence presented in Carlson (1975), we expect that it would not, in which case we would only have succeeded in modeling the inefficiencies of the current educational system. This observation, coupled with the usual problem of specification errors has probably kept most researchers away from attempting to estimate the university's production function by any such direct means.

Another approach to modeling the joint production process is given in Chapter 5 of Hopkins and Massy (1981). Built upon the earlier work of Oliver and Hopkins, this effort was largely undertaken for the purpose of developing a framework for computing the full costs (direct and indirect) of instructional outputs. The model recognizes five general classes of faculty activity (classroom teaching and preparation, teaching outside the classroom, joint teaching and research, pure research, and administration) and develops a weighting scheme for apportioning total faculty effort among the five categories. Implicit in the weighting scheme are the trade-offs among teaching effort, research effort, and joint teaching/research effort. Yet, since these tradeoffs are all expressed in terms of faculty *input* variables, they do not tell us anything about the real tradeoffs of interest, namely those between teaching and research *outputs* for fixed faculty inputs.

Another means of examining the tradeoffs implicit in the university production function is through the formulation of economic models of individual faculty behavior. One such model is described in Becker (1975). In this paper, it is assumed that a professor chooses an allocation of his or her time among teaching (T_1), research (T_2), and leisure activities (T_3) to maximize a utility function $U(. \ , \ . \ . \ , .)$ defined in terms of teaching output (Q_1), research output (Q_2), and leisure-time consumption (Q_3). The output variables are further assumed to be linear functions of the time allocations and, in the case of leisure time only, of income (Y):

(7) $$Q_1 = a_{11}T_1 + a_{12}T_2$$

(8) $$Q_2 = a_{21}T_1 + a_{22}T_2$$

(9) $$Q_3 = bY \quad + a_{33}T_3$$

(Note that the "technology" coefficients, a_{ij}, do incorporate the 'jointness' between complementary teaching and research activities.) Finally, income itself is expressed as a function of the individual's output levels of teaching and research:

(10) $$Y = w_1Q_1 + w_2Q_2$$

This system of equations is easily manipulated to yield an expression for the professor's production possibilities for a fixed amount of total time (T) in the form:

(11) $$T = h_1Q_1 + h_2Q_2 + h_3Q_3$$

where the h are themselves explicit functions of the coefficients a_{ij}, b and w_i. From this expression we can see that the terms of trade among Q_1, $(Q_1$, and Q_3 are fixed (e.g., $dQ_2/dQ_1 = -H_1/h_2$) and the respective signs of the coefficients h_1 and h_2 indicate whether teaching and research are economic complements or substitutes. No data are reported by Becker, however, that would enable us to determine the nature or magnitude of these trade-offs in output space.

Directly related to Becker's work are several published reports that evaluate the relationship between research productivity and teaching effectiveness at various specific institutions (see, for example, Bresler, 1968, Hayes, 1971, and Voecks, 1962). The specified intent of these articles was to determine whether performance in research is positively or negatively correlated with teaching effectiveness in the university setting. Unfortunately, the results are quite mixed, with some schools exhibiting a positive correlation, others a negative one, and still others none that is measurable. Once again, the difficulties inherent in defining and measuring such concepts as 'teaching effectiveness' and 'research performance' in stark statistical terms have hindered us from compiling the evidence that we seek.

In summary, it is widely believed that teaching and research are at least partially complementary activities in the university setting. Several different attempts have been made to model the exact nature of the interaction between the two, yet none has succeeded in quantifying this joint production relationship in a way that would permit one to draw hard inferences based on real data about the economic 'terms of trade.' Thus, the empirical evidence we seek concerning the exact nature of the joint production process for instruction and research is still missing.

Economies of Scale in Higher Education

We turn finally to a review of the evidence concerning economies of scale in the production function of higher education. It is necessary to state at the outset that no evidence about economies of scale in university research could be found in any of the literature reviewed for this paper. Several efforts have been made, however, to determine whether economies of scale exist in the overall instruction process.

The Carnegie Commission on Higher Education (1971) explored the relationship between costs per student and institutional size using a national data base of colleges and universities. The data revealed a generally declining trend in unit costs with increasing size; this trend was especially pronounced for institutions with enrollment of under one thousand. These results suggest that there are increasing returns to scale for dollar expenditures as a function of enrollment over quite a broad range of institutional sizes.

This work was extended by Radner and Miller (1975) who performed a number of cross-sectional and longitudinal studies of faculty-student ratios as a function of institutional size. Here, again, student enrollment was used for the scale variable, while inputs were faculty rather than expenditures (although one would expect these two input variables to be highly correlated). Results were obtained from a national data set, stratified by type of institution, and showed for undergraduate-only institutions definite increasing returns to scale up to an enrollment level of between three and four thousand students. These increasing returns to scale were more pronounced for the private schools in the group than for the public ones. On the other hand, despite the authors' heroic efforts to examine the relationship of faculty inputs to size and several other variables at the public and private Ph.D.-granting institutions in their sample, they could discern no significant economies or diseconomies of scale for such institutions with respect to either undergraduate or graduate enrollment levels.

In addition to the studies already cited which draw evidence concerning economies of scale from a data set consisting of a wide number and variety of institutions, there is one published study that focuses on the same issue at a single institution. Dunworth and Bottomley (1974) formulated a cost model in terms of enrollments for the University of Bradford in the United Kingdom. Total costs per student were broken down into teaching costs, administration, student services, library costs, and capital and maintenance costs. The authors then studied potential economies that might arise from such strategies as increasing enrollments in specific disciplines with slack teaching capacity, increasing the utilization of expensive facilities (e.g., laboratories) through enrollment increases, and introducing alternative teaching structures (such as reduced number of contact hours per student, fewer courses offered, etc.). The analysis assumed that the proportions of time given to research and public service activities by the faculty were to remain constant at their then-current levels. Using real data from their university campus, the authors found a considerable potential for economies of scale that could be realized through better utilization of facilities and course offerings. These results should serve as yet another reminder of the hazards of using actual data on faculty-student or cost-student ratios as measures of efficient technological production possibilities in higher education!

In summary, there is considerable evidence for the existence of economies of scale for instruction in institutions of higher education. This evidence is found both in studies of aggregate 'productivity ratio' statistics across a large number of institutions, and of single institutional data on utilization of facilities and courses of instruction. The evidence is stronger in the case of institutions devoted primarily to teaching than it is for universities having a major research objective as well. Although intuitively one knows that economies of scale probably exist in research as well, no quantitative studies of the nature and extent of such economies were found in the higher education literature.

Suggestions for Further Research

Clearly, additional research is needed before we can specify the production function for higher education, or at least characterize the important trade-offs among inputs and outputs in anything other than purely qualitative terms. The following suggestions are made in recognition of the extreme complexity of the task, including the immeasurability of the many of the critical factors that enter into or result from this particular production process.

First, it does not appear particularly fruitful to examine any further the direct relationship between instructional output, as measured by student test scores, and variables representing attributes of the student and the institution. Such efforts have failed in the past due to inaccurate or incomplete specification of the 'learning model,' and any empirical basis for correct specification is still missing. Unless and until educational psychologists can reduce the learning process to quantitative terms with a high degree of accuracy, efforts by economists in this area will remain largely empty exercises in statistical manipulation.

We do not mean to suggest, however, that institutions should refrain from studying the impact of variables within their control upon the results of educational and research activities. Indeed, one area in which further research is badly needed is in the cost-effectiveness of alternative technologies for instruction. For example, in spite of the fact that many labor-saving devices have been recently invented in the form of computer-assisted learning devices and educational programming on television (including videotaped lectures given by the foremost authorities in many fields), the technology of instruction at most colleges and universities has scarcely been affected. The predominant mode of instruction remains face-to-face interaction between an instructor (often a highly paid member of the senior faculty) and a group of students in the classroom setting. Given the recent innovations referred to above, it would seem important to study alternative configurations of the instructional process in order to determine whether economies can be achieved through some substitution of capital for faculty labor without reducing educational effectiveness. Of course, one must be careful in conducting such studies to evaluate the quality of instruction along with its more measurable attributes.

Second, it is apparent that our ability to identify and measure the outputs of research is even more limited than in the area of instruction. Undoubtedly, this will remain an elusive problem, especially where the results of basic, as opposed to applied, research are concerned. We simply lack the means to define a 'quantum' of knowledge produced which could serve as a common unit of measurement across all disciplines or even across research teams operating in the same discipline. Neither can we anticipate future payoffs from basic research.

Here, again, we would propose less ambitious and more discrete efforts to study the production function for research in higher education. For example, we should be able to collect data on specific research teams in different disciplines within the same university. These teams are typically comprised of one or more faculty senior investigators, professional research associates, postdoctoral fellows, technicians, and students. What role is played by these actors individually and collectively and how do they contribute to the overall result? What would it mean to