WASHINGTON UNIVERSITY

(NASA-CR-154286) IDENTIFICATION OF AN EDUCATIONAL PRODUCTION FUNCTION FOR DIVERSE TECHNOLOGIES (Washington Univ.) 214 p HC A10/MF A01 CSCL 05I

N77-31004

Unclas G3/80 45603

CENTER FOR DEVELOPMENT TECHNOLOGY

REPORT NUMBER
CDT-77/2

JANUARY, 1977

IDENTIFICATION OF AN EDUCATIONAL PRODUCTION FUNCTION FOR DIVERSE TECHNOLOGIES

ROGER L. McCLUNG

WASHINGTON UNIVERSITY / ST. LOUIS / MISSOURI 63130

CENTER FOR DEVELOPMENT TECHNOLOGY

WASHINGTON UNIVERSITY
SAINT LOUIS, MISSOURI 63130

Report Number
CDT-77/2

January, 1977

IDENTIFICATION OF AN EDUCATIONAL PRODUCTION FUNCTION FOR DIVERSE TECHNOLOGIES

ROGER L. MCCLUNG

This study was supported by the National Science Foundation under Grant No. EC-38871 and by the National Aeronautics and Space Administration under Grant No. NGR-26-008-054, and was taken from a dissertation prepared for the Doctor of Philosophy degree in Economics. The views expressed in this memorandum are those of the author and do not necessarily represent those of the Center for Development Technology, Washington University, or the sponsoring agency.

ABSTRACT

Over the past ten years, viewing schools as though they were firms that use a variety of resources to produce desired outputs has become common practice. This analogy has inspired numerous estimates of educational production functions, which have become a practical way of investigating the effectiveness of schools and the productivity of educational resources. At the same time, instruction by technology has grown more sophisticated and expanded into more institutions. This thesis uses production function analysis to estimate the cost effectiveness of three alternative technologies in higher education: traditional instruction (TI), instructional television (ITV), and computer-assisted instruction (CAI).

The methodology of the educational production function is explored at both the conceptual and operational levels. The discussion of efficiency in education questions the existence of an educational production function and suggests weak behavioral assumptions in the educational sector. The characteristics of multiple production in education are described. The importance of the objective function is stressed, along with the insurmountable problems of estimating it. Several sources of bias are outlined that result from the researcher being forced to estimate a single equation model for what is truly a complex multiproduction activity.

With respect to operational issues, the multiple regression and frontier estimation (linear programming) techniques are compared.

Criteria and selection of a functional form are outlined and a general discussion of variable selection and measurement is presented.

The educational production function literature is reviewed with a summary of results on selected variables. Then, the expansion of ITV and CAI in higher education is sketched, followed by a survey of the literature on the effectiveness of these two technologies.

A unique set of data was collected for this research, consisting of students in comparable first-semester accounting courses at three universities. Each course was taught by a different technology: TI, CAI, and ITV. The result on an accounting examination especially prepared for this study was used as the output measure. Stepwise regression was the primary analytical tool.

The major findings of the production function analysis were:

(1) About a third of the variation in student performance was explained by the five variables representing program/technology and student ability. Student background, experience in related courses, and other standardizing variables did not appear important.

(2) The marginal productivity of ability appeared to be an increasing function with respect to mathematics and a decreasing function for verbal ability.

(3) There was evidence of interaction between ability and the technology inputs. Specifically, verbal ability was far more important than mathematics ability for traditional instruction, but this situation was reversed for both the ITV and CAI technologies. Students under either technology can be expected to perform at least as well as, and often 2 to 3 percent better than, the traditional student.

For the systems in this study, both CAI and ITV have lower unit cost than TI. As a result, CAI and ITV are more cost effective than traditional methods. If class size is small, CAI is more feasible, while ITV is preferred for larger classes. The assumed system characteristics--such as number of courses offered, the lifetime of the courseware, and enrollment per section--turn out to be critical to the cost-effectiveness results.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I am most indebted to the members of my thesis committee--Theodore Bergstrom, Robert P. Morgan, Barry D. Anderson, and especially the chairman, Edward Greenberg--for their insightful comments and generous donation of time. In recent months, Lee Benham has also contributed a great deal of assistance in the preparation of this thesis.

In addition, I would like to thank Harold J. Barnett for suggesting the research topic and Robert Morgan for arranging financial support. This research was funded by NSF grant No. EC38871 and NASA grant No. NGR-26-008-054.

Conversations at the early stages with Dean Jamison, William T. Garner, J. Alan Thomas, and Steven Klees helped shape the project design. From time to time, I have also benefited from conversations with Wesley Mellow.

However, this research could have never taken place without the generous cooperation of the following people during the data collection efforts: Molly Murray of Colorado State University, James C. McKeown and James Burton of University of Illinois at Urbana-Champaign, and David R. Ganz of the University of Missouri at St. Louis.

Finally, I appreciate the tireless efforts of Renae Dutkowski in editing the manuscript and Anita Knapp at the typewriter.

CHAPTER ONE

INTRODUCTION

Interest in the study of school effectiveness[1] has mounted over the last 20 years. In the 1950's and 1960's, researchers began to examine the connection between educational resources and program outcomes.[2] But those efforts were, and still are, hampered by the lack of measures for many important educational outcomes and by the absence of program effectiveness criteria. In general, a school or program will be judged effective if the inputs expected to make a difference show an independent effect on the performance measure or if the school or program compares favorably with similar programs.

The results of the Equal Educational Opportunity Survey (EEOS) in 1965[3] heightened interest in school effectiveness analysis and attracted the attention of economists. Around 1968, researchers began to view education as though schools were engaged in a production process using resources to produce outputs in much the same way a firm manufactures goods.

---

[1] School effectiveness study, input-output study, and educational production function studies are used synonymously.

[2] See S. Goodman (1959), Mollenkopf and Melville (1956), J.A. Thomas (1962), S. Marklund (November 1963), M. Johnson and Scriven (1967), and M. Nachman and S. Opochinsky (October 1958). See also the literature review in Chapter 3 of this thesis.

[3] Hereafter referred to as the Coleman Report after its principal author; J.S. Coleman et al. (1966).

This conceptual analogy has stimulated empirical literature offering various estimates of educational production functions.[4]

Although literature shows little quantitative agreement about any one input, it does reveal enough about the relative impact of educational resources to be useful to policy-makers.[5] However, this area of research has much to accomplish before it can adequately answer the following questions: What expenditures can be trimmed from an education budget without adversely affecting quality? Or, alternatively, what changes will result in the greatest improvement for the least cost? The research presented in this thesis is an attempt to improve the educational production function methodology.

Another significant development of the 1960's was an increased emphasis on communication technologies for student instruction, such as the widespread use of simple technology like overhead projectors. But even more important, instructional television was taken seriously at all levels of education and computers had begun to play a tutorial role in the classroom (for example, simulations and drill and practice).[6] Technology made it possible to present the same instruction to students at different locations. With some systems, students could play back material, thus allowing more individualized instruction. It was also thought that television

---

[4]Important examples are Bowles (1970), Katzman (1967), Kiesling (1967), Hanushek (1970), Levin (1970), Michelson (1970), and Burkhead (1967).

[5]Such results are examined in Chapter 3.

[6]For an account of new technology in education see P.H. Saettler (1968), Carnegie Commission (1972), G.C. Chu and Schramm (1967), H.J. Skornia (1968), and R.E. Levien et al. (1972).

might improve large group instruction. And it was hoped that technology would reduce the costs of instruction over the long run.

At the same time, educators were eager to evaluate the impact of instruction-by-technology on student performance. Researchers generally used the control-experimental group technique to examine the effects of technology within individual schools.[7] However, this thesis uses an alternate procedure.[8] It reviews the methodology of estimating educational production functions and uses that approach to determine the cost-effectiveness of three instructional technologies: traditional instruction (TI), instructional television (ITV), and computer-aided instruction (CAI). The research concentrates on a narrowly defined subject area: the outcome of a first semester course in college accounting. The subjects of this study are students at three colleges, each of which offers a comparable course in accounting, one by the traditional approach, another by ITV, and the last by CAI. The output is each student's performance on an accounting exam specially prepared for this study. Input information was obtained through student questionnaires and selected figures from student files.

The remainder of this thesis is organized as follows: Chapter 2 applies production theory to the study of education resource effectiveness. Special attention is given to the empirical procedure: the estimation technique, the form of the production function, and the measurement

---

[7]For a survey of this literature see D. Jamison et al. (1974), G.C. Chu and W. Schramm (1967), and R. Dubin and R.A. Hedley (1969).

[8]Two other studies using the production function approach to evaluate technologies at different schools are S. Wells (1974) and S. Klees (1975). These are discussed in Chapter 4.

of variables in the equation. Chapter 3 reviews the school effectiveness literature. Chapter 4 summarizes the use of technology in higher education with emphasis on ITV and CAI. The results of research on the effectiveness of technology in education are also reported.

The empirical work begins in Chapter 5 with a description of site selection and data collection techniques. This is followed by a discussion of the sample in the first part of Chapter 6. The primary statistical tool is regression analysis.[9] An attempt is made to isolate the impact of the three instructional technologies on educational outcomes by using different model specifications and various sampling stratifications. The rest of Chapter 6 is devoted to reporting results of the estimated equations on a variable-by-variable basis.

To compare the cost-effectiveness of the three technologies, chapter 7 calculates their unit costs and combines those costs with the estimates of the educational production functions from the previous chapter. Chapter 8 presents the recommendations resulting from the educational production function analysis and summarizes the major research findings.

---

[9]Principle component and factor analysis were also used, but to a lesser extent than regression analysis.

CHAPTER TWO

PRODUCTION AND EFFICIENCY IN EDUCATION: THEORY

Definitions and Overview

The production function is the fundamental concept in economics that describes the maximum output attainable from specified combinations of inputs. The output, or product, is any good or service whose generation requires one or more scarce resources. The inputs or factors of production are the resources used in the production process. Inputs may take the form of ingredients that are consumed during production or they may be durable factors, such as capital and labor, that offer a flow of services during production even though the inputs themselves are not consumed. The production function is always defined for a specified period. The length of the production period and the distribution between variable and fixed inputs are key elements in production theory. Over a sufficiently long planning period, all inputs are variable. In the conventional short-run period, however, one or more inputs must be regarded as fixed.

The production function is estimated to provide a measure of each resource's importance to the production process. The amount of additional output obtained by increasing an input by one unit is the marginal product of the input, or simply that factor's productivity. In any production period, the school administration must decide how much of each resource to purchase. That decision is subject to budget constraints and the cost of each factor.

In the conventional case, the size of the marginal product varies with the amount used.[1] Output will be maximized if the school purchases the combination of resources where the ratio of the marginal product to cost is equal for all factors.[2] This defines optimum resource allocation; when it occurs, the school is said to be allocatively efficient. A school that fails to equate the factor productivity per dollar at the margin is said to be allocatively inefficient. The school can become more efficient by trading some of the resources it possesses for others with higher productivity-cost ratios.

Cost effectiveness is a special application of the allocative efficiency concept. Rather than dealing with the factor productivity, cost effectiveness concerns the change in productivity associated with production process changes, such as different programs, methods, administrations, or technologies.

Including a shift parameter variable in the education production function, which changes for observations from different technologies, will measure the effect of technology on output, adjusted for the differences caused by other specified resources. This is the effectiveness of the technology relative to traditional instruction.

To determine which is most cost effective, a ratio of the effectiveness to unit cost is formed for each technology (effectiveness of traditional instruction equals one) and the ratios are ordered by size. The technologies with the larger ratios are the most cost effective.

---

[1]Whether this in fact is the case is an empirical matter.

[2]This is the result of the constrained output maximization problem for the single product firm. It requires a convex production set. See J.M. Henderson and R. Quandt (1958) for the proof of this result.

Realizing that the estimate of a production function is a serviceable piece of information, one may question whether computing such a relationship is feasible for education. Estimating the production function for any sector is no simple task. But the difficult issues of estimation technique, functional form, measurement of variables are reserved for a later section.

Efficiency in Education for the Single Output Case

At this point, a special problem must be discussed; that is the existence of an empirically estimable production function for schools. In other words, what assumptions about a school's behavior are necessary to estimate the input-output relationship and interpret that relationship as a production function? The following discussion of efficiency in education initially supposes that schools produce only one homogeneous output. After describing the problems in this simple case, multiple products are introduced and additional complications are pointed out.

Before going on, however, another concept of efficiency must be defined. Where allocative efficiency involved gaining more output (or reducing costs) by exchanging one group of resources for another, technical efficiency pertains to the application of resources within one's possession. A school is technically efficient if it produces the maximum possible output with the resources at hand. In other words, a school on the frontier of its production set is technically efficient. Another school operating at an interior point produces less than the maximum attainable and is termed technically inefficient.[3]

---

[3]Levin takes issue with the definition of technical inefficiency, and views it as a special case of allocative inefficiency. His argument runs

No evidence exists that schools are technically efficient. In general, they probably are not. Part of the problem lies with the fact that little is known about the production of learning. Not even a well-defined list of the resources involved is available.

However, suppose that we knew exactly what resources were required and that education produced only one output instead of multiple outputs. Under what conditions would a production function estimated across schools reflect the true underlying relationship?

The key issues emerge if the production function for education is compared with that for a manufactured product. In the case of a manufactured good, the product is well-defined, it has natural units (like boxes, loaves, cans, etc.), and the price determined in the marketplace provides a measure of value per unit. The profit motive offers an incentive for the manufacturer to allocate his resources very efficiently.

---

as follows: Presuming that a production process conforms to the laws of physics, matter cannot be "lost", nor "created." It must be possible then to map outputs onto inputs for any production process. Therefore, one must always be on the production frontier. Although one school may appear to employ the same amount of resources as another, but produce less output, it is only producing less of the desired or highly valued output but more of such outputs as leisure of students and teachers, ability to communicate with the opposite sex, self defense, and personality development. Thus, every case of technical inefficiency reduces to allocative inefficiency. See footnote 5, p. 22 in H. Levin (January 1974).

Subscribing to Levin's view of efficiency does not alter the conclusions to a production problem, but it has dire consequences for conventional application of production theory. Briefly, all production becomes multiproduct production; the theory of the single output firm is disallowed. Second, since one is always on the production frontier by definition, interpretation of the production function is easier, although measurement of each output may be difficult. Furthermore, the objective function must be estimated to derive shadow prices for each output in the production set. There seems no obvious advantage in trading ease of interpretation on the production side for a more complex objective function. Therefore, in our analysis, we take the conventional view that technically inefficient observations are possible.

Estimating the production function in education is not as clear-cut. The most important requirement to the analysis is that schools practice some approximation to maximizing behavior that will result in desired educational outcomes. The problems are obvious. Responsibilities for making production decisions may be fragmented so no one feels accountable for the results. Since the school planner is a manager and not a proprietor, he will not necessarily be motivated to minimize costs (or maximize output). As a result of the generally underdeveloped incentive structure in education, many schools undoubtedly are technically inefficient. However, if the inefficiency is neutral among inputs,[4] the estimated function will mirror the true frontier. Finally, a point interior to the production set could be symptomatic of a school's different definition or valuation of output. If there are proper behavioral incentives and agreement on the definitions and measurement of outputs across schools, the estimated function will represent the true relationship.

Efficiency in Education for the Multiple Output Case

Now that efficiency in education has been explored for the single output case, education can be looked at as a multiproduct endeavor. A school's objectives may be represented by cognitive goals, such as learning particular facts, developing certain skills, or getting all students to read above a certain level, and by noncognitive outcomes, such as imparting democratic values, shaping tastes and moral attitudes, or personality and physical development.

---

[4]That is, a school may become technically less efficient but maintain the same factor ratios. However, it is more likely that schools are more inefficient with some inputs than others. This issue is discussed at length in a later section.

The production relationships are enormously complex and interrelated. In thinking about modeling the process, one would expect some inputs to be necessary in the production of virtually all outputs, while others would be useful in only certain areas (a beaker for a science course or a slide rule for mathematics). The production of different outputs will not only be interrelated, but some of them, such as learning certain geometric principles and developing mathematical skills, will be joint products.[5]

Other outputs will be technically complementary (for example, knowledge of mathematics and physics). And still others will be technically competing outputs (for example, learning to translate Latin and dissecting an earthworm).[6] The production processes will be so interdependent, in fact, that it is difficult to think of a case where the production functions are separable.

Determining the production of any one output would require information about the levels of other outputs, the inputs used, and the effect of other production techniques on production of the output under consideration.

Educators consciously make production decisions among the various outcomes realizing that, frequently, more of one outcome can only be

---

[5]That is, they occur together. One could not learn geometry without increasing his mathematical skills.

[6]Two outputs, x and y, are technically complimentary with respect to input k if increasing the output of x increases the marginal product of k in y. They are technically competing if the marginal product declines, and technically independent if it is unaffected. Obviously, the same two outputs could be complimentary with respect to one input and independent or competing with respect to another. For a discussion of the model see S. Carlson (1939).

achieved by cutting back another. Thus, the study of resource allocation in the multiproduct case requires information on the relative values of the different outputs. In the private economy, the market establishes prices on goods which in turn determines their relative values. Since the cognitive and noncognitive educational outcomes are not priced directly in the marketplace, one might try revealed preference as an indirect way of establishing relative values. If one knew that certain schools produced different combinations of outcomes and the costs of education at those schools, the difference in demand arising from the various outputs could be isolated by observing the numbers of students purchasing education at each school. However, this method of calculating relative output prices would not work unless the schools defined the outputs in the same way and each output could be measured. Unfortunately, this is not the case.

The inability to define, measure, or set relative values for different outcomes makes it impossible to say how resource allocation could be made more efficient in the multiproduct case. Therefore, any study of efficiency in education must use the single output model. Researchers generally realize that proxies are not available for all outputs.[7] They cope with this problem by either ignoring the noncognitive outputs or assuming that each nonmeasurable component in the output vector is produced in a fixed ratio to the cognitive output measure. The remainder of this section discusses the bias and irregularities that may result

---

[7] Usually, a test statistic is used to represent the measure of cognitive output. This is discussed at length in a later section.

from estimating a production function assuming a single output for what is really a complex multiproduct activity.

Problems in Interpreting an Education Production Function

If a production function for cognitive achievement is estimated using data for a group of schools, an observation interior to the production set could mean that the school is technically inefficient or that the school simply prefers a different combination of outputs. Furthermore, the preferences may be caused by differences in the way the school defines, measures, or values certain outcomes.

This issue is even more difficult at the college level because of the extended range of educational outputs. College and university goals extend beyond the production of cognitive and noncognitive benefits for students. The goals have important research and community service functions. As a result, a college might produce less in the way of cognitive achievement because it has stepped up its research output. In addition, the incentive structure at the college level is more complex. Research, administrative duties, and community service functions may be more highly rewarded than excellence in teaching.

The fact that a school may appear technically inefficient with respect to the output being studied may result from different preferences, but it could also be caused by a lack of similarities in production techniques between schools. This happens when production processes are interdependent and the technique used to produce other outputs affects the productivity of inputs. To determine the conditions governing the choice of production techniques, the role of fixed inputs in the single and multiproduct cases must be examined.

Fixed input costs are not important to the production decision for a single-dimensional output. The stock of fixed factors merely sets the total production capacity. However, fixed inputs play a significant role in a multiproduct situation. The different products are forced to compete for the services of the fixed inputs. The price of the fixed inputs is the opportunity cost plus the cost of switching from one production line to another. Thus, the fixed inputs in the multiproduct case form a separate class of variable inputs and an equilibrium condition, which the single output model does not have, governs the efficiency of both variable and fixed inputs among the possible outputs. The techniques or factor ratios will be determined at the point where the value of the marginal product per dollar of input cost is equal <u>for all uses</u> for both variable and fixed inputs. Further, within each line of production, the physical productivity per dollar of input must be equal for all variable and fixed inputs.[8] As a result, the perceived efficiency of the technique that a school uses to produce achievement may be influenced by its choice of technique in producing other outputs, which may in turn have been caused by different endowments of fixed factors or different costs associated with fixed factors.

The assortment of fixed factor endowments, regulations, contract limitations, and locational and environmental advantages may mean that the schools will be subject to different production possibilities with respect to achievement production. Therefore, the educational production function for achievement is not defined here as a universal mapping of

---

[8]For a derivation and explanation of these conditions see R. Pfouts (1961).

production possibilities for changes in technologies, but rather as a production process function.

By claiming the presence of price ratios that could induce a school to jump from one process function to another, one might argue that education should be viewed as a single production function instead of separate production process functions. However, since the existence of institutions is often defined according to location, it does not always make sense to discuss prices at which production techniques can be transferred. For example, the University of Florida probably has a comparative advantage over the University of Nebraska in teaching marine biology.

Although the following quotation pertains to agricultural production functions, it nonetheless expresses a relevant concern for the problems associated with a single output model:[9]

> Obviously the simplest type of production process - at least in economic terms - is one that can be fully represented by a single unilateral causal relation. However, few real-world production processes of import can truly be represented by a single equation model...To a large extent, the single equation approach has been used because of its computational simplicity, the implicit hope of the researcher being that the single equation estimates are not greatly biased...In general however, the researchers have had no idea of the extent of the bias introduced by not using a multi-equation model. The single equation approach has been used without any appreciation by the researcher that a system of equations might be more appropriate - at least theoretically although perhaps not computationally. Such a situation is evident with regard to nearly all of the work that has been carried out with plants and animals.

It is also probably true that education researchers have not fully appreciated the differences between single and multi-equation models. It is hoped that this section has increased that awareness.

---

[9]E.O. Heady and J.L. Dillon (1960) pp. 201-2.

Estimating an Educational Production Function

Estimating an educational production function raises three fundamental issues: (1) Which estimation technique is appropriate? (2) What is the suitable functional form? (3) Which variables should be selected and how can each be measured? Each issue will be discussed in turn below.

Choosing an Estimation Technique

The production function underlying the education process may be either deterministic or probabilistic. If the relationship is deterministic, production is not subject to random effects, or $A = f(X_i)$. If the true production relation could be estimated, the production frontier would be precisely defined. With a probabilistic production function, given inputs result in a distribution of possible outputs such that $A = f(X_i) + u_i$. The distribution depends on random factors, such as rainfall in the case of agriculture. The production frontier is given by the mean and variance of the probability distribution. Since the education process is subject to uncontrollable random components the underlying production function is probabilistic.

With virtually no prior information about education's true function, the researcher must select the most useful estimation technique. The following paragraphs consider two possible methods: regression analysis, which computes the results for the average school, and the frontier estimation technique,[10] which uses linear programming to compute production coefficients based on the most efficient data points.

Most educational research has taken the former approach. Regression analysis evaluates all observations in the data set, ignoring the fact that some schools are technically less efficient than others. Since

---

[10]For an explanation of this approach see D.J. Aigner and S.F. Chu (September 1968) and C.P. Timmer (1969).

regression analysis fits a line by minimizing the sum of the squared deviations, the result is not a production function in the traditional textbook sense. Furthermore, the first derivative of such an expression does not represent the maximum additional output from an added unit of input. Rather, it must be interpreted as the expected marginal productivity-in-use of the input.

It is important to question whether the frontier and regression methods generate the same implications for allocative efficiency. For the moment, assume perfectly specified production model that is devoid of the usual measurement problems. Figure 2.1 depicts one production possibility set where a number of schools are using various amounts of inputs X_1 and X_2 to produce the same level of output (individual observations are not shown). Clearly, any school northeast of the production frontier is at a technically inefficient point and uses excess resources to generate the unit of desired output. Two types of functions have been estimated, a frontier function, A_1, and an average function, A_2. Allocative efficiency occurs at points B and C, where the price lines, G_1 and G_2 are tangent to the production functions. Since both B and C lie along the same ray from the origin, allocative efficiency is reached at the same ratio of inputs for each function.

If the slope is equal at each intersection of a ray drawn through the unit production contours, technical inefficiency is neutral among inputs. If the slope of the contours changes at these intersections, technical inefficiency is non-neutral among inputs.

Figure 2.2 illustrates the analogous case to Figure 2.1, except that technical inefficiency is non-neutral between inputs. According to the frontier estimate, equilibrium at B implies the allocatively efficient

Figure 2.1*

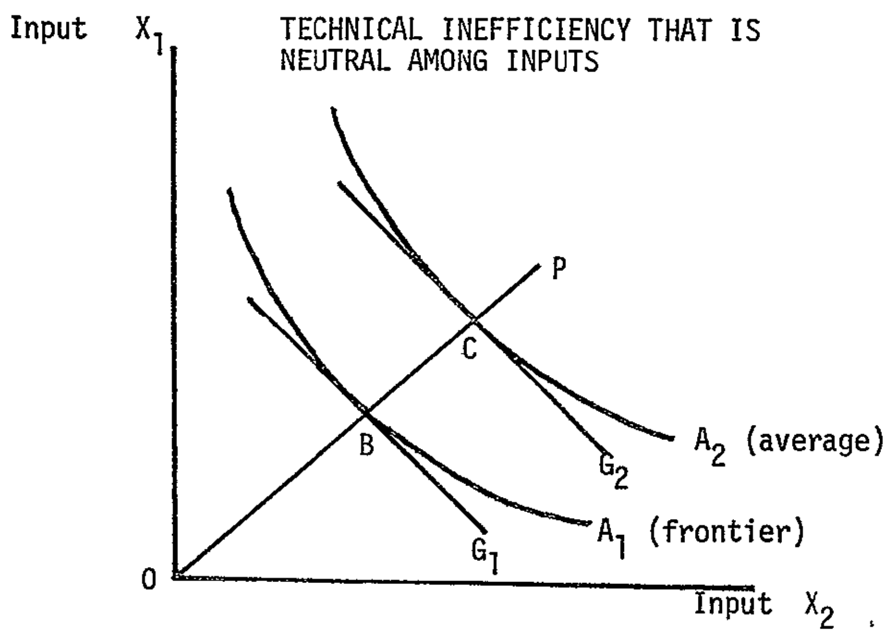
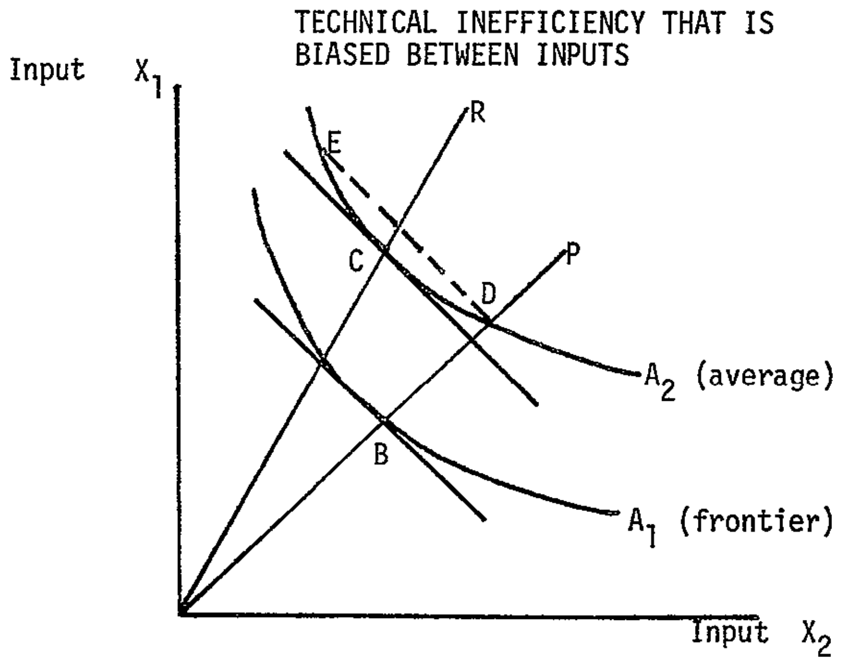


Figure 2.2*



Source: H. Levin, "Efficiency in Educational Production," Public Finance Quarterly (January 1974) pp. 11 and 12 respectively.

* In each diagram, A_1 and A_2 are unit production functions.

input ratio given by the ray OP. However, for the average school, this means that point D is allocatively inferior to the optimum allocation represented by point C and inferior as well to any combination within the shaded area enclosed by DCE.[11]

Each estimation method reaches the same conclusions with respect to allocative efficiency only if the technical inefficiency of schools is neutral among inputs. There is no reason to believe that the apparent technical inefficiency of schools will be distributed neutrally among inputs. If technical inefficiency in non-neutral, the two estimation techniques arrive at different recommendations. The input combination that will be optimal for the frontier schools actually leads to a reduction in allocative efficiency for the average school.

The situation portrayed in Figure 2.2 is supported by Levin's estimation of frontier and average functions for the same data set. Few of the coefficients on the frontier function bear any resemblance to those for the average function. For example, the marginal product of teacher's verbal ability at the frontier is 3.16 times its value for the average function. In the quotation below, Levin plainly echoes concerns expressed in the previous section on multiproduction:

> It appears that when student achievement is used as the criterion of educational output, so called frontier schools are more efficient in the use of some inputs and less efficient in the use of others. This suggests that the production isoquants for schools of different efficiencies with regard to the production of student achievement may be intersecting within the relevant ranges of factor substitution. This characteristic is probably attributable to differences in output mixes that are ignored in this

---

[11]The arguments and figures are credited to H. Levin (January 1974) pp. 2-23.

type of analysis. In other words, the optimal set of input proportions will vary from school to school depending upon its other priorities.[12]

In light of the possibilities for different results, which estimation is preferred? The frontier technique is a closer analogue to the textbook definition of a production function. However, determining the coefficients with only a few observations makes the model specification and measurement of variables even more important than they are in regression analysis. As discussed above, requiring measurement accuracy does not favor applying the frontier technique to education research. Furthermore, the random unmeasured factors make the frontier schools appear to exhibit superior efficiency and use fewer inputs for the same output, when, in fact, they may simply be experiencing a relatively higher level of beneficial random factors. Given the characteristics of the frontier process, regression analysis is the preferable technique for use with educational production data.

Specification of the Function

Some prior assumption must be made about the mathematical form of the function "f" in the general equation $A_t = f(X_1 \ldots X_m)$. For simplicity and ease of interpretation, most previous studies have assumed a linear form,[13] and the following equation is estimated using ordinary least squares (OLS):

(2.1) $$A_i = \sum_{j=0}^{m} \hat{b}_j X_{ji} + e_i$$

---

[12] Ibid., p. 19-20.

[13] See Kiesling (1967), Bowles (1970), Hanushek (1970), Levin (1970), and others. See Table 3.1.

for observations $i = 1 \ldots q$ and where $X_{oi} = 1$. With this specification, the marginal product of input j is equal to the coefficient $\hat{b}_j$. Hence, the linear form implies marginal products that are constant and independent of any input levels. Constant marginal products in turn imply perfect substitution between inputs. These are unlikely properties in education and most other production activities. Yet, researchers justifiably adopt the linear form, not for its descriptive realism, but because it often yields reasonably close estimates of production coefficients for more complicated, unspecified functions. If marginal products of the true underlying function were, in fact, constant, applying the condition for optimum resource allocation leads to the untenable implication that education should be produced using only one input. This is easy to see if the condition is written as:

$$(2.2) \quad \frac{\hat{b}_a}{p_a} = \frac{\hat{b}_b}{p_b} \quad \text{for all } a, b = 1 \ldots m, \text{ so long as } X_a, X_b > 0 \text{ and}$$

p_a, p_b are the prices of the inputs.

There will be one input whose marginal product per dollar is greater than any other input. In this situation, the school manager would maximize output for a given expenditure by using only that input in the education process. On the other hand, if factor productivity declines as more is used, other inputs soon become economical. The optimum combination of inputs occurs when the ratio b_j/p_j is equated for all inputs (except for the corner solution where $X_j = 0$).

Obviously, the results of a production function study largely depend on the initial assumptions about the functional form. Educational research provides little insight on the way educational resources may be combined. With no sound a priori information and no theoretical reason to favor any particular specification, researchers must resort to whatever functional form each considers reasonable. We will want to select a mathematical function with certain characteristics.

The following statements outline the desirable properties for an educational production function:

(1) Zero Output. The production possibility set should include zero. Whenever input levels are zero, output should also be zero.

(2) Essentialness of an Input. Some functional forms require a positive amount of each input for production to take place. But in education, it may be possible to eliminate one or more inputs (film strips and tape recorders, blackboards and chalk, or even books) and still have some learning take place. Thus, if just one input equals zero, the output should not be forced to zero.

(3) Marginal Products. This thesis will examine the direction and magnitude of different marginal products, the change in magnitude at different levels of input, and the extent to which one marginal product depends on other input levels. Although no evidence supports the allegation, declining marginal productivity of educational resources is intuitively reasonable. For instance, the first few years of teaching experience probably contribute more to a teacher's effectiveness than between the twelfth and fourteenth years. We obviously prefer a mathematical function which will not constrain the magnitude or changes in the marginal products.

(4) Elasticity of Substitution. The elasticity of substitution between i and j is the percentage change in the input ratio X_i/X_j that

would be induced by a 1 percent change in the ratio of their marginal products. Mathematically,

$$\sigma_{ij} = \frac{d(X_i/X_j)}{X_i/X_j} \div \frac{d(f_i/f_j)}{f_i/f_j} \quad .$$

If the elasticity is infinite, the inputs are perfect substitutes; if it is zero, input substitution is impossible. Intuition suggests that the educational process is not characterized by either extreme. For reasons that will become obvious, this thesis makes no attempt to quantitatively estimate the elasticity of substitutions in education. Nevertheless, a functional form that does not artificially bound but rather accommodates a variable elasticity would be preferable.

The empirical work that follows will be carried out for two models: the linear production function described above and a modified Cobb-Douglas function representing the class of non-linear models. The advantages of the linear model are its simplicity and its conformity to the non-essential input property. As previously noted, the major drawbacks to the linear form are the constant and independent marginal products. In spite of its shortcomings, the linear form is by far the most popular specification in education studies. In fact, much of the important work on educational production functions has been conducted exclusively with the linear model. It is included here to provide comparisons between past results and the outcomes of this study. Furthermore, a linear model can provide a close approximation of the underlying nonlinear function for cases exhibiting limited input variation.

The Cobb-Douglas function is an example of an intuitively more appealing nonlinear model. It is given by equation

$$(2.3)\qquad A = a \prod_{j=1}^{m} X_j^{b_j},$$

where a is a parameter. With this specification, the marginal products are variable and a function of the input levels. The marginal products are always positive and, as the input level ranges from zero to infinity, they decline monotonically. However, the Cobb-Douglas form does impose the constraint that cross partial derivatives be positive. Where $\partial A/\partial X_i$ and $\partial A/\partial X_j$ are positive, $\frac{\partial A}{\partial X_i\,\partial X_j}$ must be positive. In other words, high ability students will experience a greater impact from a given increase in teacher quality or school facilities than low ability students. The Cobb-Douglas function also restricts the substitution possibilities among inputs by forcing a constant elasticity of substitution equal to unity.

Probably the most distressing property of the Cobb-Douglas form is the essentialness of each input. If any $X_j = 0$, output must equal zero as well. In education studies, inputs will frequently be dichotomous variables measuring school quality such as teacher degree level or certification. The essential input property can be alleviated while preserving the character of the Cobb-Douglas function (2.3) when modified as follows:

$$(2.4)\qquad A = a\, e^{\sum_{k=1}^{p} b_k X_k} \prod_{j=p+1}^{m} X_j^{b_j},$$

where X_k are qualitative inputs, with $k = 1 \ldots p$, and where X_j are quantitative inputs, with $j = p + 1 \ldots m$. Thus, a qualitative variable may equal zero without forcing output to do the same.

The Cobb-Douglas model can be estimated by regression techniques once it is translated to the logarithmic form:

$$(2.5) \quad \ln A_i = \ln a_i + \sum_{k=1}^{p} b_k X_{ki} + \sum_{j=p+1}^{m} b_j \ln X_{ji} + e_i$$

for observations $i = 1 \ldots q$, and where e_i is the log of lognormally distributed error term.

When different functional forms are judged according to the production properties outlined above, no particular form demonstrates a clear superiority. Wells (1974) supports this conclusion:

> While functions which allow for variation in marginal products are intuitively more appealing than the linear function, there are no a priori theoretical reasons for choosing a particular functional specification.[14]

Hence, the criteria for preference among functional forms must be determined on the basis of statistical fit. In a recent study, Wells (1974) evaluated the transcendental-log production function and a general variable transformation form in addition to the two models presented above. He concluded that there is no convincing statistical basis for selecting any one form over another. Since there seems to be no reason to estimate a wide range of functional forms, the empirical work in this thesis is limited to two models, one linear and one nonlinear, which are estimated using equations 2.1 and 2.5 respectively.

<u>Specification and Measurement of Variables</u>

The goal of this research is to devise a model that can explain a student's outcome at a particular point in time. At the <u>conceptual</u>

---

[14] Page 117.

level, one might propose a function like that used by Eric Hanushek:[15]

$$(2.6) \quad A_{it} = f(B_{it}, S_{it}, I_{it}),$$

where A_{it} is a vector of educational outcomes for the ith student in period t and B, S, and I are input vectors relevant to student i and cumulative to period t. B comprises the individual and family background characteristics, S the school inputs, and I the student's initial endowment or ability.

The attractiveness of Hanushek's formulation compared with other input-output studies rests in his effort to incorporate the substance of an achievement theory. Specifically, the student's performance is a product of his inherent ability and the cumulative amount of capital embodied in him by his family and school. Few researchers would quarrel with the model at this level of abstraction. Most would agree that educational performance is determined by the direct effect and interactions of factors pertinent to these categories. However, differences of opinion arise once the variables are defined with the precision needed for empirical estimation. Defining and measuring the input variables will be discussed later in this section. The question of defining and measuring outputs will be dealt with next.

Measurement of Output

Since resource allocation is being studied here, educational output must be viewed from the perspective of the school administrator, budget

---

[15] See Hanushek (1970). Hanushek's formulation also includes variables to account for the impact of a student's peers and community on his achievement. Because the sample is inappropriate, these are not analyzed in the present research.

director, or educational consultant. Therefore, the output measures selected for a model must reflect the goals of the educational institutions involved. The preceding discussion on multi-output production defined educational output according to the customary cognitive and noncognitive dimensions. With few exceptions,[16] school effectiveness research has been limited to the cognitive outputs of education. This should not be taken as evidence that researchers consider the cognitive outputs more important than the noncognitive aspects. On the contrary, some researchers such as Bowles and Gintis[17] argue the opposite. In their view, achievement is a byproduct of noncognitive outputs--attitudes and values--that the school generates to socialize students into the existing socio-political structure, thus perpetuating the class structure. Virtually all educators would agree that the noncognitive benefits are important. Yet, because of difficulties in their definition, measurement, and valuation, these factors are infrequently studied. Neglecting the noncognitive outputs is obviously not a solution to the problem but, for the present state of the art, it seems the only recourse.

The remainder of this chapter discusses the guidelines for selecting a dependent variable to measure cognitive ability. Typical measures are scores on the Stanford Achievement Test, the Iowa Test of Basic Skills, and the College Entrance Examination Board Scholastic Aptitude Test.

---

[16]Several researchers have attempted proxies for noncognitive outcomes with less than promising results. The problem with interpreting proxy variables presented at the end of the preceding chapter applies here. For examples see, Burkhead et al. (1967), Levin and Michelson (1970), Averch and Kiesling (1970), and M.T. Katzman (1971).

[17]See S. Bowles (1972) and H. Gintis (February 1972).

Although standardized examinations are widely accepted and probably represent the best measure of general abilities, they are, at the same time, widely criticized. The researcher must be aware of the problems with standardized tests and make case-by-case judgments about which are appropriate instruments for effectiveness studies.

Three possible problem areas in standardized tests are their cultural bias, reliability, and validity. Cultural bias produces testing error resulting from the exclusion of ethnic or minority viewpoints from the test design. Reliability refers to the fact that test scores involve measurement error. In a standardized test, the dependent variable could account for only 81 percent of the variation in "true" student achievement.[18] The final problem pertains to test validity. Validity questions the similarity in the objectives of the test makers and the goals of the program being evaluated. Do the test questions measure the same attributes that the school administrators have in mind? If not, such test results should not be used to evaluate the program.

The problem with evaluating technology-based education is obvious. Standardized tests are designed for wide distribution and are most useful for the so-called "average" school. Hence, they are probably inappropriate measures of the effectiveness of new innovations in education. As the quotation below shows, Harnett agrees:

> These tests are almost always constructed so as to be widely appropriate and sufficiently general in nature to ensure their appropriateness for many educational experiences. Yet herein lies part of the evaluation problem. Criterion measures designed to be broadly applicable may well be too general in

---

[18]Averch, et al. (1972) p. 37. It is not possible to predict the reliability of the test measure used in the present research.

> nature to measure the specific outcomes of educational experiences at a local level.[19]

At the same time, an examination must not be so specific to one school that it loses its utility as a standard of comparison and evaluation between schools:

> Measures of a general nature yield little or no interinstitutional variation, while measures geared to the program of a specified department or institution do not allow for multi-college comparisons.[20]

The solution to the problem is to get away from general output measures and to move toward multiple, disaggregated evaluations of schools:

> Educational evaluators may have to turn to achievement examinations geared especially to syllabi used in specific college courses if they are to turn up indexes of college effects.[21]

This is the approach used in the present research.

After selecting an examination as the output instrument, one must decide which test statistic makes the best dependent variable. One may use an absolute measure such as a raw score, or a relative measure like a percentile rank, a grade, or an age-equivalent score. In some instances, the program objective may be a modification of the score distribution in a class or perhaps a percentage of students performing above some minimum level.

While percentile and equivalent scores give information about student performance relative to other students, they are poor indicators of school performance. Such measures tell nothing about the absolute level of performance. It is possible for a relative measure to remain the same or

---

[19] R.T. Harnett (1971) p. 17.

[20] _Ibid._, p. 18.

[21] _Ibid._, p. 18.

even decline during a period when a student's actual performance improved considerably. For this research, an absolute measure, such as the student's raw or percentage score, is preferable to any of the relative measures.

Intuition tells us that a school's effectiveness should be judged according to a "value added" criterion. Ideally, a researcher would strive for longitudinal data on students that incorporated both pre- and post-test measures of student achievement. One way to proceed would be to use the calculated gain score as the measure of school effectiveness. The inputs during period t would be related to the incremental gain in performance over the period by the production function $(A_{ti} - A_{t-1}) = f(X_1 \ldots X_m)$. Despite its intuitive appeal, this form offers no advantage over the reduced form $A_t = (A_{t-1}, X_i \ldots X_m)$.[22] Moreover, the reduced form is superior because it enables one to distinguish between equal gains for high and low ability students and, in general, allows the estimating procedure to select an appropriate coefficient on the pre-test measure. When longitudinal data are not available, it may be possible to use a proxy variable as a benchmark for aptitude in a subject area. Or one may be able to select a sample where students are likely to have no prior knowledge, such as first courses in foreign languages, chemistry, physics, and economics.

The last important issue is the appropriate level of data aggregation. Studies using the school or school district as the unit of analysis have been widely criticized as opposed to those using data on an individual

---

[22]For discussion see L.J. Cronbach and Furby (1970) and C.W. Harris (1963).

student basis. Resources are not distributed equally across students within schools, and the results of a study using mean student achievement as the dependent variable provide no information on how resources might be redistributed internally to increase efficiency. Averaging student scores within schools collapses the distribution of student output and provides less variation to be explained by the independent variables. As Averch put it:

> Roughly 30% of the variation in student's outcomes is variation among schools. Thus, an analysis of individual student's outcomes that uses school resources or peer group influence data aggregated to the school level can at best account for about 30 percent of the variation in student's outcomes. Analysis that use data aggregated to the district level are even more restricted because the variance in student's outcomes between districts is smaller yet--even more information is "averaged out" of the analysis.[23]

Therefore, a study in which the inputs explain one-half the variation in mean student test scores is really only accounting for 15 percent of the variation in true student knowledge (0.5 x 0.3 x .81).

For some of the reasons noted above, standardized examinations were not used in the present research. Instead, a special experimental examination was constructed, which is discussed in Chapter 5.

The Definition and Measurement of Educational Inputs

Considerable advancements have been made toward understanding learning in a controlled laboratory situation. For example, researchers are beginning to grasp the interacting physiological and psychological phenomena that enable one to read. However,

> The reader should realize at the outset that classroom and laboratory studies differ greatly in their objectives and

---

[23] H. Averch et al. (1972) p. 39.

> approaches. Classroom studies have not generally produced highly definitive results. Laboratory studies, however, have produced many significant and consistent results, but their relevance for classroom learning is often not clear.[24]

In the classroom, the participants are a less homogeneous group, the working conditions are not subject to the laboratory-like controls, and the variables themselves differ. Perhaps we do have a primitive theory of learning under certain rigid conditions, but we do not have a theory of instruction to help justify selecting relevant inputs to an educational production function. Educational researchers have little recourse but to use whatever input data happens to be available. Consequently, the measures for resource and background inputs vary widely among different studies. This will become evident below, as the discussion returns to the conceptual model presented earlier to define the input side of the production function (equation 2.6).

(1) Background. A student's educational achievement is strongly conditioned by the cumulation of background factors. Aside from providing for his physical welfare, the family implants a set of attitudes and values. Those attitudes and values affect the student's potential through the quality and quantity of verbal interaction and the nature of the home surroundings. Background influences have been shown to be highly correlated with the socio-economic status of the family. Socio-economic status is usually approximated by one or more of the following variables: family income, parents' education, parents' occupation, the presence of key durable goods in the home, and possession of dictionaries, encyclopedias, and daily newspapers in the home. Family size and mobility variables are

---

[24]Averch, et al. (1972) p. 51.

sometimes included to allow for variation in the emotional environment of the student.

(2) School Inputs. School inputs to the production function measure the quantity and quality of interaction between teacher and student plus the learning facilities at the student's disposal. A school's most important and expensive resources are its teachers. Some common proxies for teacher quality are average salary, years of experience, degree level, quality of college attended, socio-economic background, responses to attitude questionnaires, and tests of verbal or mathematical ability. Scale variables such as pupil-teacher ratio or total enrollment are sometimes included. Typical measures of facilities are the age of the school building, the number of library books per pupil, the existence of counseling services, or number of science laboratories.

(3) Innate Ability. Hanushek carefully explains that within the conceptual model,

> Innate ability refers to a pure genetic input that should not be confused with I.Q. or any other common measure of ability. Though we do not know of any satisfactory method of measuring this elusive concept, its inclusion in the conceptual model of the educational process is nevertheless of utmost importance.[25]

If we had a variable denoting the genetic component of I.Q. and comprehensive cumulative variables representing the influence of the environment, a separate proxy for ability would clearly be redundant. But, since there is no operational proxy for genetic input and background variables are likely to be inaccurate or incomplete, a separate measure of I.Q. or ability is usually included. This action is justified by the frequent

---

[25]E. Hanushek and J. Kain in Mosteller and Moynihan (1972) p. 123.

importance of student aptitude as a determinant of achievement in the presence of variables representing student background.

The fundamental problem of specifying and measuring educational inputs may be summarized as follows: First, the inputs to the production process must be defined at a conceptual level without the support of a definitive theory of learning. Second, to achieve an estimable model, proxies must be found to represent the domain and range of each selected conceptual input.

Most of the problems on the input side of the production function can be traced to the necessary reliance on data from "natural experiments." Student achievement varies simultaneously with changes in variables representing student background and school inputs. For the most part, students from advantaged backgrounds attend schools with larger budgets, which at the same time employ the most experienced and highly educated teachers. The proxy variables for socio-economic status and school quality are not independent of each other, although the regression analysis treats them as though they were. In this situation, the explanatory effect of one variable may be attributed to another correlated with it. For example, if a student ability variable is not included, the uncaptured effect of ability will probably exert an upward bias on the estimated effect of student background. If multicollinearity is severe, the standard errors will expand, variables will appear to be insignificant, and one would not be able to place confidence in the coefficients.

In the earlier discussion of output measurement, the importance of using the student as the unit of analysis was stressed. While many researchers have employed student background and ability measures on an individual basis, school resource input is generally not available for the individual student. Typically, data such as the number of teachers with more than five years' experience or the number of books in the library are collected at the school level and each student is assumed to receive an average amount.

Two points must be emphasized. First, as Garner (1973) correctly argues, the relevant inputs are not the stock of school resources but rather the services rendered as measured by each student's use of available school resources. Second, resources measured at the school level or averaged across students present a distorted pircture of the actual distribution of factors. The course selection process on the part of students and the student assignment practices of schools both contribute to a skewed distribution of school inputs among students. Moreover, failure to match school resources with students leads to specification bias where the complete effect of school inputs is either not revealed or appears as an upward bias in variables such as the background inputs that are positively correlated with school resource levels.

Still another source of bias occurs when a proxy variable is correlated with an omitted variable or unknown influence outside the model. Suppose, for example, that a significant and positive coefficient is associated with having a dictionary in the home. One would not want to give this a literal interpretation because it is surely not the dictionary itself that makes the difference. However, assigning any special interpretation to the coefficient would simply force the variable into a role based on ad hoc conjecture. A more reasonable conclusion is that possession of a dictionary is a surrogate for some uncontrolled attribute or combination of factors in the home that has a favorable effect on student performance. Failure to recognize the cases where inputs represent uncontrolled surrogates for other important, but unknown, factors leads to seriously misleading interpretations.

The next chapter contains a general survey of the school effectiveness literature and summarizes the results found there for the key variables of interest to this study.

CHAPTER THREE

SCHOOL EFFECTIVENESS STUDIES: A REVIEW OF THE LITERATURE

In the 1950's, school cost-quality studies generally used the number of graduates or percentage going on to college as a measure of success or output, and assumed that expenditures per pupil measured school quality. As this chapter shows, the major weakness of those early early efforts was their lack of control for other influences, especially student background.

The ambitious Equal Educational Opportunity Survey (EEOS) of 1965 provided the research community with one of the most extensive compilations of educational data ever assembled. The official survey document, the Coleman Report (1966), generated interested in the data and prompted a great deal of much needed research. In addition, its major finding on the overwhelming influence of socio-economic status and the relative unimportance of school inputs touched off a debate that still continues.

In the late 1960's economists began to study the school effectiveness issue. Kiesling (1967), Burkhead (1967), and Katzman (1967) were the forerunners, but it was Bowles (1970) who first outlined a detailed methodology applying the theory of the firm to educational institutions.

During the decade since the Coleman Report (1966), no less than 40 distinct input-output studies have appeared. Table 3.1a lists the general school effectiveness studies in chronological order by author. The table provides a frame of references for the summary of research results to permit some comparisons of data and to show the author's reliance on previous data. Table 3.1b provides detailed references for the school effectiveness studies listed chronologically on Table 3.1a.

In the three years following the Coleman Report, studies were conducted at the rate of about eight per year. This contrasts sharply to the slower pace of two or three studies per year since 1970. Notice also that, while early research generally relied on aggregate data at the school or district level, more recent work uses data on the individual student. The relative merits of the different tactics have been discussed earlier.

Another important observation from Table 3.1a is the propensity of authors to take advantage of convenient existing data sets. Ten used EEOS data and seven used the Project TALENT sample. Approximately nine authors used a questionnaire to collect at least part of their data, but most of the research was accomplished with information previously compiled for other purposes. Too frequently, researchers allowed existing data to determine the variables and relationships to be studied instead of letting the model dictate the data requirements. Critics now suspect that to be the most serious shortcoming of prior work. Garner's (1973) study is the only exception; he takes an experimental approach to estimating an educational production function.

Virtually all studies of school effectiveness have occurred at the elementary and secondary levels. The only exceptions are Astin (1968), Wells (1974b), and Polachek (1975). One would expect a marked difference in pedagogical processes and significance of different inputs between

TABLE 3.1a

A CHRONOLOGY OF SCHOOL EFFECTIVENESS STUDIES

Author(s)	Date of Study; (subseq. pubs.)*	Unit of Analysis**	Description of Sample	Source of Data
1. Mollenkopf and Melville	1956	S	9th grade in 100 schools; 10th grade in 106	Questionnaire Special tests
2. Goodman	1959	D	New York, 70,000 7th and 11th grade, male and female in 102 districts	Special data collection in 102 selected districts (QMP)
3. Thomas	1962	S	206 schools 10th and 12th grade	Proj. TALENT; Census
4. Benson, et al.	1965	D	California, 5th grade 249 districts	Census; district records
5. Coleman, et. al.	1966	I/S	645,000 students in about 3,100 schools	Huge survey by HEW/OE of nations elem & secondary schools (EEOS) in 1965
6. Kiesling	1967 (based on 1965 dissertation)	D	97 school district in New York	New York QMP data (Goodman)
7. Shaycoft	1967	I	6,500 students in 118 schools; longitudinal data on 12th graders (from grade 9)	Proj. TALENT
8. Burkhead, et. al.	1967	S	39 Chicago schools 22 Atlanta schools 180 small community high schools from Proj. TALENT	District records; Proj. TALENT

TABLE 3.1a
(continued)

Author(s)	Date of Study (subseq. pubs.)*	Unit of Analysis**	Description of Sample	Source of Data
9. Plowden Report	1967	I	Stratified random sample of primary school students in England	Central Advisory Council on Education sample
10. Katzman	1967 (Spr. 1968) (1971)	S	56 Boston elementary schools	Local records
11. Bowles	1968a (1970)	I	1,000 Black 12th graders	EEOS
12. Astin	1968	I	669 students at 224 accred. 4 yr. colleges in 1961-1965	American Council on Education questionnaire; E.T.S. (Princeton); National Merit Scholarship Corp.
13. Cohn	1968	D	377 high school districts in Iowa; 12th graders; (372 - one school)	District sources
14. Raymond	1968	D	5,000 entering at W.Va. Univ. Sept. 64-66 from W.Va. school districts	District and State records
15. Ribich	1968	D	Bottom quintile by SES 12th grade males for national sample of districts	Proj. TALENT
16. Bowles & Levin	1968b	S	12th grade Black and white students	EEOS

TABLE 3.1a
(continued)

Author(s)	Date of Study (subseq. pubs.)*	Unit of Analysis**	Description of Sample	Source of Data
17. Hanushek	1968 (1970a; 1972)	S	471 schools w/5 or more white 6th graders; 242 schools with 5 or more black 6th graders	EEOS
18. Levin	1968 (1970a)	S	Same as Hanushek (1968)	Hanushek's (1968) results (EEOS)
19. Bowles	1969	I	207 black male 12th graders (USOE Reg 1, 2 & 3); 1,000 black seniors; EEOS national sample	Proj. TALENT; EEOS
20. Fox	1969	S	39 Chicago high schools	Burkhead, et al. study
21. Kiesling	1969	D	97 N.Y. school districts	QMP data (see Goodman)
22. Guthrie et al.	1969 (1971)	I	5,284 6th grade students in Michigan	EEOS
23. Kiesling	1970	D	5th and 8th grade pupils in 86 N.Y. school districts in 1964-65	District records
24. Levin	1970b	I	597 white 6th graders in 36 schools in large Eastern city	EEOS
25. Michelson	1970	I	597 white & 458 black 6th graders in large Eastern city	EEOS

TABLE 3.1a
(continued)

Author(s)	Date of Study (subseq. pubs.)*	Unit of Analysis**	Description of Sample	Source of Data
26. Bassett and Miller	1970	I	1,000 high school and 278 elementary pupils in one Seattle area suburban district	Local records
27. Perl	1970 (1971; 1973)	I	3,265 1960 high school graduates	Proj. TALENT
28. Hanushek	1970a (1970b; 1972)	I	1,061 3rd graders in large urban California school district in 1969	District records; questionnaires
29. Averch and Kiesling	1970	S/I	5,000 9th graders at 746 high schools; 820 9th graders from above group	Proj. TALENT
30. Levin	1971 (1974)	I	597 white sixth graders in large Eastern city	EEOS (same as Levin 1970 and Michelson 1970
31. Carnoy	1971	I	182,000 students in Puerto Rico strat by sex, urban-rural and SES	Survey 1/3 school district in Puerto Rico
32. Kiesling	1971	S	42 schools strat random sample out of 700 Title I projects in Calif. in 1969-70	Survey questionnaire
33. Tuckman	1971	S	1,001 public high schools national sample	Census Bureau 1965 CPS questionnaire
34. Smith	1972	I/S	Northern 6, 9 & 12th grade subsamples from EEOS	EEOS

TABLE 3.1a
(continued)

Author(s)	Date of Study (subseq. pubs.)*	Unit of Analysis**	Description of Sample	Source of Data
35. Winkler	1972 (1975)	I	388 white and 385 black high school students from large urban district in Calif. Longitudinal records for gr. 1-8 test scores, school inputs, & peer group	Local 1964-65 records; student questionnaires
36. Garner	1973	I	62 and 48 white 8th gr. pupils in two Chicago suburb elementary schools (1970)	Survey & experimental results
37. Bieker and Anschel	1973	I	226 11th graders in five rural Kentucky high schools	Local records
38. Barnett	1974	S	School data on 7th graders in Michigan classified by district; sample stratified by geo. region and type of community (metro, city, town, rural)	Michigan Education Assessment Program 1970-71
39. Wells	1974	I	a) CAI: 446 male and female 5 & 6th graders from three schools in northern Calif. dist. b) 1,510 students enrolled in first year economics at one of 37 U's in Great Britain	a) Survey and program data b) Econ Educ. Project Survey of Heriot-Watt University

TABLE 3.1a
(continued)

Author(s)	Date of Study (subseq. pubs.)*	Unit of Analysis**	Description of Sample	Source of Data
40. Summers and Wolfe	1974	I	627 6th graders in 103 elem. schools; 553 8th graders in 42 jr. high schools; 716 12th graders in five sr. high schools in Phila. in 1970-71, 71-72	Local survey
41. Klees	1975	I	1,101 9th grade students in traditional secondary schools (Ensenanza Directa) in Mexico; 1,236 9th graders in the ITV system (Telescundaria) in Mexico. 23 and 58 classes respectively; total sample from four geor. regions near Mexico City.	Collected in 1972 by survey, questionnaire, on-site visits
42. Brown and Saks	1975	D	38 city districts 116 suburb districts 365 rural districts	Michigan Education Assessment Program 1970-71
43. Polachek et al.	1975	I	227 students in 1st semester economics at U. of N.C., Chapel Hill	Student questionnaire

TABLE 3.1a
(continued)

Notes: * Date of Study is the earliest known account of the research. Frequently, this will be the author's dissertation or a mimeographed paper. Dates appearing in parentheses are subsequent versions of the same work appearing as papers, articles, or books. All known published references are cited; but, in some cases, mimeographed papers which were largely duplicates of the original were omitted. Complete citations are given in Table 3.1b.

** Unit of Analysis is the unit of the observations in the statistical analysis; D = school district; S = school data; I = individual student.

TABLE 3.1b

REFERENCES FOR A CHRONOLOGY OF SCHOOL EFFECTIVENESS STUDIES

1. Mollenkopf, W. and S. Melville. (1956). "A Study of Secondary School Characteristics as Related to Test Scores." Princeton, New Jersey: E.T.S., (RB-56-6).

2. Goodman, S.M. (1959). The Assessment of School Quality. Albany: New York State Department of Education.

3. Thomas, J.A. (1962). "Efficiency in Education: A Study of the Relationships Between Selected Inputs and Mean Test Scores in A Sample of Senior High Schools." PhD Dissertation, Stanford University.

4. Benson, C.S. et al. (1965). "State and Local Fiscal Relationship in Public Education in California." Report of Senate Committee on Revenue and Taxation.

5. Coleman, J.S., et al. (1966). Equality of Educational Opportunity. Washington, D.C.: Health, Education and Welfare, Office of Education (OE-38001).

6. Kiesling, H. (1967). "Measuring a Local Government Service: A Study of School Districts in New York State." Review of Economics and Statistics 49 (August 1967): 356-367.

7. Shaycoft, F. (1967). The High School Years: Growth in Cognitive Skills. Pittsburgh: American Institute for Research and University of Pittsburgh, Project Talent.

8. Burkhead, J., et al. (1967) Input and Output in Large City High Schools, Syracuse, New York: Syracuse University Press.

9. Plowden Report. (1967). Children and Their Primary Schools, Great Britain.

10. Katzman, T. (1971). The Political Economy of Urban Schools, Cambridge: Harvard University Press.

11. Bowles, S. (1969). "Educational Production Functions." (Final Report) Washington, D.C.: Health, Education and Welfare/Office of Education.

12. Astin, A.W. (1968) "Undergraduate Achievement and Institutional Excellence." Science, 161 (August 16, 1968): 661-668.

13. Cohn, E. (1968). "Economies of Scale in Iowa High School Operations." Journal of Human Resources 3 (Fall 1968): 422-434.

TABLE 3.1b
(continued)

14. Raymond, R. (1968). "Determinants of the Quality of Primary and Secondary Public Education in West Virginia." Journal of Human Resources 3 (Fall 1968): 450-470.

15. Ribich, T.I. (1968). Education and Poverty, Washington, D.C.: Brookings Institute.

16. Bowles, S., and H.M. Levin. (1968b). "The Determinants of Scholastic Achievement--An Appraisal of Some Recent Evidence." Journal of Human Resources 3 (Winter 1968): 3-24.

17. Hanushek, E. (1972). Education and Race: An Analysis of Education Production Process, Lexington, Mass.: D.C. Heath Lexington Books.

18. Levin, H. (1970a). "New Model of School Effectiveness." In Mood, A.M. Do Teachers Make A Difference? Health, Education and Welfare, Office of Education (OG-58042).

19. Bowles, S. (1969). "Educational Production Functions." (Final Report) Washington, D.C." Health, Education and Welfare, Office of Education.

20. Fox, T.G. (1969). "School System Resource Use in Production of Interfependent Educational Outputs." Joint National Meeting of American Astronautical Society and Operation Research Society.

21. Kiesling, H.J. (1969). "The Relationship of School Inputs to Public School Performance in New York State." Review of Economics and Statistics 49 (August 1967): 356-367.

22. Guthrie, J.W. et al. (1971). Schools and Inequality, Cambridge, Mass.: M.I.T. Press.

23. Kiesling, J.H. (1970). "A Study of Cost and Quality of New York School Districts." Washington, D.C.: Health, Education and Welfare, Office of Education.

24. Levin, H. (1970b). "A Cost Effectiveness Analysis of Teacher Selection." Journal of Human Resources 5 (Winter 1970): 24-33.

25. Michelson, S. (1970). "The Association of Teacher Resourcefulness with Children's Characteristics." In Mood, A.M., ed., Do Teachers Make A Difference? Health, Education and Welfare, Office of Education.

26. Bassett, L.R. and R.L. Miller. (1970). "The Production of Education and Socio-economic Inputs." Institute for Economic Research Discussion Paper #127, University of Washington, Seattle.

TABLE 3.1b
(continued)

27. Perl, L.J. (1973). "Family Background, Secondary School Expenditure, and Student Ability." Journal of Human Resources(Spring 1973): 156-180.

28. Hanushek, E. (1970b). "The Value of Teachers in Teaching." (RM-6362-CC/RC), Santa Monica, California: The Rand Corporation.

29. Averch, H., and H. Kiesling. (1970). "The Relationship of School and Environment to School Performance: Some simultaneous Models for PROJ TALENT High Schools." Santa Monica, California: Rand Corporation (unpublished).

30. Levin, H. (1974). "Measuring Efficiency in Education Production." Public Finance Quarterly 2 (January 1974): 3-25.

31. Carnoy, M.G. (1971). "A Systems Approach to Evaluating Education." Stanford University, mimeo.

32. Kiesling, H. (1971). "Input and Output in California Compensatory Education Projects." R-781-CC/RL, Santa Monica, California: Rand Corporation.

33. Tuckman, H.P. (1971). "High School Inputs and Their Contribution to School Performance." Journal of Human Resources 6 (Fall 1971): 490-509.

34. Smith, M.S. (1972). "Equality of Educational Opportunity: The Basic Findings Reconsidered." In Mosteller and Moynihan, On Equality of Educational Opportunity, New York: Random House.

35. Winkler, D.R.. (1975) "Educational Achievement and School Peer Group Composition." Journal of Human Resources 10 (Spring 1975): 189-204.

36. Garner, W.T. (1973). "Identification of an Education Production Function by Experimental Means." PhD dissertation, University of Chicago.

37. Bieker, F. and K.R. Anschel. (1973). "Estimating Educational Production Functions for Rural High Schools: Some Findings." American Journal of Agricultural Economics 6 (August 1973): 515-519.

38. Barnett, W. (1974). "Production Function for Education: An Empirical Study." PhD dissertation, Michigan State University.

39. Wells, S. (1974). "Technology Efficiency and Educational Production." PhD dissertation, department of Economics, Stanford University.

TABLE 3.1b
(continued)

40. Summers, A. and Wolfe. (1974). "Equality of Educational Opportunity Quantified: A Production Function Approach," presented at the winter meetings of the Econometric Society, December 27, 1974.

41. Klees, J. (1975). "Instructional Technology and its Relationship to Quality and Equality in Education." PhD dissertation, Department of Economics, Stanford University.

42. Brown, B. and D. Saks. (1975). "Production and Distribution of Cognitive Skills Within Schools." Journal of Political Economy 83 (June 1975): 571-594.

43. Polachek, S. et al. (1975). "Educational Production Functions: A Microeconomic Analysis." Chapel Hill, North Carolina: University of North Carolina.

college and lower educational levels. Since the present study is concerned with the college level, the results of effectiveness research at lower levels will not be discussed in great detail. Instead, results for some of the popular inputs will be summarized. The reader who wishes a fuller treatment may consult one of the surveys of the literature.[1]

The predominant impression gained from a study of the literature is the inconsistency of results among the different studies. This is a consequence of the problems discussed above: collinearity in the explanatory variables combined with the variety of inputs selected by different researchers. It is also possible that different schools (or different data sets) are actually subject to different production possibilities. Or, each may desire different mixtures of cognitive and noncognitive outputs. Nevertheless, the lack of corroboration in the results is indicative of the need for further refinement of research methods.

Quantitatively, "production function studies are rarely better than 15 to 20 percent accurate, and are often far less accurate (in predicting student achievement levels)."[2] "Although production functions estimated thus far are helpful in understanding student outcomes, the amount of help they offer (in explaining performance levels or gains) is relatively small."[3]

---

[1]For a general survey, see Guthrie (1971) or Averch, et al. (1972). For effectiveness of alternative media, see Jamison, Suppes, and Wells (1974). For a thorough discussion of methodology and comparison of 98 exemplary programs for the disadvantaged, see Hawkridge (1969).

[2]Averch, et al., p. 40. The 15 to 20 figure refers to the percentage variation in the dependent variable explained by the independent variables.

[3]Ibid., p. 40.

Background Characteristics

The importance of student background or socio-economic status is the most common finding throughout the literature. Averch (1972) concluded that socio-economic status accounts for roughly 15 percent of the variance in student scores. The Coleman Report (1967) touched off a controversy over the relative importance of background characteristics and school inputs, and the interaction of the two on student achievement.[4]

A study by Bassett and Miller (1968) demonstrated that background influence as measured by fathers' occupation is an important determinant of the level of achievement but not of the change in achievement. The range of socio-economic data available varies considerably for different studies, as does the statistical procedure employed. Some selected one or two variables and discarded the rest; others included a wide range of background inputs, thereby possibly introducing multicollinearity. Still others choose to stratify their sample along SES lines by race, occupation, or urban-rural to allow for interaction effects.

Re-analysis of the Coleman results by Bowles and Levin (1968) and a thorough reworking by Smith (1972), indicate that using the stepwise process results in overemphasis of the background variables at the expense of the school inputs. However, researchers still believe background factors to be the single best predictor of student achievement at the elementary and secondary level. Studies by Bowles (1969), Kiesling (1970,

---

[4]For a full account of the debate, see Bowles and Levin, Journal of Human Resources (Winter 1968) 3-24; reply by Coleman (Spring 1968). Comments by Smith, Cain and Watts, and Bowles and Levin (Summer 1968). Also see Cain and Watts (April, 1970) 228-41; Hanushek and Kain in Mosteller and Moynihan (1972); and Smith in Mosteller and Moynihan (1972).

Levin (1970), Perl (1970), Hanushek (1970a), Winkler (1972), and Summers and Wolfe (1974), all relatively rich in SES data, substantiate this view.

Initial Aptitude, Ability, or Innate Endowments

In estimating the effectiveness of schools, it is obviously important to control for the differences in ability among students. One's abilities are developed over a lifetime as a product of genetic endowment and environment. Thus, by including variables for student background, we simultaneously account for at least part of the difference ability among students.

To an extent, some authors sidestep the problem of adjusting for ability by structuring the model to explain the value added to achievement; the achievement gain score is used as the dependent variable. But even for this case, one may argue that students with higher ability will learn faster and exhibit a greater gain. Summers and Wolfe (1974) and Bieker and Anshel (1973) both employed a value-added framework and used an I.Q. test on the right-hand side of the equation as well.

Wells (1974) chose to use post-test scores as the dependent variable and included pre-test scores plus a measure of I.Q. with the explanatory variables. Klees (1975) and Hanushek (1970) used the pre-test score from the same test series as a measure of initial achievement. Bassett and Miller (1970) and Winkler (1972) used only I.Q. tests as the measure of ability. Some of the typical instruments found in the literature are the Iowa Test of Basic Skills, Stanford Achievement Tests, Lorge-Thorndike I.Q. test, and the California Mental Maturity Test.

Ability and pre-test measures are statistically significant in all studies. In work by Wells (1974), the pre-test score was by far the dominant input and the I.Q. measure did not seem to matter. The quantitative magnitude of ability inputs was implied by the elasticity of achievement level with respect to ability, which ranged from 0.3 to 0.8. The variation over this wide range may have been caused by differences in the sample characteristics, the sensitivity of the instruments, and their correlation with background and other inputs.

School Inputs

In a recent survey of school effectiveness studies, Averch (1972) observed "School resources are seldom important determinants of student outcomes."[5] To a large extent, the quantitative importance of the school inputs hinges on the specification of the background characteristics. In the absence of adjustment for socio-economic status, measures of school inputs have been shown to account for approximately 5 percent of the variance in student scores; but if the school variables are added after controlling for student background, school impact adds only about 1 percent to our ability to predict student achievement. The difference between 1 and 5 percent stems from the correlation between school resources and student background discussed earlier.

With the great variety of data and no theory of learning to guide the selection of school inputs, considerable variation is found in the school inputs included by different authors:

> Almost every study finds one or two or three school resources to be significantly related to student outcomes. But these studies generally examine a large number of school resources. Along with the two or three resources that are found to be significant many are found to be insignificant. And, when we compare the results of various studies, we find that the same

[5]Averch et al., p. 44.

> resources do not appear among the lists of significant variables. For that matter, it is not unusual to find a research report in which the students have been divided into a number of groups by some stratification rule, with separate analyses yielding distinctly different results with respect to the significance of school resources for each group.[6]

Some studies have used school budget figures as inputs. Total expenditure per student has occasionally been a significant input for such outputs as dropout rate or percent going on to college, but it generally has not had an influence on student achievement. Administrative expenditure per pupil and the value of school property have at times been included with no apparent significance. The only financial measure with a consistently strong impact on student achievement has been average teacher salaries or starting salaries. This was demonstrated in earlier studies by Thomas (1962), Kiesling (1969), Cohn (1968), Burkhead (1967), and Benson (1965), and in more recent work by Perl (1973) and Winkler (1975).

Since the major portion of the educational budget goes toward purchasing the most and best instruction possible, we are especially interested in the relationship between class size and such teacher characteristics as years of experience and degree level, and how they affect the level of student achievement. Using pupil teacher ratios as a proxy for class size, several early studies concluded that class size within a "normal range" does not influence achievement. Bowles and Levin (1968) pointed out that faculty teaching loads vary from four to six hours per day, which implies as much as a 50 percent variation in class size for

---

[6]Ibid., p. 45.

schools with identical pupil-teacher ratios. More recently, Averch and Kiesling (1970) found a significant negative influence of class size on pupil achievement and a positive impact on teacher transfers. Klees (1975) discovered class size to be a relevant factor in the public educational system of Mexico. Summers and Wolfe (1974) found low achievers to be the only group adversely affected by large classes.

A variety of teacher characteristics have been examined in the literature. Probably the most puzzling input has been years of experience. The evidence is about evenly split on its relevance to student performance. Frequently, when a researcher has stratified his sample, years of experience will be significant in some equations but not in others. Furthermore, no consistent pattern has emerged. It appears relevant in recent work by Levin (1971), Wells (1974), and Klees (1975), but research by Perl (1973) and Hanushek (1970) revealed teacher experience to have a negative coefficient. One finding of Summers and Wolfe (1974) offers an explanation:

> High achieving pupils do best with more experienced teachers, but these teachers lower the learning growth of low achievers--these students do best with new, relatively inexperienced teachers, who perhaps, have an undampened enthusiasm for teaching those who find it hard to learn.[7]

Thus, average years of teacher experience may frequently appear insignificant or even negative with respect to the average student.

In contrast to the ambiguous results obtained for teacher experience, fairly potent and consistent results have been achieved with direct measures of teacher quality. In Bowles (1969, 1970), Hanushek (1972), and

---

[7]Summers and Wolfe, p. 12.

Guthrie (1971), the teacher's score on a brief verbal test was usually significant and the often the single most important school input.

Years of teacher graduate training or percent with MA or percent certified have not distinguished themselves as important factors in student achievement.[8] This is a notable result given the fact that teacher salary scales are based in part on these criteria.

The quality of the teacher's undergraduate college is one measure not widely employed by the authors. But, whenever included, it yielded valid results.[9] Undergraduate school quality is usually measured by type of institution (university, college, teacher's college) or by the school Gourman rating.[10] "It seems clear that, in some segments of the school system, teachers from colleges with higher ratings are more effective teachers."[11]

In retrospect, these results are not necessarily as inconsistent as they at first appear. It may be that school resources act differently on different people. Probably the single most notable characteristic of a study by Summers and Wolfe (1974) is the importance the authors place on interaction effects. They feel many education inputs appear unimportant because most research has looked at the effect of particular inputs on the "average" student. If many of the inputs act in different ways on different people, benefiting some and hindering others, they will

---

[8]With the exception of Wells (1974).

[9]See Levin (1970, 1971), Michelson (1970), and Summers and Wolfe (1974).

[10]Gourman (1967).

[11]Summers and Wolfe, p. 13.

appear to be insignificant on the average. Many of Summers' and Wolfe's findings support that hypothesis.

Brown and Saks (1975) go a step beyond Summers and Wolfe (1974) by arguing that school inputs may be just as significant for their effect on the distribution of student outcomes as for their effect on mean performance. Their research is the first attempt to build preferences of the school into the model of the production process. With "welfare" defined in terms of the mean and standard deviation of a composite achievement score, Brown and Saks discover that teacher experience (in years), degree level, and student-teacher ratios are significant determinants of the standard deviation in some cases. One wishes that the results had been more robust and that a wider selection of school inputs had been available. Although their innovation is an advancement in educational model building, they concede that the added complexity cannot reveal anything about the productivity of inputs unless one is able to independently estimate the parameters of the objective function. The magnitude of the factor coefficients may be indicative of the input's productivity, or simply a result of school preferences.

The research method employed by Garner (1973) deserves special attention. Garner adopted an experimental procedure whereby data were generated to estimate an educational production function for a learning mastery model. Instead of achievement level, the dependent variable was the amount of time required for a student to reach a specified level of mastery in a particular subject. "An educational production function based on observations of individual students in a controlled setting was estimated and found to explain more than 60 percent of the variance in in student time to criterion performance levels, at a high level of statistical significance."[12]

Do school resources make a difference? Some do, but frequently not those directly attributable to school expenditures. The lack of consensus in the literature on the importance of school inputs is probably symptomatic of the significant interaction effects and poor research methodology. To correct past deficiencies, researchers must continue efforts to model multiple outputs, investigate the plentiful interaction effects, and endeavor to develop appropriate experimental data, rather than rely on remote proxies for school and teacher quality.

School Effectiveness Studies at the College Level

As mentioned earlier, almost all production function studies have been conducted at the elementary and secondary levels. This must result in part from the increased complexity of educational objectives at the college level. Research and community service become important in college, and cognitive outputs cover a broader range of topics. Thus, measuring outputs is more difficult for higher education than at lower levels, expecially when the goal is to compare the quality of different schools. "Of the more than 1,000 studies of college impact recently reviewed by Feldman and Newcomb (1969), only a handful used measures of cognitive outcomes and, of these, virtually all used psychological rather than behavioral (test) measures."[13] Alternatively, college is typically viewed as an investment, and its effectiveness is evaluated by the job

---

[12]Garner (1973), p. 121.

[13]Astin (1973), p. 113.

market success and augmented income profiles of its graduates.[14] However, since this approach is not suited for exploring the impact of different educational resource allocations, input-output research in higher education is still sorely needed.[15]

Probably the closest thing to a general production study was conducted by Astin (1968). Using 669 student GRE exams in three subject areas as the output measure, Astin first controlled for student ability, SES, motivation, and school environment. He then examined the relevance of eight measures of school quality: ability level of students (selectivity), expenditures per student, academic competitiveness, library size, books per student, faculty-student ratios, percent faculty with PhD, and a composite of the above. Student ability, as measured first by the NMSC test scores and second by the high school GPA, was the single most important determinant of student achievement in college. No single measure of institutional quality seemed to have an effect on student achievement: "There is apparently no 'value added' from attending a highly selective institution, at least with respect to cognitive performance as measured by the GRE."[16] There is some evidence that differences in college quality exert an effect on aspects of college output other than achievement scores such as dropout rates, satisfaction with later

---

[14]See L. Solomon, and B. Chiswick, in Solomon and Taubman (1973). For additional references, see the bibliography to these articles.

[15]For a discussion of problems and review of the meager research completed in this area, see Astin, in Solomon and Taubman (1973). Also, Walhaus (1975), and Hanushek (1975), in New Directions for Institutional Research series (Winter, 1975).

[16]Astin (1968), p. 124.

life, and income returns over the life cycle. But, with respect to student achievement levels, we are faced with the familiar cliche: Differences in college 'quality' have not been found to exert any independent influence on the performance of students.

The remainder of effectiveness studies in higher education have investigated the production of knowledge in a specific course. As one of five separate empirical studies in his dissertation, Wells (1974) analyzed the production of joint economics knowledge in micro-, macro-, and international economics, and student attitudes on the usefulness of economics for a sample of 1,510 students at 27 universities in Great Britain. Because of the number of models estimated, it is difficult to summarize the results. For example, the first set of tables presents estimates for both linear and Cobb-Douglas forms by two statistical techniques (OLS and unrestricted stepwise) for each of five dependent variables, total of 40 regressions with 44 independent variables in each. Several other separate analyses were performed, including determinants of student attendance and study and determinants of school resource distribution. The variety of data is particularly rich, and as one might expect, particular inputs rarely behaved consistently in all regressions. Nevertheless, it is of interest to point out some of the especially robust inputs as well as those that were conspicuous for their lack of influence.

By far the most important input was the pre-test measure of initial understanding in economics. Taking the advanced course in economics in high school also resulted in better college performance by some. Background factors were poorly measured because no socio-economic data were available. However, student I.Q. as measured on two intelligence tests probably picks up part of the SES influence as well as the importance of ability. The control for the type of high school attended can also be a surrogate for background factors. Of the wide range of educational process variables, only the percent of lectures attended was important. Of the university variables, three were significant: the number of pupils enrolled in economics, the school size, and the average intelligence of the student body. Some of the variables that would be expected to influence economics education but that failed to do so include the student's major field, age, SES status, and years of experience of lecturers and tutors, course balance between macro and micro, the textbook, the hours spent studying per week, and such attitude measures as course rating and lecturer ratings. Research by Wells (1974) on the effectiveness of technology in education will be reported in the next chapter.

Other efforts have also been made to study the resource effectiveness within schools. A study by Polachek, Kniesner, and Harwood (1975) used grade as the output to investigate the allocation of three resources (ability, study time, and lecture attendance) by 227 students taking Macroeconomic Principles in the spring of 1975 at the University of North Carolina at Chapel Hill. Nonlinear maximum-likelihood techniques were applied to estimate a constant partial elasticity of substitution (CPES) production function. Although this functional form does not permit the elasticities of substitution between factors to vary, its advantage is that it allows the estimation process to select the factor substitutability ratio. The marginal products of SATM, one hour of lecture, and one hour of study were 0.05, 0.3 and 0.3, respectively. Students appear to equalize the marginal productivity of time spent in class and study outside of class. Furthermore, the average student may overcome the advantage of his "smarter" classmate by trading 15 hours of extra study for each 100 points yielded on the SATM. The use of grade as the output measure must certainly restrict this approach to a limited range of resources within a given class. Nevertheless, the methodology and estimation technique of Polachek, et al. (1975) provides an interesting contribution to the literature.

Numerous other studies have been conducted within schools for a particular course. However, each is usually interested in some particular input such as the effectiveness of graduate assistants, teachers with advanced degrees, teachers with high scores on an attitude or psychological test, or some educational innovation like programmed learning, television, or computer-aided instruction. These studies can only be considered research on the educational production process depending on the amount of control for the various intervening inputs.

Results on the effectiveness of television and computer-aided instruction are reviewed in the next chapter, following a discussion of their use in higher education.

CHAPTER FOUR

TECHNOLOGY AND EDUCATION

Technology in Higher Education: Definition, Utilization, and Benefits

In the last 20 years, a great deal of attention has focused on innovations in education. In 1967, Eric Ashby predicted a Fourth Revolution[1] in education over the next several decades, with communications technology becoming an indispensable tool equal in importance to the part played by books at present. This chapter will outline the extent and use of technology in higher education and review the existing literature on the effectiveness of ITV and CAI technology.

Although the empirical work in this thesis concerns two rather specific technologies, it is important to be aware of the available alternative technologies. Media-related educational technologies fall into three general categories: less complex technologies, televised forms, and computerized instruction. Table 4.1 presents a more detailed breakdown of alternative configurations.

The less complex technologies vary widely in their capabilities and costs--from the economical film strips and slides to 16mm sound movies. Instructional television may be either taped or live, with or without

---

[1]See The Fourth Revolution (N.Y., McGraw Hill, 1972), Carnegie Commission on Higher Education. The other three revolutions were (a) task of education shifted from parents to schools; (b) acceptance of written word as tool of education in the classroom, and (c) with printing press and movable type, widespread availability of books.

TABLE 4.1

MEDIA BASED TECHNOLOGIES

I. Less Complex Technologies	a. Film b. Recordings (tape and cassette) c. Sound motion pictures d. Instructional radio
II. Instructional Television	a. Videotaped and Videocassette TV b. Closed Circuit TV c. Instructional Television Fixed Service (ITFS) d. Point to point microwave TV e. Broadcast TV f. TV by phone line g. Cable TV
III. Computer Assisted Instruction	a. Small scale system for data processing, problem solving, simulations, and games b. Large scale interactive networks

interactive talkback between student and instructor. It may be a basic "talking head," typical classroom lecture by TV or a complicated production. Either taped or live presentations could be transmitted by any of the configurations listed "b" through "g". Each differs primarily with respect to the mode by which the signal is transmitted (cable, broadcase, microwave, phone line, etc.), and each involves different hardware and system costs. The engineering characteristics give rise to different operational capabilities in terms of the number of channels, the geographical ranges, the type and number of users served and educational services delivered.

Computer-assisted instruction technologies are simply broken down according to scale. However, the large-scale CAI could be further categorized by mode of delivery (cable, leased phone line, microwave, or satellite).

How widespread is the use of ITV and CAI at the college level? A NEA survey of closed-circuit television in 1967 revealed its use in over 200 schools. In addition to these on-campus systems, a recent survey by Wong (1974) lists approximately 20 TV networks directed to off-campus audiences. Table 4.2 from the NEA closed-circuit study gives a picture of the distribution of ITV by subject area. TV is used with the highest frequency in education, speech, drama, and life science areas.

Of a total 2,477 schools, only 978, or 39 percent, reported access to computers in 1966-67. The 59 percent without computers enrolled about 1.7 million, or 25 percent of the college students. "It has been estimated that the figure was 1,100 in 1968, 1,255 in 1969--44 and 50 percent. Even with these increases, about half of the nation's colleges and

TABLE 4.2

USE OF CLOSED CIRCUIT TELEVISION IN HIGHER EDUCATION
BY SUBJECT AREA

Field of Study	Number of Institutions Reporting Use		
	Lower Division	Upper Division	Graduate
Agriculture	9	2	2
Business Administration	36	23	7
Economics	28	8	1
Education	73	129	78
Engineering	33	32	18
English	71	36	5
Fine Arts	63	41	13
Foreign language	34	17	5
History	43	20	2
Humanities	47	22	9
Health/physical education	62	32	5
Life Sciences	83	36	14
Mathematics	44	14	7
Military training	11	12	3
Political science	29	16	4
Physical science	71	30	7
Social science	95	49	24
Speech and drama	123	96	28
Technical and vocational	38	20	11
Professional			
Dental			47
Law		1	6
Medical			63
Nursing	1	1	58
Theology			14

Source: A Survey of Institutional Closed Circuit Television, National Education Association, 1967, pp. 34-5.

universities were entirely without computer service as the 1970s began."[2] Table 4.3 shows the distribution of computers among public and private schools by size and type of institution. Notice that the larger the school and the higher the degree offered, the more likely that a computer is present.

It is also interesting to see how the breakdown of computer activity differs by class of school. These divisions, expressed in percentages, are shown in Table 4.4. The instructional share ranges from about 7 percent to a full 100 percent. The use of computers for instruction is heavier at larger schools and at schools below the doctorate level. On a national average, about 30 percent of computer utilization is for instruction.

Table 4.5 shows the division of computer instruction by academic field in 1966-67. These figures demonstrate how heavily computer instruction is centered in engineering, computer science, and business. Unlike ITV, which almost always presents instruction as a substitute for on-site lectures, computerized instruction encompasses several activities. George Comstock (1972)[3] classifies computer instruction by five categories:

1. Data processing and computer science: the teaching of computer skills.

2. Student problem-solving and research: teaching about the computer as a tool for use in some field outside computer science.

3. Tutorial: use of computer as a medium to present instruction directly to the student.

4. Simulations, demonstrations, and games: use of the computer to simulate, in part, social and physical phenomena.

---

[2]Levien (1972), p. 143.

[3]Page 201.

TABLE 4.3

COMPUTER USE BY INSTITUTIONS OF HIGHER EDUCATION, 1966-67*

Degree Level and Enrollment	Public Institutions Using Computers			Private Institutions Using Computers		
	Total	Number	Percent	Total	Number	Percent
Associate						
Below 500	100	10	10	187	5	3
500-2,499	256	120	47	87	15	17
2,500-9,999	116	80	69	4	2	50
10,000-19,999	23	19	81	**		
Over 20,000	**			**		
Bachelor						
Below 500	10	1	14	247	8	3
500-2,499	65	31	48	461	112	24
2,500-9,999	24	21	86	22	15	69
10,000-19,999	1			1	1	100
Over 20,000	**			**		
Master						
Below 500	4	2	50	97	8	8
500-2,499	40	17	41	156	63	40
2,500-9,999	133	116	87	58	46	80
10,000-19,999	8	8	100	2	2	100
Over 20,000	10	10	100	**		
Doctorate						
Below 500	32	6	18	38	4	11
500-2,499	53	38	72	48	29	61
2,500-9,999	49	47	96	45	43	94
10,000-19,999	53	52	98	17	17	100
Over 20,000	24	24	100	6	6	100

* Source: G.A. Comstock, "National Utilization of Computers," in Levien (1972) p. 143.

** Indicates no schools in category; other absent data indicates information unavailable.

TABLE 4.4

USE: THE SHARE FOR MAJOR USES OF COMPUTERS BY CLASS OF SCHOOL, 1966-67*
(in percent)

Degree Level and Enrollment	Public Institution				Private Institution			
	R**	A	I	O	R	A	I	O
<u>Associate</u>								
Below 500		30.0	70.0		10.0	20.0	70.0	
500-2,499	3.9	28.4	63.2	4.4	6.7	53.3	33.3	6.7
2,500-9,999	3.2	37.6	57.5	1.6			100.0	
10,000-19,999	1.8	30.9	67.3		***			
Over 20,000	***				***			
<u>Bachelor</u>								
Below 500	40.0	40.0	20.0		26.3	47.4	26.3	
500-2,499	8.8	30.0	53.8	7.6	12.1	34.0	52.1	1.9
2,500-9,999	15.1	36.5	39.7	8.7	8.0	56.0	36.0	
10,000-19,999					20.0	20.0	60.0	
Over 20,000	***				***			
<u>Master</u>								
Below 500					40.0	50.0	10.0	
500-2,499	17.1	45.7	31.4	5.7	16.2	39.6	39.1	5.1
2,500-9,999	14.5	43.4	41.1	1.0	14.2	40.8	28.5	6.6
10,000-19,999	18.2	34.8	47.0		13.3	33.3	40.0	13.3
Over 20,000	19.7	32.9	43.4	3.9	***			
<u>Doctorate</u>								
Below 500	73.3	20.0	6.7		50.0	25.0	25.0	
500-2,499	55.7	22.6	21.0	0.6	48.7	24.6	25.6	1.0
2,500-9,999	36.6	28.2	30.1	5.1	51.7	22.5	21.8	4.0
10,000-19,999	41.6	28.6	28.1	1.7	54.7	26.3	16.5	2.6
Over 20,000	62.7	17.0	18.8	1.3	36.4	38.2	25.5	

* Source: <u>Ibid.</u>, p. 154.

**Note: R = research; A = administration; I = instruction; O = other use by noneducational institutions.

***No schools in category; other absent data unavailable.

TABLE 4.5

DISTRIBUTION OF STUDENTS AND EXPENDITURES FOR INSTRUCTIONAL COMPUTER USE BY ACADEMIC FIELD, 1966-67
(in percent)

Field	Level of Instruction					
	Undergraduates		Graduates		Total	
	Students	Expend.	Students	Expend.	Students	Expend.
Engineering	35.2	23.9	33.0	25.7	34.9	24.5
Computer Science	23.2	28.5	16.3	14.6	22.2	24.0
Business and Commerce	26.0	20.3	2.6	13.2	22.6	18.0
Mathematics	2.2	15.3	12.8	7.0	3.7	12.6
Physical Science	4.8	3.7	8.2	16.3	5.3	7.8
Social Sciences	2.6	2.2	7.4	4.8	3.3	3.0
Psychology	1.2	1.8	3.2	4.2	1.5	2.6
Education	0.9	1.3	6.8	3.9	1.8	2.2
Agriculture and Forestry	1.3	0.9	3.6	3.8	1.7	1.8
Biological Sciences	0.8	0.7	1.3	2.8	0.9	1.4
Health Professions	0.1	0.2	3.1	2.4	0.6	1.0
Humanities	0.4	0.5	0.5	0.5	0.4	0.5
Military Science	0.1	0.3			0.1	0.2
Architecture	0.3	0.1	0.6	0.3	0.4	0.2
English and Journalism	0.1	0.1	0.4	0.1	0.1	0.1
Law	0.0	0.0	0.01	0.2	0.0	0.1
Home Economics	0.6	0.0	0.03	0.1	0.5	0.05

Source: *Ibid.*, p. 178.

5. Teacher's aide: use of the computer to assist the teacher in recording grades, attendance, assignments, givine exams.

In a survey of computer use at 200 college and universities in California, Comstock (1972) found processing and computer science to be the most common areas of instructional use (68 percent), followed by teacher's aid and student problem solving (50 percent), and research (45 percent). Only 10 percent of the institutions with computers reported tutorial use.

This thesis concerns ITV and CAI as alternatives to traditional classroom instruction. Course instruction at the college level by ITV and CAI is probably more widespread than the layman would expect, and it is undoubtedly more widespread now than the figures and tables from 1967 show.

What does technology-based instruction have to offer the university or college? There are many answers to this question. What ITV or CAI can or cannot do fundamentally depends on the approach and effort behind course and software development. Hardware engineering characteristics are important and ought to be viewed as constraints on the system, limiting its range, capacity, computer response time, picture quality, and the like. But ultimately, if technology is destined to make a substantial impact on higher education, it will be the courseware that makes the difference.

With the emerging importance of accountability in education, innovation in education must be justified by its effect on educational benefits[4]

[4]Technically, according to its effect on all educational outcomes. However, because of the measurement problems discussed earlier, benefits, as used here, refers only to those cognitive outputs that can be measured at present.

and costs. One benefit for ITV and CAI is the potential for net gain in teaching effectiveness. Educational technology is effective if the student performs at least as well on an achievement test as he would have by traditional instruction. If it results in a savings of student or instructor time, or other indirect costs, or a net gain in achievement, then the technology is more effective.

Another advantage, especially with CAI, is individualized instruction and increased flexibility. Classes would no longer have to proceed in a lockstep fashion at the same speed. Each student could advance through course material at his own pace and have the flexibility to schedule lessons, when, and to some extent, where, he or she chooses. One would expect that some subject areas would be more suited to technological delivery than others. This matter must be answered by well-controlled research.

If it turns out that the introduction of technology causes no significant difference in effectiveness (student performance), another possible source of attraction could stem from its effect in reducing the costs of education. This is discussed at length in Chapter 7. At this point, it's enough to note that the prospects are not as far-fetched as one might think. Although communications hardware and software development costs are enormous, the unit cost of education can be lowered to tolerable levels by spreading the fixed costs among a sufficient number of users. Moreover, while the costs of teachers and other labor--which constitute 70 percent of the traditional education dollar--may continue to rise, the unit cost of communication services are expected to fall over the next few decades. Therefore, under certain conditions, technological innovation in education may bring about cost savings.

ITV widens the range of courses available to students as a result of course sharing between campuses. The courses may be either taped or live. Frequently, an upper-level course will be taught live to several campuses simultaneously in a metropolitan area when sufficient course enrollment might not be generated on any one campus. This enables students from one school to take courses from especially talented faculty at other schools. In addition, the savings in transportation costs are obvious.

Technology in education suggests possibilities for another form of sharing--between the campus and noncampus populations. When ITV or CAI involves a network extending beyond the campus, the potential benefits expand in two dimensions. First, the instructional material reaches a nontraditional audience: people off campus who, for one reason or another, do not have access to on-campus instruction. This opens possibilities for providing college education to people living in rural areas, prison inmates, industrial workers, handicapped individuals, and others. Second, a communications network originally justified on educational grounds may be suited to the delivery of numerous other services, including virtually any two-way information transfer to the home. Examples of these services are given in Table 4.6.

As you can see from the table, the evolution of educational methods along technological paths may bring about other externalities or secondary benefits with significant social, political, and economic implications.

<u>Effectiveness of Technology in Education</u>

Most evaluation studies define the effectiveness of technology by its impact on student achievement as measured by some type of objective

TABLE 4.6

PROJECTED 1989 MARKET FOR SELECTED TWO-WAY HOME INFORMATION SERVICES

	Services	Dollar Value (millions)
1.	Plays and movies from a video library	2,829
2.	Computer-aided school instruction	2,047
3.	Cashless-society transactions	1,810
4.	Person-to-person (paid work at home)	1,713
5.	Computer tutor	1,414
6.	Adult evening courses on TV	1,131
7.	Correspondence School	943
8.	Dedicated newspaper	849
9.	Answering services	743
10.	Computer assisted meetings	707
11.	Household mail and messages	707
12.	Secretarial assistance	707
13.	Shopping transactions	584
14.	Banking services	566
15.	Grocery price list, information and ordering	566
16.	Special sales information	354
17.	Consumer's advisory service	354
18.	Daily calendar and appointment reminder	292
19.	Legal information	285
20.	Weather bureau	228
21.	Newspaper, electronic, general	200
22.	Past and forthcoming events	130
23.	Fares and ticket reservations	124
24.	Message recording	106
25.	Index, all services served by home terminal	106
26.	Library access	95
27.	Bus, train, and air schedulings	79
28.	Access to company files	46
29.	Restaurants	35

Source: P. Baran, "Potential Market for Two-Way Information Services to the Home, 1970-90." Institute for the Future (R-26), December 1971, p. 121.

test. Typically, the results of a media-instructed group are compared with those from a "control" lecture group. Many studies simply test for a significant difference between the sample means; a few attempt to measure and adjust for the influence of other intervening factors. Summarizing the literature in this area raises problems similar to those encountered in the previous chapter's discussion of school effectiveness literature. There are far too many studies to discuss individually, and, for several reasons, they are not easily generalized.

Studies have been conducted at all educational levels: elementary, secondary, and college.[5] Within the ITV and CAI groups, there is considerable variation in system capabilities as well as subject matter and how and to what extent the technology is used. These differences have the effect of generating many categories of mutually incomparable research results. Consequently, a discussion of research results is cumbersome and few valid generalizations are possible. With these limitations in mind, the following two sections summarize the research on the effectiveness of ITV and CAI.

Effectiveness of ITV

Several years ago, two independent surveys were conducted on the effectiveness of televised learning: Chu and Schramm, Learning from Television: What the Research Says (1967), and Dubin and Hedley, The Medium May be Related to the Message: College Instruction by ITV (1969). The literature search connected with the present research did not disclose

[5]For a survey of the literature on the effectiveness of traditional instruction, instructional radio, programmed instruction, ITV, and CAI, see Jamison, Suppes, and Wells (1974).

any findings that contradict the general results of those two comprehensive surveys.

Researchers attempting to generalize the literature have come to the same conclusion: ITV can be expected to result in student test scores at least as high as those resulting from traditional instruction. Chu and Schramm (1967) found evidence that television instruction is likely to be more effective at elementary and secondary levels. That result is borne out in Table 4.7, which reproduces some of their "case count" results. There are several potential explanations for this phenomenon. Younger children may be more receptive to the television medium since it has played a larger role in their experience than it has for older children. Furthermore, a TV teacher may hold special prestige for young pupils and thus be in a better position to motivate and stimulate learning. The higher the educational level, the more difficult the material and the more reluctant the teacher to surrender this role to outside control. Finally, the deficiencies of the system, such as lack of inadequate interaction and feedback capabilities, may be more serious with complex subject matter at the college level. Out of 202 comparisons at the college level, 14 percent favored traditional instruction and 11 percent favored TV.

In other evidence reproduced from Chu and Schramm (1967) Table 4.8 reveals some unexpected evidence on relative effectiveness for different subjects. While some subjects are more easily adaptable to television, we can, on the average, expect students to learn almost any subject by television with equal effectiveness. The survey two years later by Dubin and Headley reached the same conclusions.

From 79 studies at the college level, Dubin and Hedley (1969) based their evaluations on 348 comparisons between an "experimental" and

TABLE 4.7

RESULTS OF 421 COMPARISONS BETWEEN ITV and TI[1]

Level	Number of cases of[2]		
	No significant difference	ITV more effective	TI more effective
Elementary	50	10	4
Secondary	82	24	16
College	152	22	28
Adult	24	7	2
	308	63	50

1. Source: G.C. Chu and W. Schramm, Learning from Television: What the Research Says (Washington D.C.: National Association of Education Broadcasters, 1967), p. 7.

2. This Table reports 421 comparisons of ITV with TI from 207 separate studies.

TABLE 4.8

RELATIVE EFFECTIVENESS OF ITV AND TI, BY SUBJECT MATTER*

Subject	Number of comparisons	Percentage of comparisons in which ITV did as well or better than TI
Mathematics	56	89.2
Science	100	86.0
Social studies	77	89.6
Humanities	45	95.5
Languages	77	88.3
Skills	26	96.1
Miscellaneous	40	75.0

* Source: Ibid., p. 10.

"control" class. An overall comparison of ITV and TI slightly favored face-to-face instruction. However, examining the separate effects of systems with different configurations overturns face-to-face advantage. When the method of instruction (i.e., lecture) was held constant, one-way television was demonstrated to be as good as face-to-face instruction and a two-way interactive system was overwhelmingly inferior. The idea, of course, is that two-way communication should improve ITV by approximating a live lecturing situation. For the majority of the two-way systems in this sample, these expectations were not fulfilled. Dubin and Hedley (1969) offer two explanations: (1) technical difficulties with the equipment, and (2) the reluctance or inability of students and professors to adapt to the procedures involved.

A brief digression is warranted at this point to draw a clear distinction between media and methods. Dubin and Hedley (1969) define the instructional medium as a "total configuration of the technology and interaction between teaching and learning."[6] Differences in system configurations obviously distinguish TI from TV and CAI. "A teaching method on the other hand, is a recognizable procedure employing a given medium of instruction." "It is essential to insure that possible variations in teaching method do not confound the media comparison."[7] The point that teaching method should be held constant to examine the effect of different technologies is well taken. This is a straightforward and unambiguous requirement as long as one adopts gross definitions of "method," as in

---

[6]Dubin and Hedley, p. 3.

[7]<u>Ibid</u>.

the case of Dubin and Hedley's survey, which classifies teaching method as either lecture, discussion, or demonstration. Needless to say, there are different possible teaching methods within each of these categories and the meaning of this condition or requirement is not as clear-cut as it first appears.

The deeper one explores the educational television literature, the more one becomes aware of the importance of teaching method and the impact on effectiveness of fairly small, subtle changes in method. In the polar case, television may be used to simply reproduce exactly what would normally happen in the lecture hall:

> The camera picks up a professor lecturing as he would "live". The only dimension added by television broadcasting is that the audience becomes unlimited in size, unrestricted in location, and scheduled for listening at any time a video tape may be broadcase of the original live performance.
>
> By contrast, in instructional television, an attempt is made to design the performance to use the visual potentialities of television as an integral part of the instructional process. Instructional television may, therefore, significantly modify the performance aspects of lecturing, and produce a learning retention effect, measurable on examinations, that reflects the combined audio and visual impacts of the instruction's performance.[8]

Jamison, et al., point out that

> When highly stringent controls are imposed on a study [such as in the polar case cited above], the nature of the controls tends to force the methods of presentation into such similar formats that one can only expect the 'no significant differences' that are in fact found. When ITV is used in a way that takes advantage of the potential the medium offers-- we would expect more cases of significant differences between the experimental group and the 'alternative treatment' group.[9]

---

[8]Dubin and Hedley, p. 12.

[9]Jamison, et al., p. 28 (1974).

The experimental-control model is typical of the technology effectiveness literature, and, usually, such "stringent controls" are imposed.

In view of this situation, the present research employs a rather loose condition for holding the "teaching method" constant. Instead of an experimental-control group format, the sample is composed of different schools employing different technologies. As a result, a comparison of the media encompasses more than just the communication equipment's ability to reproduce a situation on tape to different audiences, locations, and times. It also includes adjustments in teaching or delivery techniques, plus adjustments in other inputs--such as discussion sections, course outlines, or textbooks--that are complementary to the basic instruction. It is hoped these adjustments contribute to the effectiveness of the media. The complementary components of instruction and other inputs are, in fact, an element in the definition of the technology medium.

In recent research, Klees (1975) used the production function approach to compare the effectiveness of the rural ITV system of secondary education in Mexico (Telesecundaria) with the traditional school system (Ensenanza Directa) for ninth grade achievement in mathematics and Spanish. After reporting results for the total sample and for partitions by sex and ability levels, Klees concluded that the treatment variable was significant in all regressions at the 1 percent level. However, the coefficient magnitudes varied widely in the different regressions. Part of the problem stemmed from other differences between the two programs that accounted for the significant shift variable.

In another TV study, Wells (1974) found no significant differences attributable to system characteristics. In one case, a production function was estimated for 26 MBA students studying economics off campus by closed circuit TV at Golden Gate University. No particular input appeared significant in any of the five models estimated, and since there were no non-ITV observations, it was impossible to isolate the effects of technology. For another sample of 180 students enrolled in a graduate microeconomics course at Stanford University, TV was used to broadcase live lectures into an adjacent classroom to the overflow from the lecture hall. Students had the option of selecting the lecture or the TV room, and there was no effort to control the amount of TV each student received. As a result, almost half (43.8 percent) never saw a TV lecture, and the average student attended 43 times as many lectures in the lecture hall as the overflow room. It is hardly surprising that the three technology variables failed to be significant in all six models.

Before going on to discuss the effectiveness of CAI, a comment should be made about attitudes toward ITV. Dubin and Hedley (1969) reported that students have a considerably more favorable attitude toward ITV after they have experienced it than before. There is also evidence of interaction between class size and attitude. Students are likely to opt for TI over ITV except in the case of large lecture classes, where "typically over half of the students preferred ITV."[10] Chu and Schramm (1967) note that "Liking ITV is not always correlated with learning from it."[11] Their conclusion coincides rather well with this author's own informal observations and conversations with students during the data collection phase of the research for this thesis.

---

[10] Dubin and Hedley, p. 80.

[11] Chu and Schramm, p. 67.

Effectiveness of CAI

As noted earlier, the primary instructional use of CAI centers around drill and practice or games and simulations--some type of supplement to regular class instruction. Only in the last four years has CAI reached the developmental stage where entire courses have been committed to the system and teacher contact becomes voluntary through scheduled discussion sessions or informal office visits. Since it is such a recent development, research evaluations of CAI courses are limited to working papers and conference proceedings.

Jamison, et al. (1974) provides a discussion of the CAI effectiveness literature. As with ITV, the experimental-control group method predominates. Most research makes no attempt to provide data for complete productivity analysis. Numerous evaluations of drill and practice CAI programs at the elementary and secondary level have been conducted in mathematics [Suppes and Morningstar (1969), and Wells, Whelchel, and Jamison (1975)] and reading [Atkinson (1968), and Fletcher and Atkinson (1972)]. Generally, strong evidence exists that TI augmented by CAI results in a higher level of achievement compared with traditional methods alone.

Wells (1974) used the production function approach to study the effectiveness of CAI for 446 fifth and sixth grade students in three Northern California schools. CAI was administered in two of the three schools. The score in mathematics on the California Test of Basic Skills (CTBS) was used as the output measure. With the sample stratified by grade and sex, regression results indicated that the CTBS pre-test score, the number of CAI sessions during the year, and an attitude score representing self-image were significant in that order (usually at the 5 percent level or better). Variables representing teacher verbal score, experience, and degree level did not appear important. The coefficient on CAI sessions was 0.004:

> If this coefficient is valid over the entire range of the variable as the use of the linear model assumes then the student with 100 sessions will have increased his post-test score by .4 grade equivalents. This is an impressive gain considering that 100 sessions represents the usage of the computer on an average of slightly over every other day for a 10 minute session.[12]

It was not possible to determine the effect of the distribution of CAI sessions on student achievement.

Beyond the potential gains in achievement, even larger improvements in efficiency may be possible through student and faculty time savings made possible by CAI. In addition, CAI has such desirable features as self-paced instruction and system flexibility that are not present in other modes of delivery.

In the early 1970s, substantial CAI development occurred at several universities in the areas of physics, chemistry, foreign languages, accounting, nursing, library science, mathematics and economics. Several effectiveness studies[13] of college-level CAI have been undertaken and many more are undoubtedly underway. Most employ the experimental-control group methodology. The following general conclusion is strongly supported by evidence accumulated through 1973: no significant differences were

---

[12]Wells (1974), p. 179.

[13]See Hansen, Dick and Lippert (1968) and Lawler (1971) all of Florida State U.; Adams (1969), Adams and Morrison (1969), all of SUNY Stony Brook; Grandey (1974), Axeen (1967), McKeown (1974), Bitzer and Boudreaux (1969), Coombs and Peters (1971), all of U. of Ill.; Castleberry and Lagowski (1970), Homeyer (1970), Judd, Bunderson and Bessent (1970), all of U.T. at Austin; Suppes and Morningstar (1969) at Stanford.

noted in the performance of students in the two groups. In some cases, the CAI group was found to reach the same level of competence as the TI group in substantially less time.

In view of the sample selected for the empirical work in this thesis, it will be useful to report the results of J.C. McKeown (1974) for CAI in elementary college accounting. The system under scrutiny is the well-publicized PLATO IV. A full description of the system operation is included in the next chapter.

After a year of development and debugging, a test of the effectiveness was conducted by McKeown in Fall 1973. He used a Latin Squares research design to test two hypotheses:

H1) That students can be brought to at least as good a performance level with less class time and significantly less total student time spent on the course using PLATO compared with conventional teaching.

H2) That students can be brought to a significantly better performance level using PLATO compared with conventional teaching.

Four sections of approximately 40 students each were involved, two meeting at 10 a.m. and two at 11 a.m. Students were randomly assigned to control and experimental groups for the two time slots with two different instructors, each teaching one control and one experimental section. The test results are reported below on Table 4.9.

With the Latin Squares design, the effects of instructor or meeting times are factored out. Results for the first three examinations indicate no significant difference for either method, while on the final examination, performance is about 6 percent better than the control group at a 1 percent level of significance. McKeown suggests that the difference between the earlier exams and the final may be accounted for by

TABLE 4.9

TEST RESULTS FOR PLATO IV IN ELEMENTARY ACCOUNTING*

	Exam Means				F Statistic
	10 AM Section		11 AM Section		
	Exper.	Control	Exper.	Control	
Exam 1	80.9%	84.1%	80.2%	80.6%	1.06
Exam 2	65.8	67.8	63.1	66.2	.90
Exam 3	88.5	89.6	85.6	86.4	.25
Final (Comprehensive)	81.2	76.1	76.2	70.7	9.12**
Students missing from final	1 prior B grade	1 prior F grade	1 prior B, C grade	1 prior F grade	

* Source: J.C. McKeown, "Computer-Assisted Instruction for Elementary Accounting," (Mimeo, Working Paper, College of Commerce and Business Administration, University of Illinois at Champaign-Urbana, June 1974), p. 14.

** Sign. at.01.

different sensitivity in the output instrument (fewer multiple choice questions on the final, and a longer exam), by accumulated student experience with the PLATO medium, and improved system reliability later in the semester.[14]

McKeown's study of student time and homework completion is also interesting. First, PLATO students completed a higher proportion of their assigned work. The experimental sections completed 20 and 15 percent more of their homework problems than the control students. Moreover, the experimental groups spent an average of 10 and 14 hours less time in this activity, respectively. In McKeown's view both hypotheses, H1 and H2, are supported. Not only do the results for the PLATO system coincide with the result in the literature that CAI-educated students will perform at least as well as TI students, but the demonstrated savings in time for the computer student is particularly impressive.

More research needs to be undertaken to determine the tradeoffs between CAI and TI with respect to time savings and other resources. This is especially important and attractive in view of the accelerating development of this medium, the desirability of individualized instruction, and the expected declines in relative cost of CAI technology compared with projected costs of traditional instruction.

---

[14]McKeown, p. 11.

CHAPTER FIVE

PROGRAM SITE SELECTION AND DATA COLLECTION PROCESS

Site Selection

When this research was originally planned, a data set for estimating the effect of educational technology was not available. The entire project was contingent on the location of an appropriate sample of schools and students, and the successful collection of data. Although those efforts proved time consuming, they also provided some of the most interesting and educational aspects of the research effort.

In its initial stages, the objective of this research was to study the role and effectiveness of communications technology for off-campus, post-secondary education. The potential for technology in the "open university" or "university without walls" was of particular interest since they sought to shift the focus of college education from structured, on-campus activities to flexible, unstandardized programs emphasizing individualized instruction. A survey of institutions was conducted by telephone to determine the extent of the technology in use. On-site visits were made to six institutions in the East including Empire State College, the Open University at the University of Maryland, and the New York Regents External Degree Program. Several facts soon became apparent: For the most part, technology was not used heavily in off-campus college programs. The learning units of these institutions do not have an analogue within traditional systems with which they can be compared. Even though the university without walls may represent a fertile opportunity for the expansion of technology, the dissimilarities with traditional instruction in terms of objectives, structure, and content make effectiveness study impractical to say the least. At this point, the search for data shifted from nontraditional institutions to a survey of technology utilization by traditional colleges and universities.[1]

Another problem that arose early in the research was the apparent impossibility of gathering enough observations for students who had been subjected to technology and had also taken a standardized test, like the GRE and Medical or law boards, that could readily serve as the output measure. It gradually became obvious that site selection would be subject to the condition that data be collected for on-campus students in a particular course taught by alternative technologies at different schools. Furthermore, a suitable output measure would have to be devised to fit the situation.

In early 1974, the universities and colleges known to offer courses by ITV or CAI were surveyed in the hope of finding a course in mathematics, economics, business or a foreign language similar in content but taught by different technologies. An accounting course was located at Colorado State University. It was taught by ITV and similar to one taught

---

[1]Some of the institutions contacted were CUMBIN (CUNY), Stanford University, Colorado State University, University of South Carolina, Case Western Reserve, Pennsylvania State University, TAGER (Southern Methodist University), GENESYS (University of Florida), University of Southern California, University of Connecticut, Chicago TV College, Oklahoma Higher Education Televised Instruction System, University of Illinois (PLATO), and Indiana Higher Education Telecommunications System.

by CAI on the PLATO system at the University of Illinois. The faculty involved were contacted and agreed to participate in this research. The cooperation of an accounting professor at the University of Missouri in St. Louis was solicited to represent the traditional school. With cooperation assured from schools with three different technologies, the next step was to develop an output measure and collect the data.

Program Description

Colorado State University is recognized as one of the forerunners in off-campus televised instruction, especially in the field of engineering. At present, CSU maintains an extensive graduate education program known as SURGE offering over 30 courses to employees of 27 different industries in Colorado, an ITV project with 14 other colleges in the state (CO-TIE), and a program to provide college courses to high school teachers and students (HI-TIE). The CSU network is a videotape technology.[2] Actual classroom lectures are taped along with impromptu discussion and shipped by courier to various remote locations, later to be returned and retaped.

In contrast to most of the ITV activity, accounting (BA200), the subject of our attention, took place on campus. Referring to Table 5.1 for a detailed comparison of the programs, you can see that the course met for approximately one hour, three days a week for nine weeks. On Monday and Wednesday, the students watched a 25-minute videocassette lecture, followed by discussion. Friday was a lab period devoted to

---

[2]For more system information, see L.V. Baldwin (September, 1973). pages 172-92, or M. Wong (August, 1974) pages 145-56.

TABLE 5.1

PROGRAM PROFILES

	Institution		
	U. Missouri St. Louis	Colorado State U.	University of Illinois
1. Technology	traditional lecture	ITV videocassette	Large scale interactive CAI network (PLATO IV)
2. School schedule	semester system (15 wks)	quarter system (9 wks)	semester system (15 wks)
3. Course	Fin Accy 140	BA 200	ACCY 101
	(all 3 are comparable first semester college acctg.)		
4. Faculty	David Ganz	Molly Murray	J.C. McKeown
5. Text	Finney & Miller (Chapt. 1-12)	Pyle & White (Chapt. 1-8)	Schattke, Jensen & Bean*
6. Class Meeting Information	55 min. class (8:10 - 9:10) Spring 1974: M,F lecture W lab Summer 1974: 5 days/wk	50 min. class - - M,W a 25-20 min. taped lecture & discussion F lab	90 min. class (8:00 - 9:30) T-Th lecture M,W,F PLATO lab (optional)
7. Total class time**	33 hours	27 hours (incl. ITV)	21 hours (not incl. CAI time)
8. Presiding Instructor	L - Dr. Ganz Lab - Grad. Asst.	Grad. Assts.	Grad. Assts.
9. Class Enrollment	150 (Spr. '74)*** 68 (Sum. '74)	239 in 12 Sec. (Spr. '74)	38 (Sum. '74) 114 in 4 Sec. (Fall '74)
10. Net Data Collected (# observations)	26 (May '74) 27 (July '74 + mail follow-up)	190 (May '74)	29 (July '74) 66 (Dec. '74)

* Summer only; a different text was used in the fall.

** Unfortunately student data were unavailable as enrollment figures were not recorded.

*** Survey was conducted during an optional review session only attended by about one-fourth of the students. Those at the review session also may be a biased group.

homework problems and questions. Lectures were taped by different faculty members who presided over their individual specialties. The same tapes were used in successive terms except that they were updated where necessary (approximately every two or three years). The tapes were produced in a studio-classroom equipped with three cameras (one overhead for transmitting desk notes and illustrations). The accounting class consisted of 12 sections of approximately 20 students each, with an MBA graduate student in charge.

University of Illinois: PLATO IV CAI system

Programmed Logic for Automatic Teaching Operation, better known as PLATO, is a sophisticated interactive CAI network developed since approximately 1960 at the Computer-Based Education Research Laboratory (CERL) of the University of Illinois. The system comprises a central computer at Urbana-Champaign and student terminals at remote locations. As of 1972, the system had grown from only 40 terminals in service to 1,000, one third of which are on campus and the rest dispersed at colleges and military and commercial locations throughout 24 states.

Communication with the student is by means of words, figures, graphs, or pictures transmitted on a television-like display panel composed of ionized gas sandwiched between two thin sheets of glass. Each point in the 512 x 512 element screen can be selectively stimulated by electrodes at the edge of the panel. The result is a bright display that can be successively modified without retransmission. PLATO's ability to generate a high quality display that combines inherent memory with a relatively inexpensive unit requiring low data transmission rates is a particularly significant development.

The student responds to the computer by typing replies on a keyboard very similar to a typewriter or, in some cases, by touching the panel at the appropriate spot. A student experiencing difficulty may press the "help" key for a complete explanation of the problem, a review of concepts, and hints of the correct solution. The computer is programmed to accept answers containing typographical errors so typing inaccuracy does not inhibit student progress. The system allows the student to enter an algebraic expression as an answer to a question, enabling him to concentrate on the method of solution rather than being encumbered or frustrated by arithmetic errors. The student may sign on the computer and call up any particular lesson he wishes. He may work during appointed lab sessions or during any unrestricted periods. The computer automatically keeps a student record of lessons completed, time on, days on, plus any other information of interest to the faculty member.

The versatility of PLATO has enabled it to adapt to education at all levels, though its use centers on post-secondary courseware. With the exception of some courses in fine arts, it has the capability to transmit instruction in almost any area. "About 4,000 hours of instructional material are now available in more than 100 subject areas, from elementary reading and math to language, music, science, business and social sciences."[3] The CAI curriculum materials generally are developed in subject units of lessons on a particular topic. They may involve simulations, games, drill and practice, or programmed learning. For most university courses using PLATO, the CAI materials will perform the major

---

[3]E. Jenkins (1976).

instructional roles at certain times; otherwise, course instruction is performed by traditional lectures. One might classify CAI use roughly into three categories: first, those cases where CAI is purely supplementary because it reinforces material transmitted primarily through lectures. Second are those courses where CAI functions as a major instructional component for some class topics. Third are the courses where CAI is relied on as the primary source of instruction and virtually supplants traditional instruction. Nationwide, most CAI is supplementary and falls into the first group. The majority of the PLATO curriculum is represented by the second category; a few courses fall into the third group, the only case where CAI approaches being a substitute for traditional instruction at the course level.

Accounting 101 is one of the courses where PLATO performs the primary instructional function for the duration of the course. It consists of 25 discrete lessons plus an initial PLATO orientation lesson. Later lessons build on the knowledge of prior exercises, but the lessons are sufficiently separable to be worked out of sequence. A lecture session manned by a graduate assistant is held for two 90-minute periods per week; these function primarily as discussion sections. Students may use the computer terminals located in a central laboratory at any of three reserved periods per week or during any unrestricted time at their own discretion. The student who reads the appropriate sections of the text and actively pursues the CAI lessons will generally use lecture sections to discuss solutions to questions that arise during independent study and to maintain faculty contact on administrative matters.

University of Missouri at St. Louis: Traditional Instruction

The accounting course representing traditional instruction is the Fundamentals of College Accounting (140) taught by Professor David Ganz. Every detail--from the course syllabus and lecture to homework assignments--was carefully structured and developed by Professor Ganz through several years' experience with the course.

Data were collected in two different terms. In the spring of 1974, the lecture met on Monday and Friday with a lab on Wednesday in the hands of a graduate assistant. In the summer of 1974, Dr. Ganz lectured five days a week and incorporated the discussion section into regular class periods.

Development of the Output Measure

After the sites for the study had been identified, selection and validation of an output measure was the next concern. A sample accounting examination was prepared from an instructor's manual and former examinations. Two work-sheet problems and 43 multiple choice questions covered the material common to the courses at the three schools. This examination was circulated among accounting professors at Washington University and the three participating universities for criticism and comment.

Professor Murray of Colorado State was the first to respond. She not only approved administering the examination to the ITV accounting students, but also suggested that it be made compulsory by including one of the work-sheet problems and all but five of the multiple choice questions as the comprehensive section of the course final (roughly half of the final examination grade).

With the exception of two questions, the University of Illinois accounting department agreed to give the same examination to its PLATO section as the comprehensive portion of its final examination. The traditional school consented to give three-quarters of the questions in the form of two quizzes and suggested fifteen points on their course final, which examined the same skills and concepts as the work study problem that was not administered.

In this manner, a standardized examination was agreed on and administered to 239 ITV students in the spring of 1974, and 35 CAI students and 68 TI students in the summer of 1974. The similarity of the three output measures is evident in Table 5.2, which also shows the conversion of raw scores into a standardized dependent variable. A copy of the examination itself is included in the appendix. As the selection of an output measure was being settled, a sample student questionnaire was also circulated among the participating faculty. A copy of the final version appears in the appendix.

Three implications of the output activities deserve mention. First, considerable effort was expended in the hope of constructing an identical output measure acceptable to the three schools involved. As Table 5.2 shows, the measures for the ITV and CAI students are nearly the same since the University of Illinois deleted only 4 percent of the questions from the examination that had been administered at Colorado State. The instrument was less suitable at UMLS, however, where only about 60 percent of the output measure was composed of questions identical to those given at Colorado State. Under the circumstances, the resulting output measure was the best possible approximation of a standardized output

TABLE 5.2

STRUCTURE OF THE OUTPUT MEASURE

		Exam Questions: Multiple Choice	Exam Questions: Work Sheet	Total Points	Dependent Variable (percentage)
Sample Exam:		43	2		
CSU*	Questions selected	38 of 43	1		Student Raw Score / 106 x 100
	Scoring	76 pts. (2 ea.)	30 pts.	106	
UI (PLATO)*	Questions selected	36 of 38	1		Student Raw Score / 51 x 100
	Scoring	36 pts. (1 ea.)	15 pts.	51	
UMSL**	Questions selected	30 of 38	None		Student Raw Score / 45 x 100
	Scoring	30 pts. (1 ea.)	15*** pts.	45	

* These exams represented the comprehensive section of the course final exam. In each case, this section comprised about half of the grade on the final exam.

** The instructor was not eager to incorporate these questions into his final exam because the same exam is administered each term to generate records on student performance over time. As a second-best alternative, 30 multiple choice questions were given in the form of two 15 point class quizzes. Only students who were present for both quizzes and who completed the questionnaire are included in the sample.

*** The worksheet problem from the CSU and UI final was not given at UMSL. As a substitute, 15 questions (15 points) were scored separately from the course final exam which seemed to best examine the same skills as the worksheet problem. These questions were determined by Dr. Ganz.

measure. Nevertheless, the differences between the outputs and the selective deletion practiced by many schools may bias the statistical analysis.

Second, Colorado State made the initial selection of questions from the sample examination. Faculties at the University of Illinois and University of Missouri at St. Louis were denied this privilege. Instead, they were asked to adopt the same examination as Colorado State from which each made selective deletions. In terms of the experimental results, this could lead to a bias in favor of students at Colorado State.

Third, the faculty members involved were aware that the examination would be used as an evaluative tool. Not only did they approve its use for this purpose, but they each elected to use the examination results in their own grading. This suggests a confidence in the program validity of the output measure that most school effectiveness studies lack. In summary, the conscious composition of an output variable in an "experimental" situation undoubtedly resulted in an instrument superior to measures that could have been selected from existing data.

Data Collection Procedures

The data collection process itself proceeded as follows. A personal visit was scheduled to each school near the end of the term. An attempt was made to talk to the students on a day when class attendance was at a normal level. A questionnaire and statement of compliance[4] were distributed to each student, followed by a brief explanation of the research project goals and procedures. Students willing to participate were asked to read the statement and sign it to verify their voluntary participation

---

[4]A copy of the statement appears in the appendix.

in the study. The statement also elicited the students' permission to access their files for additional information.

Students were instructed to place their name and student number only on a small piece of paper stapled to each questionnaire. Fifteen minutes of class time were provided to complete the questionnaire. The completed questionnaire and signed statements were both collected.

After a conference with the Dean of Admissions and presentation of the signed releases, permission was granted to enter the student files and record high school class rank and standardized test scores for verbal and mathematics aptitude, which were noted by student name and number. Once the course finals had been administered and scored, the pertinent raw scores were recorded by name and student number. For the PLATO students, CAI time on, days on and number of lessons completed were recorded by name and number. The data were then transcribed onto the student questionnaires and the name tags were removed, resulting in anonymous data sets for the different schools. An index number was assigned to each questionnaire so the original questionnaire could be matched with the computer-coded observations.

Table 5.3 gives a full list of the information collected for each student.

The next chapter begins with a description of the sample and a complete list of all variables included in the regressions.

TABLE 5.3

FULL LIST OF DATA COLLECTED[1]

Output

Achievement on accounting exam (percentage score)[2]

Socioeconomic Data

Age
Sex
Racial - ethnic group
Occupation of father
Occupation of mother
Total years formal education of father
Total years formal education of mother
Name of high school and its location (where student graduated)
Average number of hours per week spent working for a wage

Student Educational Background

Class
Major subject area
Minor subject area
Course load this term
College cumulative GPA
Grade average last term alone
Standardized Achievement Test scores (SAT, ACT)[3]
    math
    verbal
High school class rank[3]
Number of college credits accumulated
    in business
    in economics
    in mathematics

Accounting Course/Program Characteristics

Technology: Traditional, ITV, CAI
Estimated study time on course LAST WEEK outside of class
    (not including computer study)
Estimated study time on course during an AVERAGE WEEK ourside of class
Student self rating of study habits in general (Scale: 1 to 5)
Student self rating of study habits with respect to this course (1 to 5)
Student response to question: "Would you take accounting again?"

TABLE 5.3
(continued)

Data Stored As "Student Record" on Computer for CAI Students Only[4]

Name
Date record started
Last day on
Total hours on
CPU used
Days on
Sessions on
Lessons completed (recorded by lesson number)

---

1. All data was collected by questionnaire unless indicated otherwise (see notes 2 and 3).

2. Collected from class records.

3. Collected from admission files.

4. Information was collected for hour on, days on and lesson completed only.

CHAPTER SIX

EFFECTIVENESS ANALYSIS: EDUCATIONAL PRODUCTION FUNCTION RESULTS

The data for this study were collected on five separate occasions. The first opportunity occurred at the traditional school in May 1974, when 26 observations were collected for students in the spring term at University of Missouri in St. Louis (UMSL). The output measure for this group was the percent score on the final examination. This data set is designated by the label LECT1.

Shortly afterwards, the standardized output measure was accepted by the ITV institution. On a visit to Colorado State University during May 1974, 190 cases were collected for the spring term. That visit was followed by two collections later that summer at other schools. Table 6.1 presents the breakdown of the sample by technology and output measure. The two samples in summer 1974 provided 27 observations by traditional instruction (TI) and 29 by CAI; these were labeled with set names LECT2 and CAI2, respectively. The union of LECT2, ITV, and CAI2 composes the data set called POOL, with the key distinction that each participant was subjected to the same output instrument.

Observations were included if the students were present on the day the questionnaire was circulated, were willing to participate, and had standardized test scores in their admission files. Thus, students with

TABLE 6.1

PROFILE OF SAMPLE*

Output Measure***	Technology**			
	TI	ITV	CAI	TOTAL
Non-Standard Exam	LECT 1 N = 26 Spring '74	0	CAI 1 N = 66 Fall '74	ADD 92
Experimental Exam	LECT 2 N = 27 Summer '74	ITV N = 190 Spring '74	CAI 2 N = 29 Summer '74	POOL 246
TOTAL	LECTUR 53	ITV 190	CAI 95	ALL 338

* Key to tabular information:

(Data Set Name)
(Number of Cases)
(When Collected)

** Applicable technologies are traditional instruction (TI) at University of Missouri at St. Louis, instructional television (ITV) system on campus of Colorado State University, and the computer-assisted instruction (CAI) network at University of Illinois (known under acronym, PLATO).

*** Three groups took an (almost) identical exam that had been constructed as the output measure for this study. (See Table 5.2 for exam details.) In the other cases, results on the course final were used.

poor attendance records, transfer students, or foreign students were probably underrepresented in this sample.

Since enrollment in the summar PLATO course was comparatively small, it was decided to return to the University of Illinois in the fall for additional data. Consequently, in December 1974, an additional 66 cases were gathered under the name CAI1. By that time, however, all course examinations including the final were being administered by the computer system. Therefore, it was not possible to use the standardized output instrument. Like the initial collection effort at UMSL, the percent score on the course final represented the best output measure available. The latest CAI sample and the initial TI set were collated with the POOL sample to form a composite set named ALL, consisting of 338 observations.

The variables used in the regression analysis are described in Table 6.2. The socio-economic variables are represented by a four-category classification of father's occupation, father's education, and mother's education. Sex, market work (JOB), and academic load (LOAD) are standardizing variables. SAT mathematics and verbal scores, class rank, and credits in business, economics, and mathematics adjust for differences in student abilities. Given these characteristics, achievement in accounting is described by variation in student study time, motivation, and program variables. In the case of program variables, traditional instruction is used as the benchmark from which deviations associated with ITV and CAI are measured. ITV is a shift variable to measure the change in the production function intercept associated with ITV. Since information on individual student attendance or class attentiveness was not available, the use of the technology did not vary across students, and there is no

TABLE 6.2

LIST OF VARIABLES USED IN THE ANALYSIS

Var Name	Units	Mean[1]	SD[2]	Definition
1. OUTPUT	percent	77.5	12.6	Achievement score on accounting exam
2. SEX	d[3]	.38	.48	= 1 if female; = 0 if male
3. OCCUP1[4]	d	.18	.39	= 1 if father's occupation professional, 0 otherwise
4. OCCUP2	d	.44	.50	= 1 if father's occupation is manager, proprietor category; 0 otherwise
5. OCCUP3	d	.22	.42	= 1 if father's occupation is in skilled category; 0 otherwise
6. EDUCF	years	14.2	3.1	Total years education of father
7. EDUCM	years	13.1	2.5	Total years education of mother
8. JOB	hrs/wk	6.3	11.4	Average hours per week spent working for a wage
9. LOAD	cred-hrs	14.7	2.9	Course load during the current term
10. SATM	stndzd score	565	87	Scholastic Aptitude Test in math (or ACT equiv)
11. SATV	stndzd score	490	89	Scholastic Aptitude Test in verbal ability (or ACT equiv)
12. RANK	percent	74.3	19.8	High school graduation class ranking
13. CREDB	cred-hrs	4.0	4.5	College credits accumulated in business
14. CREDE	cred-hrs	4.6	4.5	College credits accumulated in economics
15. CREDM	cred-hrs	8.2	5.4	College credits accumulated in math
16. STUDY	hrs/wk	4.2	2.6	Study time on course outside of class in an average week (not incl. computer time)
17. MOTIV	d	.73	.44	= 1 if student responds that he would be willing to take another accounting course; = 0 otherwise
18. ITV	d	.77	.42	= 1 television technology; = 0 otherwise

TABLE 6.2
(continued)

Var Name	Units	Mean[1]	SD[2]	Definition
19. INRCPT[5]	d	.12	.32	= 1 if computer technology; = 0 otherwise
20. CAILSN[6]	percent	67.0	25.9	Percentage computer lessons completed by student (25 total)
21. CAITIM[6]	hrs	25.5	12.8	Total hours during term student spent on CAI system
22. CAIDAZ[6]	# days	14.6	5.4	Total number of days during the term which the student used the system

[1]Variable means for sample named POOL (n = 246). A complete table of simple statistics for all sample subsets appears in the appendix.

[2]Standard deviations.

[3]d indicates a dummy variable.

[4]OCCUP4 representing the unskilled occupation category is the base against which OCCUP differences are measured.

[5]INRCPT captures the average expected change in student performance for computer education students with CAI use set equal to zero. Hence, it will always appear in conjunction with one of the three variables CAILSN, CAITIM, or CAIDAZ to account for variation in CAI use among students.

[6]Variable pertains only to CAI students, = 0 for everyone else. The mean and standard deviations are calculated for the subset of 29 students in data set POOL who experienced CAI.

way in this analysis to explore the effect on achievement of varying amounts of ITV.

On the other hand, three measures of CAI use were available: percent lessons completed, total time on, and number of days on. Since these measures are highly intercorrelated, they could not all be incorporated in the analysis. However, using one of the measures in conjunction with an intercept term allows investigation of the difference in achievement between TI and CAI. First, CAI use is set equal to zero. Then, increases in achievement associated with increased use of CAI are observed.

Some simple comparisons of the different data collections may be worthwhile. The differences in means for output, parents' education, credits in business, mathematics, and economics for the five samples are slight. This is not so with student ability. Scholastic Aptitude Test scores range from 531 (M) and 459 (V) for the summer TI students to a mean of 624 (M) and 555 (V) for the summer CAI students. Mean class ranks vary from 65.7 to 84.7 across the five samples. All three ability measures show a clear pattern, with the traditional students at UMSL representing the lower ability level, the ITV students at CSU a higher level, and the University of Illinois PLATO participants at the top. Although the coefficient of variation is greater at UMSL, it diminishes as we proceed to schools with high ability students.

Other striking differences among the samples occur for the LOAD, STUDY, and JOB variables. These can readily be explained by the differences between summer and regular school term programs. Data on academic load collected during the summer term were standardized to a normal term, assuming a full load to be 9 credit hours in the summer and 16 in the spring or fall terms. Since the summer TI class met every day and the question concerning study time requested a weekly figure, the data for this group had to be adjusted to conform to a class that met three times a week, as was the case with the other four samples.

Study time by PLATO students seems slightly lower than the rest, which is possibly explained by the fact that these figures exclude CAI time. The students at UMSL show a stronger attachment to the workforce and a larger number of hours worked (for a wage) on the average than students at either of the other two schools. In addition, students in the sample during the summer term show a greater commitment to a job.

It is important to recognize differences among the schools in the explanatory variables, because the analysis attributes existing performance differentials between the schools to the technologies involved. When variables at different schools present different distributions, as is the case with student ability, a systematic bias may be introduced to the school technology terms. This danger is especially great when the sample is dominated by one group, as the partition POOL is by students from Colorado State. The extent of this sample problem will be investigated later.

Empirical Model and Statistical Methods

Chapter 2 discussed the relative merits of estimating two alternative models: the linear additive model and the Cobb-Douglas form. Operational identification of inputs permits estimation of the following equations:

Linear Additive

$$\begin{aligned}OUTPUT_i = {} & \beta_0 + \beta_1 SATM_i + \beta_2 SATV_i + \beta_3 ITV_i + INRCPT_i(\beta_4 + \beta_5 CAILSN_i) \\ & + \beta_6 SEX_i + \beta_7 EDUCF_i + \beta_8 EDUCM_i + \beta_9 RANK_i + \beta_{10} CREDB_i \\ & + \beta_{11} CREDE_i + \beta_{12} CREDM_i + \beta_{13} JOB_i + \beta_{14} LOAD_i \\ & + \beta_{15} STUDY_i + \beta_{16} MOTIV_i + \varepsilon_i \,. \end{aligned} \qquad [6.1]$$

Cobb-Douglas

$$\begin{aligned}\ln (OUTPUT_i) = {} & \ln \beta_0 + \beta_1 \ln(SATM_i) + \beta_2 \ln(SATV_i) + \beta_3 \ln(EDUCF_i) \\ & + \beta_4 \ln(EDUCM_i) + \beta_5 ITV_i + INRCPT_i(\beta_6 + \beta_7 CAILSN_i) \\ & + \beta_8 SEX_i + \beta_9 RANK_i + \beta_{10} CREDB_i + \beta_{11} CREDE_i \\ & + \beta_{12} CREDM_i + \beta_{13} JOB + \beta_{14} LOAD_i + \beta_{15} STUDY_i \\ & + \beta_{16} MOTIV_i + \ln \varepsilon_i \,. \end{aligned} \qquad [6.2]$$

where ε_i and $\ln \varepsilon_i$ are the stochastic disturbance terms, respectively.

The linear additive model was chosen for its simplicity and ease of interpretation. Since it is by far the most common in the literature, the results for equation 6.1 provide a ready comparison of this study with other research findings.

The main attraction of the Cobb-Douglas specification is that the marginal products are variable and a function of the input levels. The estimated coefficients are themselves the elasticities of output with respect to the inputs.[1]

---

[1]This is not true for the categorical variables or for CREDB, CREDB, CREDE, CREDM, JOB, or STUDY. These are not expressed in percent, and it was not feasible to take logs because of the presence of zero values for some observations.

The method of analysis employed in this study is a stepwise regression method, developed by James Goodnight, that uses maximum R^2 improvement as the search procedure:

> It finds first the one-variable model producing the higher R^2 statistic. Then another variable, the one which would yield the greatest increase in R^2, is added. Once this two-variable model is obtained, each of the variables in the model is compared to each variable not in the model. For each comparison, the procedure determines if removing the variable in the model and replacing it with the presently excluded variable would increase R^2. After all the possible comparisons have been made, the switch which produces the largest increase in R^2 is made. Comparisons are made again, and the process continues until the procedure finds that no switch could increase R^2. The two-variable model thus settled on is considered the "best" two variable model the technique can find. The technique then adds a third variable to the model, according to the criteria used in adding the second variable. The comparing and switching process is repeated, the "best" three-variable model is discovered, and so forth. This technique differs from the conventional STEPWISE technique[2] in that here all switches are evaluated before any switch is made.[3]

This technique is considered "almost as good as calculating regressions on all possible subsets of the independent variables."[4]

One should guard against the indiscriminate and unqualified application of regression analysis in social science research. We must be alert to the possibilities for misinterpretation, especially with the kind of data employed in this research. Outside of measurement and specification error, the greatest econometric problems are likely to stem from the presence of outlier observations, or multicollinearity among the explanatory

---

[2]Most stepwise programs use a forward selection procedure, where variables are sequentially included in descending order according to their F statistics, and those already included will be deleted if their partial F statistic falls below a specified significance level.

[3]A.J. Barr and J.H. Goodnight (1972) p. 128.

[4]*Ibid.*, p. 128

variables. The presence of outliers can be determined by examining the plot of standardized residuals. Three student observations showing 20 credit hours in the summer term created a difficulty with the LOAD variable for some sample configurations. Once these were discarded, consistent results were obtained.

When factor analysis was applied to the data, multicollinearity did not appear to be a serious problem. Moreover, the stepwise regression procedure helps to identify areas where multicollinearity is present by revealing changes in the coefficients on a given variable from one step to the next. The matrix of correlation coefficients is also of some use in diagnosing this problem.

Effectiveness Analysis: The Results

An educational production function is estimated by ordinary least squares under several alternative specifications. First, extensive regression results are presented for the linear and logarithmic functional forms for 246 POOL students who were subjected to the experimental output measure. After comparing outcomes for the two functional forms, the results for individual variables are interpreted and discussed. Information is frequently brought to bear from additional regressions and correlation matrices to enlighten the picture of individual inputs in the education production process. Once all variables have been discussed, results are reported for alternative sample stratifications. A simple three variable model is estimated for subsamples of the data. Estimating a separate regression for each technology allows for interaction effects among the variables and provides a check on the results obtained in the pooled analysis.

Tables 6.3 and 6.4 describe the regression results for the linear and

TABLE 6.3

DETERMINANTS OF PERFORMANCE IN COLLEGE ACCOUNTING
LINEAR PRODUCTION FUNCTION MODEL*

DSN = POOL; N = 246	MODEL						
Variable	Best 2	Best 3	Best 4	Best 5	Best 6	Best 7	Best 8
1. SATM	0.056 (6.7)**	0.044 (4.7)	0.043 (4.7)	0.047 (5.1)	0.044 (4.77)	0.044 (4.74)	0.045 (4.89)
2. SATV		0.026 (2.89)	0.028 (3.10)	0.026 (2.93)	0.024 (2.62)	0.026 (2.82)	0.026 (2.76)
3. ITV	8.48 (4.9)	8.88 (5.2)	9.44 (5.5)	6.34 (2.82)	5.94 (2.63)	6.41 (2.83)	6.00 (2.64)
4. INRCPT				-20.80 (3.41)	-21.78 (3.56)	-21.30 (3.5)	-21.39 (3.5)
5. CAILSN				0.235 (2.95)	0.238 (3.01)	0.230 (2.90)	0.225 (2.89)
6. RANK					0.066 (1.77)	0.065 (1.73)	0.074 (1.95)
7. STUDY							0.454 (1.48)
8. EDUCF			-0.406 (1.75)			-0.360 (1.58)	-0.342 (1.51)
9. EDUCM							
10. MOTIV							
11. LOAD							
12. JOB							
13. SEX							
14. CREDM							
15. CREDB							
16. CREDE							
CONSTANT	39.1	33.1	37.7	33.6	31.8	36.1	33.1
R^2	0.210	0.237	0.246	0.272	0.282	0.289	0.295

* Dependent variable equals percentage correct on identical examination.

** t-statistics in parentheses.

TABLE 6.3

(Continued)

MODEL							
Best 9	Best 10	Best 11	Best 12	Best 13	Best 14	Best 15	Best 16
0.045 (4.87)	0.045 (4.8)	0.045 (4.82)	0.044 (4.53)	0.044 (4.34)	0.044 (4.34)	0.044 (4.3)	0.044 (4.31)
0.026 (2.84)	0.025 (2.68)	0.025 (2.69)	0.025 (2.70)	0.025 (2.68)	0.026 (2.68)	0.026 (2.68)	0.026 (2.68)
6.06 (2.67)	7.12 (2.80)	6.01 (2.15)	6.05 (2.16)	6.12 (2.16)	6.10 (2.15)	6.14 (2.15)	6.16 (2.15)
-21.39 (3.52)	-21.38 (3.52)	-21.8 (3.5)	-21.8 (3.6)	-21.93 (3.6)	-22.00 (3.57)	-22.02 (3.56)	-21.95 (3.5)
0.225 (2.85)	0.233 (2.94)	0.234 (2.95)	0.235 (2.95)	0.237 (2.95)	0.239 (2.95)	0.239 (2.94)	0.239 (2.92)
0.075 (1.97)	0.076 (2.00)	0.073 (1.93)	0.074 (1.95)	0.077 (1.92)	0.076 (1.88)	0.075 (1.85)	0.075 (1.85)
0.490 (1.59)	0.503 (1.64)	0.524 (1.70)	0.526 (1.70)	0.536 (1.71)	0.537 (1.71)	0.532 (1.68)	0.533 (1.68)
-0.324 (1.42)	-0.311 (1.36)	-0.314 (1.37)	-0.321 (1.39)	-0.318 (1.37)	-0.321 (1.38)	-0.322 (1.38)	-0.312 (1.15)
							-0.024 (0.07)
1.99 (1.27)	1.91 (1.21)	2.09 (1.32)	2.06 (1.28)	2.03 (1.26)	2.02 (1.25)	2.08 (1.26)	2.08 (1.26)
		0.266 (0.95)	0.259 (0.92)	0.255 (0.90)	0.252 (0.89)	0.256 (0.90)	0.257 (0.90)
	0.066 (0.93)	0.069 (0.98)	0.069 (0.97)	0.069 (0.97)	0.068 (0.95)	0.07 (0.97)	0.07 (0.96)
				-0.323 (0.2)	-0.38 (0.23)	-0.43 (0.26)	-0.44 (0.26)
			0.041 (0.31)	0.037 (0.28)	0.035 (0.26)	0.04 (0.3)	0.041 (0.29)
					0.033 (0.2)	0.037 (0.2)	0.037 (0.2)
						-0.032 (0.1)	-0.033 (0.2)
30.9 0.300	30.2 0.303	27.1 0.305	27.2 0.305	27.2 0.306	27.2 0.306	27.2 0.306	27.4 0.3063

TABLE 6.4

DETERMINANTS OF PERFORMANCE IN COLLEGE ACCOUNTING
LOGARITHMIC PRODUCTION FUNCTION MODEL*

DSN = POOL; N = 246	MODEL						
Variable	Best 2	Best 3	Best 4	Best 5	Best 6	Best 7	Best 8
1. LSATM	0.403 (5.9)	0.297 (3.99)	0.337 (4.5)	0.323 (4.3)	0.3023 (4.04)	0.3157 (4.2)	0.313 (4.18)
2. LSATV		0.2047 (3.19)	0.214 (3.32)	0.206 (3.23)	0.187 (2.91)	0.183 (2.85)	0.191 (2.95)
3. ITV	0.116 (4.41)	0.122 (4.7)		0.084 (2.45)	0.078 (2.26)	0.072 (2.09)	0.077 (2.23)
4. INRCPT			-0.380 (4.25)	-0.3035 (3.24)	-0.319 (3.4)	-0.320 (3.43)	-0.315 (3.37)
5. CAILSN			0.00341 (2.76)	0.00342 (2.80)	0.00347 (2.85)	0.00344 (2.84)	0.00334 (2.75)
6. RANK					0.00101 (1.77)	0.00115 (1.978)	0.00112 (1.93)
7. STUDY						0.00692 (1.47)	0.00664 (1.41)
8. LEDUCF							-0.0547 (1.21)
9. LEDUCM							
10. MOTIV							
11. LOAD							
12. JOB							
13. SEX							
14. CREDM							
15. CREDB							
16. CREDE							
CONSTANT	1.693	1.099	0.894	0.961	1.136	1.048	1.158
R^2	0.174	0.207	0.221	0.240	0.250	0.257	0.262

* Dependent variable equals LOG (OUTPUT); t-statistics appear in parentheses.

TABLE 6.4

(Continued)

MODEL							
Best 9	Best 10	Best 11	Best 12	Best 13	Best 14	Best 15	Best 16
0.309 (4.12)	0.309 (4.1)	0.310 (4.1)	0.311 (4.13)	0.301 (3.83)	0.2934 (3.63)	0.2950 (3.63)	0.2948 (3.6)
0.182 (2.80)	0.186 (2.84)	0.186 (2.85)	0.184 (2.81)	0.191 (2.85)	0.192 (2.86)	0.193 (2.86)	0.193 (2.86)
0.095 (2.43)	0.095 (2.42)	0.0823 (1.91)	0.0806 (1.865)	0.0834 (1.91)	0.0839 (1.92)	0.0835 (1.90)	0.0840 (1.905)
-0.315 (3.37)	-0.315 (3.36)	-0.319 (3.4)	-0.325 (3.4)	-0.328 (3.46)	-0.329 (3.45)	-0.330 (3.46)	-0.330 (3.4)
0.00347 (2.84)	0.00342 (2.81)	0.00344 (2.81)	0.00348 (2.84)	0.00355 (2.87)	0.00356 (2.87)	0.00360 (2.88)	0.00361 (2.875)
0.00114 (1.961)	0.00115 (1.98)	0.00112 (1.92)	0.00113 (1.94)	0.00122 (2.00)	0.00124 (2.01)	0.00122 (1.96)	0.00121 (1.94)
0.00689 (1.46)	0.00725 (1.53)	0.00749 (1.58)	0.00750 (1.57)	0.00788 (1.63)	0.0078 (1.63)	0.0078 (1.63)	0.00783 (1.61)
-0.0522 (1.16)	-0.0494 (1.092)	-0.0503 (1.11)	-0.0651 (1.24)	-0.0632 (1.20)	-0.0642 (1.22)	-0.0648 (1.22)	-0.0648 (1.22)
			0.0309 (0.56)	0.0304 (0.55)	0.0289 (0.52)	0.0283 (0.51)	0.0284 (0.51)
	0.0207 (0.85)	0.0228 (0.93)	0.0232 (0.94)	0.0217 (0.88)	0.0210 (0.85)	0.021 (0.84)	0.0216 (0.85)
		0.0005 (0.70)	0.0004 (0.69)	0.0004 (0.64)	0.0004 (0.61)	0.0004 (0.60)	0.0004 (0.61)
0.00107 (0.99)	0.0010 (0.94)	0.0010 (0.97)	0.0010 (0.96)	0.0010 (0.96)	0.0010 (0.96)	0.0010 (0.93)	0.0010 (0.94)
				-0.0123 (0.50)	-0.0110 (0.44)	-0.0121 (0.48)	-0.0128 (0.49)
					0.0008 (0.39)	0.0008 (0.36)	0.0008 (0.38)
						0.0006 (0.27)	0.0007 (0.28)
							-0.0003 (0.13)
1.205	1.161	1.123	1.086	1.108	1.145	1.133	1.133
0.265	0.267	0.268	0.269	0.270	0.270	0.271	0.271

Cobb-Douglas (logarithmic) models, respectively. These and other selected results represent the statistics most often generated in the many trials that could not be reported.

The similarities between the linear and logarithmic regressions are more striking than the differences. Faced with the same list of 16 independent variables, each model almost always selects the same variables at each stage in the stepwise process. SATM, SATV, ITV, INRCPT, and CAILSN are five variables significant in both models at the 1 percent level using a two-tailed "t" test. With the possible exception of RANK, the next four variables--RANK, STUDY, EDUCF, and MOTIV--would be considered only marginally significant at the 5 to 20 percent level. Beyond these, significant results drop off considerably, with a t-statistic on added variables less than 1.0 and the increase in R^2 appearing in the thousandths column. With the exception of EDUCF,[5] each of the top nine or ten variables displays the expected sign.

As mentioned earlier, the interpretation of the coefficients is quite different in each case. In the log model, the coefficients are the output elasticities of the factors. To take a specific example, a 1 percent increase in SATM or SATV suggests a 0.3 percent or a 0.19 percent gain in achievement, respectively. One could calculate the elasticities of SATM and SATV at their means for the linear model as a quantitative comparison of the two models. These figures are approximately 0.26 for SATM and 0.125 for SATV. The output elasticities for these two variables do differ and appear to be slightly larger with the Cobb-Douglas specification. Nevertheless, quantitative differences of this magnitude are probably

[5]This exception will be discussed later.

not important in light of the lack of confidence shown by most researchers in the regression coefficients of school effectiveness studies.

Overall, the linear and logarithmic models seem to perform equally well using the same variables. The results here are very much in line with those reached in Wells (1974). Namely, there is no statistical evidence for selecting one model over the other.

A logical explanation can be proposed for the similarity of results between the two models. Because of the variables involved, taking logs of both sides of the equation did not involve transforming many variables. On the right side of the equation, SATM, SATV, EDUCF, and EDUCM were replaced with their logarithms. All other independent variables were left intact because they were either dummy variables or variables with numerous zero-valued observations (STUDY, CREDE, CREDB, CREDM, JOB). Discarding the observations or substituting the value "one" for "zero" to take logs were unreasonable options.

Consequently, SATM and SATV are the only two significant inputs that were transformed, and, for values in the hundreds like these test scores, the log curve is rather closely approximated by a straight line. Because of the few variables actually involved in the transformation, and the magnitude of the ones that are, the similarity of results for the linear and Cobb-Douglas models is not surprising.

Interpretation of results for inputs to the production function deserves careful attention. The next several pages are devoted to discussing individual regression variables within the four broad categories representing the effect of technology, student background characteristics, student ability, and student effort and motivation.

Although results may occasionally seem encouraging, two caveats must be kept in mind. First, since these data represent a unique sample of students in a particular course with three different technologies and only three schools in the sample, one must guard against generalizing the results to other course material or other schools. Second, although the output instrument was a carefully conceived measure, it is not identical for the three schools. Furthermore, Colorado State's option of initially selecting the questions could bias the results in favor of characteristics associated with that sample.

Technology

Three measures were available to capture the impact of different levels of CAI use by individual students: percent CAI lessons completed (CAILSN), total hours spent on the system (CAITIM), and total number of days that the student used the system during the term (CAIDAZ). Since the simple correlation coefficient between any two was around 0.9, it was appropriate to include only one measure in the regression at any given time. With maximum R^2 improvement as the selection criteria, the preference ranking for these variables is CAILSN first, then CAITIM, with CAIDAZ a decidedly poor performer by comparison and frequently revealing a puzzling negative coefficient. CAILSN dominated CAITIM and CAIDAZ in the regression, but it also provided the most logical, straightforward explanation for variation in achievement among CAI students: The more lessons completed by the student, the higher the expected level of performance.

If CAILSN and CAIDAZ were omitted, CAITIM was almost as effective as CAILSN in reducing the test score variance. Aside from that, its attractiveness lies with the units in which it is expressed. The effectiveness analysis indicates the average improvement from one hour spent on computerized instruction. Since the marginal cost of an hour of CAI can be calculated, CAITIM proves to be a convenient variable for computing the cost effectiveness of CAI.[6]

Furthermore, the use of CAITIM permits investigation of how students allocated study time. If the opportunity cost to the student on an hour of CAI and an hour reading the textbook are the same, the student will maximize his achievement by a time allocation that equalizes the marginal benefit of time in each activity.

The most interesting result is the important role played by technology in determining achievement level. The school/technology variables--ITV, INRCPT, and CAILSN--were significant in all of the regressions at the 1 percent level. The coefficients changed slightly as the stepwise process added variables, but this happened with all variables in the regressions. Moreover, the changes observed in Tables 6.3 and 6.4 are tolerable and not indicative of severe multicollinearity problems.

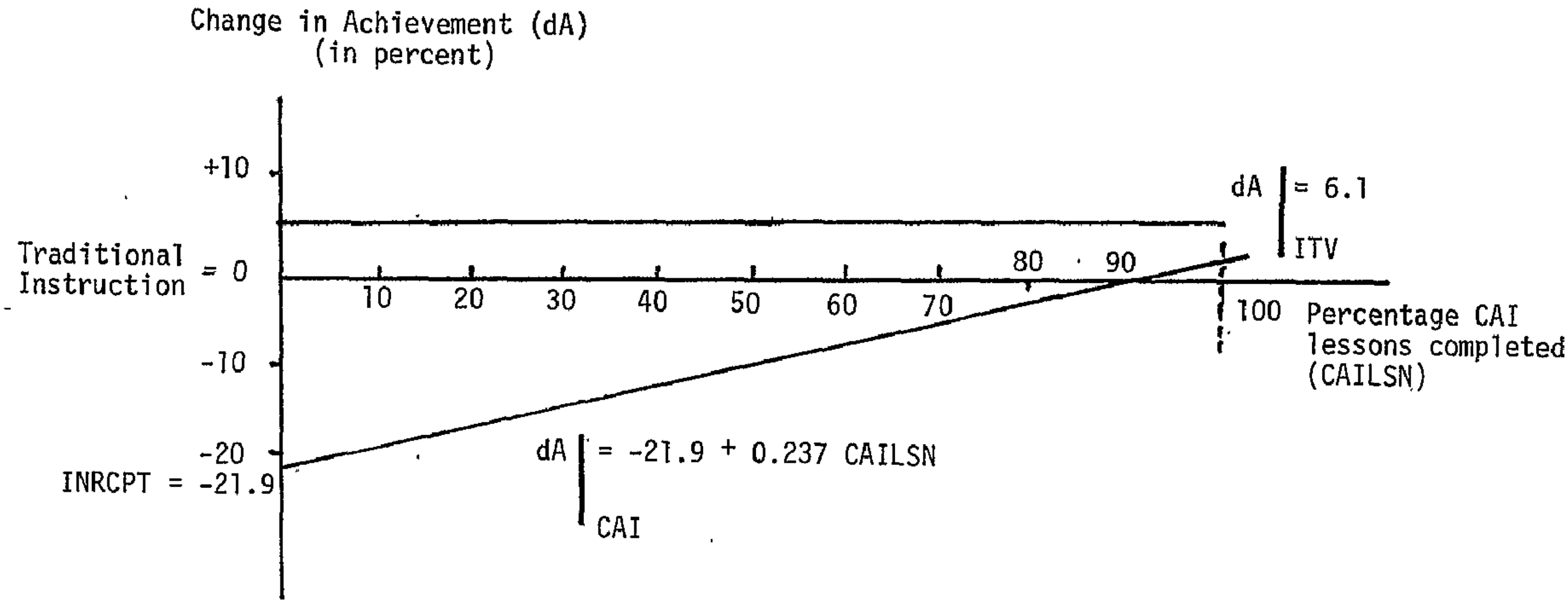
Figure 6.1 illustrates the regression results of Table 6.3 for the effect of different technologies on student achievement in accounting. Using traditionally educated students as a base, we find those subjected to ITV performed approximately 6.1 percent better on their examinations. Data describing variation in use of ITV across students were not available. The results for ITV simply imply an efficiency-improving shift of the production function for all ITV students.

The performance of the CAI student who neglects to use the computer (completes zero lessons) can be expected to fall almost 22 percentage

[6]This approach could not be pursued here because cost data were not available in sufficient detail.

Figure 6.1

THE EFFECT OF ITV AND CAI TECHNOLOGIES ON ACHIEVEMENT IN COLLEGE ACCOUNTING



points below the performance of the typical traditional student. As he completes more lessons, the student's achievement increases until, at approximately 90 percent lessons completed, the achievement of TI and CAI students is equivalent. The student who completes all 25 lessons in the system can be expected to reach a slightly higher level of performance, about 2 percent, than the traditional student. The differential between the INRCPT and CAILSN coefficients ranged from 1.2 to 2.7 percent in the linear model reported in Table 6.3. Just as Wells (1974) found the number of CAI sessions to be a significant determinant of methematics achievement of grade school children in California, CAILSN in this study demonstrates rather convincingly the importance of CAI use to student achievement in accounting.

These interpretations require some qualification, however. The sample for this study comprises students in first semester accounting at three different schools with three diverse technologies. In fact, ITV and INRCPT act as composite variables for both the change in school and change in technology. Moreover, the differences that have been attributed to the technologies could conceivably be caused by any of the factors, such as text, class size, or number of class meetings, that differ from school to school and were beyond control.[7] If it had been possible to collect data from several schools with the same technology, one could be more specific about separating the impacts of school and technology. Judging from the results above, such a study would be worthwhile.

---

[7]Consult Table 5.1 for a profile of the programs at the different institutions.

Nevertheless, through production function analysis, an unmistakable achievement differential was detected between schools. This may reflect a school effect, a technology effect, or some combination of the two. If we accept the conclusion from past research that the instructional medium has no impact on student performance, then the differential must be attributed to school effects. If we believe that "college quality" makes no significant difference in the level of student achievement, then fully utilized CAI is estimated as 2 percent, and ITV 6 percent, better than traditional instruction. Unfortunately, the data limitations of this study prevent determining which is true.

The discovery of a predictable technology/school effect must be slightly mitigated by another qualification: The differential may, in part, be generated by unique sample characteristics masquerading as school effects. First, the possibility exists that Colorado State's initial selection of questions for the output examination biased the results in favor of ITV students. Since the sample POOL is heavily dominated by students from Colorado State, the observed performance advantage for these students could be a consequence of an output measure prejudiced against the CAI and TI students.[8]

Second, in addition to having different technologies, texts, teachers, and class sizes, the schools have student bodies with different average abilities.[9] If we did not have a separate variable for student ability,

---

[8]Of course, this fact is not useful in explaining the differential between TI and CAI students.

[9]See the table of simple statistics for different samples in the appendix.

school dummy variables would have been one way to adjust for at least part of the difference in ability. If ability variables are omitted, a CAI student who completes all of the lessons shows a 10.3 percent higher achievement level than the traditional student. However, the differential falls to around 2 percent once performance are adjusted for SATM and SATV. Even with the three variables to measure student ability, it is possible that the higher ability of the ITV and CAI students relative to the traditional student caused an upward bias in the coefficients on ITV, INRCPT, and CAILSN in the analysis with the POOL sample. The estimates of simple models for various subsamples later in this chapter represent an effort to investigate the impact of these qualifications on the results.

Student Background Characteristics

Parents' education, a proxy for socio-economic status, was expected to have a positive coefficient. That expectation was based on the traditional human capital argument that education increases income, which in turn provides material advantages and a home environment conducive to learning. Moreover, more educated parents presumably take greater interest and invest more time in their children's academic development. EDUCF, years of education of the father, is the seventh input to enter the regression in Table 6.3, and the t-statistics stabilize at the 15 percent level of significance (two-tailed test). However, EDUCF reveals an unexpected negative sign. Years education of the mother, EDUCM, appears insignificant in this and most other regressions.

A possible explanation for the obverse sign on EDUCF involves assigning a new role to this variable. With the help of the simple diagram below, consider the connection between parents' education and student ability in this sample:

		Student Ability	
		Low	High
Parents' Education	Low	I	II
	High	III	IV

If the number of observations in cell II were significantly greater than cell IV, it would be possible for EDUCF to act as a proxy for student ability. If EDUCF plays this role, a negative sign would be reasonable. In any event, the distribution of student abilities among families in this sample as well as family income and the college selection process influence the partial effect of parents' education on student achievement. Therefore, a negative sign on EDUCF should not be quite so unexpected.

The variable, JOB, may have two interpretations. It may be a standardizing variable across students in its effect on the price of study time. Holding a market job affects the opportunity cost of study time, and the more hours worked, the more expensive a unit of study time becomes. Second, students who hold jobs and work many hours may reflect a low SES status. In this sense, JOB may be a proxy for student background. Either case would call for a negative coefficient. Since the coefficient on JOB is consistently positive in this study, one is inclined to offer still a third interpretation.

If market work has been substituted--not against study time, but against leisure time--JOB may be an indication of high ambition and personal motivation. To the extent that this dominates and the student is not working out of financial necessity, a positive coefficient is plausible. Even so, JOB was not a statistically significant determinant of student performance (35 percent level).

Father's occupation was used to construct a four-category set of interval variables. When the fourth category representing unskilled workers was dropped and the other three were added as regressors to the basic four-variable linear and log models (including SATM, ITV, INRCPT, and CAILSN), the results were decidedly insignificant. The increase in R^2 was 0.005 in the linear and 0.003 in the Cobb-Douglas model.

It really is not surprising that the variables comprising student background appear to be unimportant determinants of student achievement level. Astin (1968) made this observation in his popular article on undergraduate college quality. He postulated that it was due in part to the creation of a de facto track system through action of the college selectivity process. In a recent paper by Polachek, Kniesner, and Harwood (1975), parents' education was not only insignificant, but as in this study, years education of the mother appeared with a negative coefficient. In the same study, the coefficient on family income was 0.009 with a t-statistic of 0.8.

Student Effort and Motivation

To some extent, this serves as a catch-all category for variables not falling neatly into the other three categories. The negative coefficient on the SEX variable would normally mean that women can be expected to perform at a lower level in accounting than males, but little significance should be attached to that interpretation because of the low t-statistics.

LOAD is not particularly significant in either the linear or Cobb-Douglas model. A positive coefficient on LOAD would have called for an interpretation in terms of student motivation and ambition. A negative coefficient would reflect the influence of an increased course load on the opportunity cost of time spent studying accounting. MOTIV is the dummy variable declaring the impact of student response to the query, "Would you take another accounting course?" The coefficient on MOTIV is positive as expected, but only significant at approximately the 20 percent level.

STUDY has a positive effect on student performance, and is significant at the 10 percent level. The magnitude is not great, however--an extra hour of study per week is associated with a 1/2 percent gain on the final examination. This contrasts somewhat with the results obtained by others. Wells (1974) and Polachek, et al. (1975) both find a negative coefficient and get generally insignificant results with a measure of study time.[10] This variable's lack of significance is further supported in the results of Attiyeh and Lumsden (1972), Havrilesky (1971) and Paden and Moyer (1969). In view of the typical performance of this variable, a positive coefficient and t-statistic of 1.7 for STUDY are encouraging.

A priori, one would expect student experience in closely related disciplines to enhance his or her performance in accounting. However,

---

[10]In two cases, the impact of hours study was negative and significant at the 1 and 5 percent level.

CREDB, CREDM, and CREDE credits accumulated in business, mathematics, and economics are consistently the three poorest performers in the regressions. Since this information was collected by student response to one of the last items on the questionnaire, it is possible that the data are subject to inaccuracies and poor reporting. Nevertheless, the lack of significance pertaining to these variables agrees with the results of Moyer and Paden (1968) and Buckles and McMahon (1971).

With the possible exception of the STUDY input, the proxies for student effort and motivation are not among the more important variables of the production function.

Student Ability

The three measures of student ability--SATM, SATV, and RANK--are consistently among the major determinants of student performance in all regressions. Simple descriptive statistics for these measures are presented in Table 6.5. The correlation coefficients are not extremely high, but they are such that the regression coefficient on any one ability measure will be sensitive to the presence of the others in the equation. The coefficient on SATM in the linear model ranges from 0.043 to 0.056 and stabilizes at 0.045 in the presence of SATV and RANK. The coefficient on SATV varies from 0.024 to 0.028 and adjusts to 0.025 when all three variables are used. This says that a 100-point increase in SATM or SATV score can be expected to result in an average achievement gain of 4.5 percent or 2.5 percent,[11] respectively. These magnitudes are well within the limits of credibility.

---

[11]We should reserve judgment on these magnitudes until later when we consider interaction effects with partitions of the sample.

TABLE 6.5

DESCRIPTIVE STATISTICS OF ABILITY VARIABLES

Data Set: POOL N = 246

	Mean	Min	Max	St. Dev.
SATM	564	317	781	87.0
SATV	489	240	766	89.2
RANK	74.3	18	99	19.8

Simple Correlation Coefficients

	SATM	SATV	RANK
SATM	1.0		
SATV	0.477	1.0	
RANK	0.296	0.296	1.0

In similar studies, Oates and Quandt (1970) identified a range for the coefficient on SATM between 0.03 and 0.09. Emery and Enger (1972) observed a 0.05 on SATM, and Polachek, et al. (1975) found the marginal gain associated with SATM and SATV to be 0.033 and 0.026. In his study of undergraduate achievement and institutional quality, Astin (1968)[12] pronounced variations in ability to be the single most significant determinant of student achievement.

Constant marginal productivities are contrary to intuition, but it is unreasonable to expect a linear relation between student ability and test performance, or achievement. Aside from nonlinear functional forms like the logarithmic model, two direct methods within the OLS framework can be used to fit nonlinear forms for particular parameters. One is to substitute SAT and SAT squared for SAT; the other is to use a set of interval variables.

In Figures 6.2 and 6.3, the linear function is compared with the continuous nonlinear equations. Figure 6.2 plots the fitted values against SATM for the best four variable linear and nonlinear models. The nonlinear function, II, produces only a slightly better fit and does not depart appreciably from the linear form, I. It is interesting to note that the marginal contribution of a higher test score in mathematics is increasing.

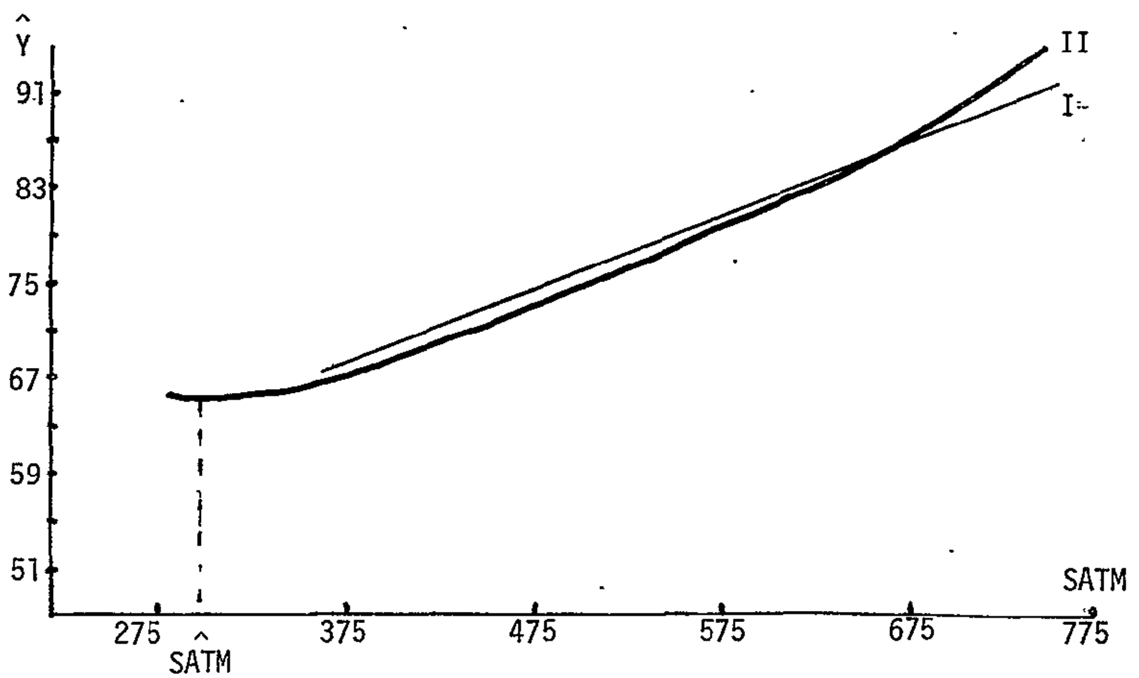
Figure 6.3 plots the fitted values against SATV for the best four variable linear and nonlinear functions. Again, the nonlinear fit is only a slight improvement, and the linear model is a good approximation

---

[12]His measure of ability was National Merit Scholarship Test scores.

Figure 6.2

EXPECTED GAIN IN ACCOUNTING ACHIEVEMENT AS A FUNCTION OF DIFFERENCE IN MATH SKILLS (CONTINUOUS FORMULATION)



I: $\hat{Y}$ = 39.6 + 6.6 ITV + INRCPT (-20.6 for 0.25 CAILSN)+0.059SATM

t stats: (2.9) (3.3) (3.14) (7.0)

R^2 = 0.246

II: $\hat{Y}$ = 71.4 + 6.8 ITV + INRCPT (-20.6 + 0.25 CAILSN)

(3.0) (3.3) (3.1)

- 0.061 SATM + 0.00011 SATMSQ

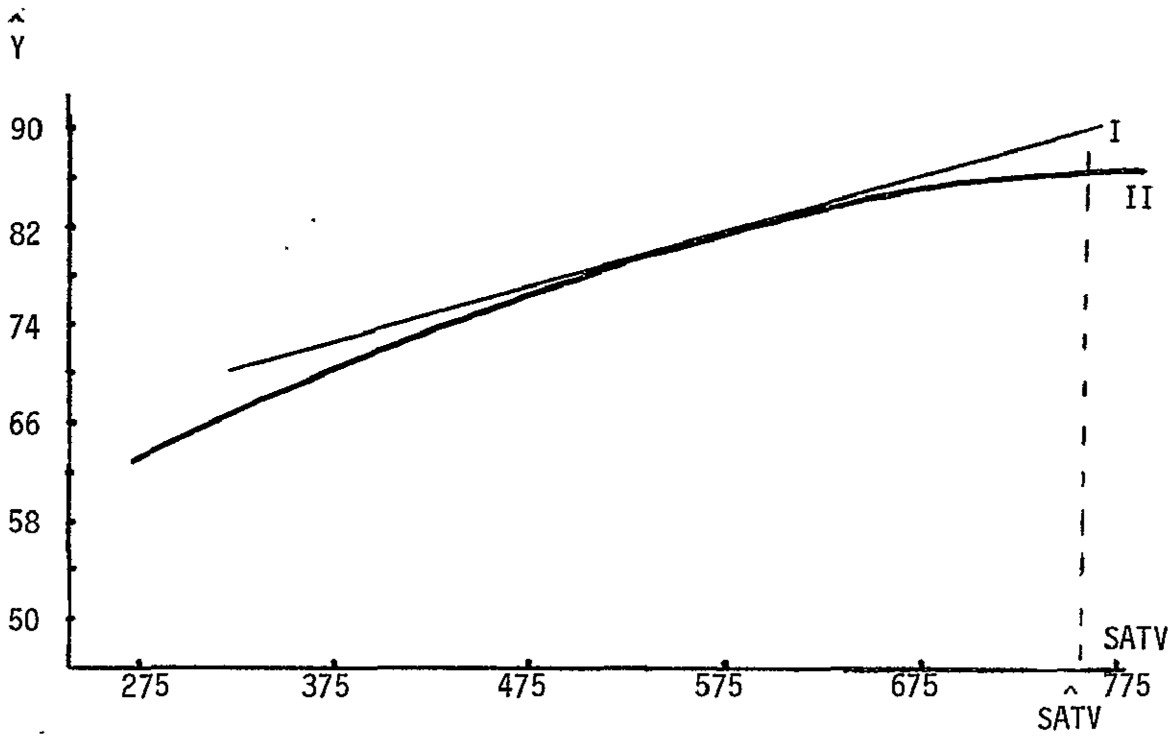
(0.8) (1.52)

R^2 = 0.253

$\hat{SATM}_{min}$ = 300 $\hat{Y}_{min}$ = 65

Figure 6.3

EXPECTED GAIN IN ACCOUNTING ACHIEVEMENT AS A FUNCTION OF DIFFERENCES IN VERBAL SKILLS (CONTINUOUS FORMULATION)



I: $\hat{Y} = 49.4 + 7.2\ ITV + INRCPT\ (-17.5 + 0.22\ CAILSN) + 0.046\ SATV$

t stat: (3.1) (2.74) (2.66) (5.5)

$R^2 = 0.193$

II: $\hat{Y} = 28.45 + 7.1\ ITV + INRCPT\ (-17.2 + 0.224\ CAILSN)$

(3.0) (2.7) (2.68)

$+ 0.136\ SATV - 0.00009\ SATVSQ$

(2.05) (1.36)

$R^2 = 0.199$

$\hat{SATV}_{max} = 755 \qquad \hat{Y}_{max} = 86$

over the large part of the SATV range. However, in this case, the marginal contribution from an increased verbal test score is declining.

These relationships have the following implications: For students with unusually low ability in both areas, verbal aptitude seems more important than mathematics for achievement in accounting. For all other students, mathematics ability overshadows the importance of verbal ability. In addition, the difference will be greater the higher the ability level of the student.

The nonlinear relation between SATM, SATV, and achievement was also investigated by using a set of interval variables to generate a step function in place of the continuous one. Table 6.6 displays the regression results. Variables 4 and 5 were created by sorting SAT scores by rank and dividing the sample as nearly as possible into equal quintiles. MATHQ1...MATHQ5 are a set of one-zero dummys indicating whether the student is or is not a member of the lowest through the highest quintile. The same applies to the SATV scores, where VERBQ1...VERBQ5 are the interval variables. To facilitate the regression, the middle quintiles--MATHQ3 and VERBQ3--were omitted.

The simple Model 1 is included in Table 6.6 to show the gain in R^2 once the mathematics and verbal ability variables are included. Model 2 is the step function analogue to the equation lebeled "II" below Figure 6.2, and Model 3 is the analogue to the continuous function "II" below Figure 6.3.

The fitted values for the interval Models 2 and 3 are graphed against SATM and SATV in Figures 6.4 and 6.5, respectively. This serves as a check on the results obtained with the continuous equation, nonlinear

TABLE 6.6

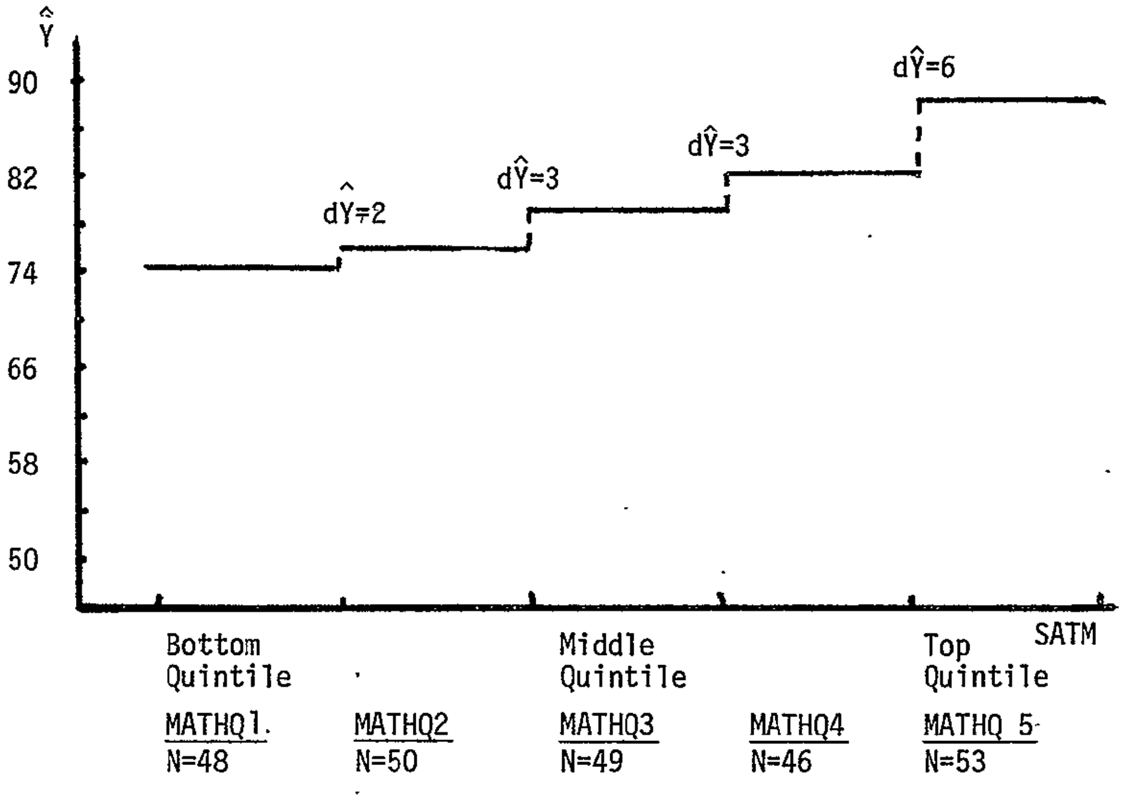
DETERMINANTS OF PERFORMANCE IN COLLEGE ACCOUNTING WITH INTERVAL FORMULATED ABILITY VARIABLES*

DSN = POOL: N = 246	MODEL			
Variable	Model 1	Model 2	Model 3	Model 4
1. ITV	8.35 (3.3)	7.21 (3.14)	6.16 (2.59)	4.83 (1.85)
2. INRCPT	-15.3 (2.26)	-20.35 (3.25)	-18.64 (2.92)	-22.50 (3.67)
3. CAILSN	0.255 (2.90)	0.26 (3.17)	0.24 (2.89)	0.244 (3.1)
4. MATHQ1		-4.83 (2.1)		-2.78 (1.18)
MATHQ2		-3.71 (1.66)		-1.81 (0.80)
MATHQ4		2.94 (1.29)		2.86 (1.28)
MATHQ5		8.99 (4.0)		7.83 (3.57)
5. VERBQ1			-8.49 (3.74)	-7.23 (3.29)
VERBQ2			-0.88 (0.40)	-1.41 (0.66)
VERBQ4			4.87 (2.1)	1.76 (0.77)
VERBQ5			2.83 (1.23)	-0.68 (0.29)
6. RANK				0.08 (2.05)
7. STUDY				0.37 (1.20)
8. LOAD				0.16 (0.57)
9. MOTIV				1.41 (0.87)
10. EDUCF				-0.27 (1.17)
CONSTANT	70.82	71.49	73.50	67.89
R^2	0.09	0.245	0.213	0.322

* Dependent variable equals the percentage correct on identical examination; t-statistic in parentheses. MATHQ1..MATHQ5 and VERBQ1...VERBQ5 are interval variables for the sample of SATM, SATV scores divided into quintiles with 1 the lowest and 5 the highest.

Figure 6.4

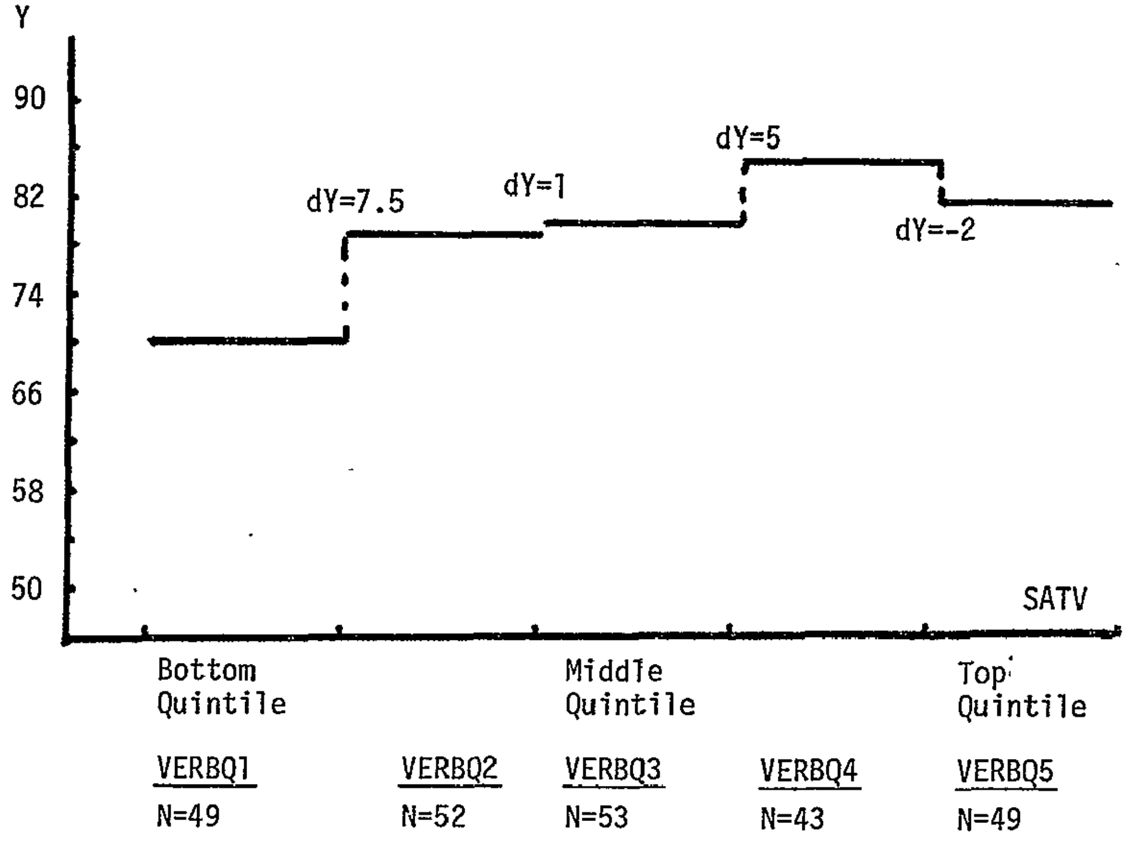
EXPECTED GAIN IN ACCOUNTING ACHIEVEMENT AS A FUNCTION OF DIFFERENCES IN MATH SKILLS (INTERVAL FORMULATION)*



* Plot of fitted values against SATM for Model 2, Table 6.6

Figure 6.5

EXPECTED GAIN IN ACCOUNTING ACHIEVEMENT AS A FUNCTION OF DIFFERENCES IN VERBAL SKILLS (INTERVAL FORMULATION)*



* Plot of fitted values against SATV for Model 3, Table 6.6.

model. Figures 6.4 and 6.5 confirm the general shape of the curves in Figures 6.2 and 6.3, respectively.[13] The increasing function for mathematics scores is clearly evident as is the decreasing relationship with respect to verbal score.

Regression Results for Different Samples

Earlier sections were concerned with the results obtained for different functional forms. This section turns to the stability of the results described over different samples. In particular, 190 of the 246 observations in data set POOL are from the ITV technology and only 29 from the CAI technology. More confidence could be placed in the results if the coefficients that were significant for data set POOL remain stable for other samples.

Table 6.7 reports the regression coefficients for a simple ability model (including the RANK variable) for each interesting sample subset[14] (columns 2 through 10). Coefficients on the ability variables change substantially when regressions are estimated separately for each technology. In two independent tests of the traditional school (columns 2 and 3), SATV proved superior to SATM but both samples at the CAI school (columns 5 and 6) show the opposite. Not only were coefficients on SATV for CAI unimportant, they were negative. As expected, coefficients on SATM and SATV for the ITV school retain roughly the same values displayed

---

[13] Computer scatter plots corresponding to Figures 6.2, 6.3, 6.4 and 6.5 are available.

[14] LECT1, LECT2, CAI1, CAI2, and ITV are the five separate data collections. LECTUR and CAI are the total observations from the traditional and computer-assisted technologies, respectively. POOL is the composite set subject to the experimental exam and ADD is the sum of two collections that were not.

TABLE 6.7

DETERMINANTS OF STUDENT ACHIEVEMENT: A SIMPLE MODEL

		Traditional			CAI			ITV		
	1 POOL = LECT 2 + CAI 2 + ITV	2 LECT 1	3 LECT 2	4 LECTUR = LECT 1 + LECT 2	5 CAI 1	6 CAI 2	7 CAI = CAI 1 + CAI 2	8 ITV	9 ADD = LECT 1 + CAI 1	10 POOL = LECT 2 + CAI 2 + ITV
SATM	.044 (4.7)	.025 (1.3)	.016 (0.4)	.020 (1.19)	.046 (2.44)	.095 (2.0)	.059 (3.3)	.042 (4.2)	.044 (3.5)	.039 (3.9)
SATV	.024 (2.6)	.073 (2.1)	.050 (1.3)	.065 (3.6)	-.005 (0.3)	-.030 (0.8)	-.014 (0.9)	.034 (3.5)	.023 (1.7)	.020 (2.05)
RANK	.06 (1.7)	-.001 (1.6)	.08 (.8)	.04 (.6)	0.1 (0.8)	.002 (.008)	.073 (0.6)	.057 (1.43)	.007 (0.9)	.055 (1.4)
ITV	5.9 (2.6)									
INRCP	-21.8 (3.5)									
CAILSN	0.238 (3.0)									
CONST (SE)	31.8	24.1 (12.0)	33.3 (14.0)	28.05 (8.8)	41.3 (12.5)	29.7 (31.6)	39.4 (12.1)	35.2 (5.7)	36.8 (8.2)	41.6 (5.4)
R^2	.28	.51	.26	.38	.13	.13	.13	.24	.22	.16
N	246	26	27	53	66	29	95	190	92	246
Means and Standard Deviation of Variables for All Five Data Collections: SATM		530 (120)	530 (95)		630 (77)	624 (66)		560 (85)		
SATV		470 (104)	460 (90)		505 (88)	555 (92)		485 (85)		
CAILSN					92.7 (12)	67.0 (26)				
OUTPUT		71.5 (13)	70.8 (12)		76.4 (11)	72.7 (15)		79.2 (11)		

* The t-statistics appear in parentheses for explanatory variables; the standard error appears in parentheses for the constant term.

by POOL. It is most interesting that the impact of different skills depends on the technologies involved. These interactions will also have an effect on the constant term. Since the shift of these constants determined the effect of the technologies, a re-examination is warranted.

It will be useful to compare the implications of the technology variables in the POOL analysis in column 1, where no interactions are allowed, with the constants actually observed when interactions are allowed (cols. 3, 6 and 8). The results of this comparison are reported in Table 6.8. As you can see, the mean performance of traditional students is 1.5 percent higher (33.3 percent versus 31.8 percent) when interactions are allowed. This obviously accounts for part of the change in the relative effectiveness of the technologies. The coefficients on the technology variable in sample POOL predicts a mean adjusted performance of 37.7 percent for the television group. The actual figure is 35.2 percent. Thus, because of the interactions, the shift variable calculated in the POOL analysis overestimates the achievement level associated with ITV by 2.5 percent (37.7 percent less 35.2 percent), and the relative effectiveness of ITV, formerly 5.9 percent, reduces to 1.9 percent (35.2 percent less 33.3 percent).

Based on the POOL analysis coefficients, the mean achievement of CAI students calculated at 67 percent lessons completed should be 26.0 percent. The actual achievement level was 29.7 percent or 3.7 percent higher than expected. The relative effectiveness of CAI depends on the average percent of lessons completed. Since the 92.7 percent of lessons completed for the CAI 1 group is based on more than twice as many observations taken during the fall term as opposed to summer school, the true

TABLE 6.8

MEAN ACHIEVEMENT BY TECHNOLOGY
ADJUSTED FOR DIFFERENCES IN ABILITY
(percent of output measure)

	Without Interactions[1]	With Interactions[2]
LECT2	31.8	33.3
ITV = 31.8 + 5.9 =	37.7	35.2
CAI2(@ 67% CAILSN) = 31.8-21.8 + 16.0 =	26.0	29.7
(@ 92.7% ") = 31.8-21.8 + 22.0 =	32.0	(35.7)
(@100.0% ") = 31.8-21.8 + 23.8 =	33.8	(37.5)

---

[1]These are calculated using the relationships implied by the coefficients in column one: $\overline{ITV} = \overline{LECT} + 5.9$; $\overline{CAI} = \overline{LECT} - 21.8 + .238\,(\overline{CAILSN})$.

[2]These are the constant terms in the regressions for the partition of sample POOL in columns 3, 6, and 8. The terms in parentheses are the approximate levels of performance we can expect if groups with a higher percentage of lessons completed do at least as well as groups with fewer lessons completed.

mean CAILSN is probably closer to 92 percent than 67 percent. If the impact of lessons completed is either not affected or increased by the presence of ability-technology interactions, then mean achievement for CAI will be at least 35.7 percent. That is, at 92 percent CAILSN, the CAI students would perform at a level 2.4 percent higher than the traditional student.

Comparisons of the two regressions for subsets within traditional and CAI schools show substantial differences in performance on the experimental versus the local, nonstandardized examinations in each school. The traditional students subject to the local examination (LECT1) fall short of the performance by students taking the experimental examinations (LECT2). Nevertheless, student achievement on the class final at the CAI school (CAI1) far exceed the performance of CAI students on the experimental examination (CAI2). It is apparently a coincidence that the union of two samples with such different parameters as LECT1 and CAI1 forms a composite set, ADD, that produces regression coefficients remarkably close to those for POOL.

The presence of interactions between the ability and technology inputs discredits the coefficients estimated in the pooled analysis. When the samples are analyzed separately, both traditional and CAI students do better than ITV students, causing the relative effectiveness of ITV to be smaller and CAI to be larger than originally expected. It now appears that ITV and CAI are each associated with a 2 percent gain in performance and CAI has potential to do slightly better.

The caveats from the earlier pooled analysis apply here, with one important addition. In the process of partitioning a sample to look at variables in more detail, degrees of freedom are sacrificed. As a result, much of the confidence in the magnitude of the regression estimates may be lost. A glance at the standard errors associated with the crucial regression constants reveals the gravity of this problem. It is obvious that the confidence intervals for the intercepts at the 99 percent level involve a considerable overlap. Although the coefficients may subtly hint at differences between the technologies, a statistically significant difference cannot be identified.

In summary, the educational production function estimates of this chapter have several noteworthy features. One is congruity. The interpretation of results for any one variable was consistent with events postulated in the rest of the equation. Second, although the Cobb-Douglas model has a more plausible interpretation, the linear model proves equally useful.

More important, the results show that 30 percent of the variation in student achievement in college accounting can be explained by a small set of variables representing differences in student ability, school, or technology. In some cases, the lack of significance for variables was almost as interesting. Student background factors failed to make a significant showing, study time attained only borderline recognition, and college credits in related fields were totally unimportant. For the most part, the results above were similar to those prevailing in the literature.

In terms of specific inputs, the most remarkable findings were the persistence of SAT scores, their nonlinear effects, and their different impacts for separate technologies. The interactions imposed a downward revision in the effectiveness of ITV to about 2 percent and suggested slightly larger effectiveness for CAI. However, the process of partitioning the sample cut the power of statistical inference to the point where only the common "no significant difference" result can be supported with confidence. Although we suspect a systematic difference between technologies, additional testing with larger samples would be required before such a claim could be supported.

The next chapter confronts the problems of measuring system costs.

CHAPTER 7

COST-EFFECTIVENESS ANALYSIS

Most evaluation of technology in higher education has concentrated on the comparative impacts of different systems on student achievement. However, educational planning should not be based on those results alone. Cost analysis is half of the cost-effectiveness issue.

We previously described cost effectiveness as a special application of the conditions for optimal factor proportions from the theory of production. But, rather than solving for the optimal proportion of resources, cost effectiveness seeks to determine the most economic production procedure or specification of feasible alternatives. For example, one may employ cost-effectiveness analysis to resolve whether shaft or strip mining is the more economic means of coal extraction. This is obviously different from defining the optimal combination of trucks and shovels. The specification yielding the highest effectiveness per dollar cost is pronounced the most cost effective. Since effectiveness is a relative measure, we express the cost-effectiveness condition as,

(7.1) $$\frac{E_i}{E_j} \gtreqless \frac{C_i}{C_j}$$ for alternative procedures i, j.
(if = holds, each is equally cost effective).

The measure of relative effectiveness from the production function study is compared with the relative unit costs of the alternatives. If we interpret the previous chapter conservatively, no identifiable difference is seen in the effectiveness of the three technologies. In this case, the most cost-effective technology is the one with the lowest unit cost.

If on the other hand, we operate on the premise that ITV and CAI are more effective than traditional instruction, the most reasonable calculations estimate both CAI and ITV, to yield about 2 percent higher performance levels. In this situation, if the cost of either CAI or ITV is less than TI, then the most cost-effective of the three technologies is still the one with the lowest unit cost. Thus, the outcome of the debate on the ability to make statistical inferences in the effectiveness analysis has no bearing on the choice of the cost-effective technology.

Before we attempt to estimate the unit cost for each of the three technologies, it will be helpful to explain the different cost concepts.

The Concepts of Cost

It is generally useful to think of total, average, and marginal costs as algebraic functions rather than numbers. The total cost of an input is the monetary cost of providing the input as a function of the amount required. In a simple case,

(7.2) $TC = TC(N)$,

$TC(N)$ is the cost required to provide a unit of education to N students.

Although cost functions dealing with production functions usually refer to the direct monetary cost of inputs, that is not the only relevant cost concept. One could speak of full cost as meaning not only the direct monetary cost to the agents involved but also the indirect or opportunity cost for inputs that are valuable but not entered in the budget constraint. An example is the opportunity cost of student and faculty time. Educational planners should regard time costs as a decision variable because a gain in efficiency in terms of time savings is just as real as a monetary gain.

Another pertinent concept is <u>social cost</u>, which encompasses the impact of production activity on individuals and resources not directly involved. It includes the net costs of production externalities (like pollution) as well as the costs to persons involved in the process (like injury at work). Social cost would be the relevant concept for a discussion of the equity effects of education.

Returning to our description of cost functions, the <u>average</u> or <u>unit cost</u> function is equal to the total cost divided by the number of units provided:

(7.3) $AC(N) = TC(N)/N.$

The <u>marginal cost</u> of an input is the added cost of providing one more unit of input as a function of the number currently provided. More precisely, it is the first derivative of the total cost function:

(7.4) $MC(N) = dTC(N)/dN.$

Input cost functions normally depend on only one variable because other factors have implicitly been held constant. In educational applications, however, this may be unrealistic, especially where planning and evaluation are concerned. The educational planner may have other constraints or objectives aside from the number of students served. This will complicate the structure of the cost functions. For example, we may want to build a CAI or ITV system to serve N students with a capacity for H hours per month and the ability to transmit K miles from the facility with user density, D. Total cost is

(7.5) $TC = TC(N, H, X, K, D)$,

where N is students, H is hours per month, X is number of channels or program capacity, K is miles, and D is user density. The marginal cost is the amount change in total cost for a one unit change in each of the five explanatory variables. Some of the average cost measures with intuitive appeal are cost per student (TC/N), cost per program or channel (TC/X), cost per hour (TC/H), and cost per student contact hour (SCH = TC/NH).

Another important cost concept is the distinction between fixed and variable costs. The time structure and length of production planning period are of prime importance. If we assume a simple linear cost function, total cost is the sum of fixed and variable costs:

(7.4) $TC(N) = F + VN$.

Fixed costs are those that cannot be altered in the short run or normal production period. Variable costs are the inputs within the discretion of the decision-maker in each production period. These distinctions are particularly important regarding the costs of education. Traditional education as well as such technologies as CAI and ITV typically involve a large initial capital investment. Before long-range decisions are made, all costs associated with technology may be variable.[1] However, after the initial capital investment, these become unrecoverable sunk costs.

When the cost function is linear, the average cost is simply equal to the fixed cost divided by N plus the variable cost:

(7.7) $AC(N) = F/N + V.$

The average cost of educational technology declines as enrollment increases, distributing the fixed costs widely across users until the average cost approaches marginal cost, when N is quite large.

Another source of confusion centers around the concept of the cost of capital. Capital refers to an input whose useful life extends beyond the usual production accounting period. All other inputs may be classified as recurrent costs. To some extent, this dichotomy is arbitrary. Labor in the form of teacher time is treated as a recurrent cost. So are materials and supplies like pencils and paper, which could technically be viewed as capital goods. A broadly accepted convention is to classify inputs with a lifetime under a year as recurrent costs and those longer than a year as capital costs. Clearly, such items as TV cameras, amplifiers, videotaping equipment, antennas, and computers are capital goods.

One important problem is the valuation of capital in cost calculations. The cost of a transmitter should certainly not be charged to the period in which it is purchased. Two factors must be taken into account. One is the depreciation cost over the expected lifetime of services, and

---

[1]Except for imputed costs for fixed overhead like classroom and building space and possibly some administrative costs.

the other is the interest cost to society of sinking resources into a piece of capital for its lifetime. In other words, the interest cost is the price one must pay to reduce current consumption to generate future output. The cost of capital must be amortized by the social discount rate to reflect this cost of tying up resources. The average annual capital charge is computed according to the following formula:

$$(7.8) \qquad \text{Annualized Cost} = \frac{P_k(1 + r)^n}{(1 + r)^n - 1} ,$$

where P_k is the price of capital of lifetime n discounted at the rate r.

Last is the issue of system cost incidence. From a resource allocation perspective, one would not want to take cost incidence into account. Hence, it does not enter the cost-effectiveness calculations. However, a significant portion of the initial cost of major educational projects (and possibly a smaller portion of recurrent costs) are financed by government agencies and foundation grants. "N.S.F. grants connected with PLATO amounted to $1 million or more between 1968 and 1972. Then a new N.S.F. contract added $7.8 million to the public investment in the system."[2] To be sure, many of the ambitious projects in educational innovation would never have begun, regardless of their expected cost effectiveness, were it not for the financial intervention of a benevolent agency or foundation. The tax incidence of education technology finance and distribution effects pose some interesting questions beyond the scope of this discussion.

---

[2] Chronicle of Higher Education, (April 26, 1976) p. 1.

The multiple input cost function discussed above is symptomatic of the variety of technology systems encountered in higher education. Designed to suit the educational needs of users in various situations, different systems may provide different services as well as exhibit differences in quality for the same service. One cannot speak of the cost of ITV or CAI. Consequently, a discussion of cost effectiveness is complicated and frequently dealt with in a shotgun fashion by researchers. A thorough discussion of the cost of technology in higher education is beyond the scope of this paper. Our goal will be to point out the especially important cost considerations, draw general conclusions where possible, and estimate the unit costs for the technologies involved.

We are not attempting to analyze the total cost of presenting an accounting course. Instead, only those costs directly attributed to the delivery of instruction will be examined. Components making up the system are included, but costs of overhead items like classrooms, maintenance, and school administration are not.

The cost calculations reflect only those items that would change when a three-credit ITV or CAI course is substituted for a similar traditional course. We assume that the conventional course includes two lectures per week and one discussion section conducted by a graduate assistant. At Colorado State and the University of Illinois, the class monitoring and paraprofessional services necessary to the ITV and CAI systems are provided by graduate assistants who would otherwise lead a traditional discussion (or lab) session. Assuming this is an even trade, the cost of paraprofessional labor need not enter the calculation of relative costs.

Cost of Traditional Instruction (TI)

For traditional instruction, the marginal cost per credit is simply the yearly salary of the professor with a full-time teaching load divided by the number of course-credits of instruction per year.[3] Dividing by average class size gives the cost of instruction per student-credit.

Differences in class size, teaching load, and salaries cause considerable variation in the costs of instruction for different schools and professors. In reality, the cost per credit is represented by a distribution of values. To estimate the cost, some reasonable, yet arbitrary, assumption must be made for teaching load and salary level. Since we want to place the burden of proof on the innovative technology, it is appropriate to select values that exert a downward bias on the cost of traditional instruction.

If a university similar to those represented in this study pays a minimum of $15,000 salary over nine months to a faculty member in accounting, and if we assume a full-time semester load of 12 hours, this results in $625. per course credit. Since it is unlikely that the teaching burden would be any greater or the pay much lower than the figures we have assumed, this is a conservative figure. However, if we perform the same calculations assuming a nine-credit load per term, the result is $834 per course-credit. For a class of 20, 30, and 40 students this amounts to $31, $21, and $15 per student-credit under the 12-hour load assumption and $42, $28, and $21 with a nine-hour load.

---

[3]We acknowledge that faculty services include research and administrative duties, but we take the position that the costs of such activities must be justified on their own grounds. For the purpose of instructional cost calculations, we assume that a faculty member teaches full-time.

There are numerous cases in the literature where researchers assume or calculate considerably higher figures. McKeown (1974) states that the "minimum cost for conventional instruction in these courses at the University of Illinois at Urbana-Champaign exceeds $1.35 per student contact hour."[4] In a 15-week semester, this says that the minimum cost for a traditional accounting course at the University of Illinois is $60.75 per student-credit. Lionel Baldwin, Dean of the CSU College of Engineering, views "$65 per quarter credit as a conservative estimate for graduate courses in engineering, and averages for all engineering courses were $37.50, the comparable figure for 71 engineering colleges was $49."[5] A study by Wilkinson (1972) bases instruction costs for a traditional system on "12 professors teaching 30 sections of a course to 1,000 students for an annual cost in salaries of $300,000."[6] For a three-credit course, this is $100 per student-credit. Wilkinson assumes a teaching load of 4.5 credits per semester, a salary of $25,000, and a class size of 34 students.

The cost estimates in the literature from $37 to $100 per student-credit make it appear that our assumptions result in unusually low cost figures. Perhaps our estimates should be inflated by assuming a smaller teaching load or a higher salary. We choose instead to retain the minimum cost estimates to force robuse results for the cost-effectiveness comparisons.

---

[4]McKeown (1974), p. 19.

[5]Baldwin, p. 187.

[6]G.L. Wilkinson, (July 1972), p. 38.

Costs of Instructional Television (ITV)

The costs of communications technology systems are usually calculated as add-on costs to a traditional system. However, this analysis examines only the cost of substituting a three-credit ITV or CAI course for a traditional course. Hence, administrative costs and building and physical plant facilities unrelated to the communications system are exogenous to the cost analysis. Presumably, these would be the same regardless of which method delivered the instruction.

Costs are broken down by component to clearly show where the major cost differences lie. The cost of ITV runs the gamut from a simple closed circuit technology, which could be started with a minimum investment of $50,000, to complex systems linking several schools, which could run into the millions. Extensive off-campus networking could also be incorporated into the system. Not only does this complicate a discussion on the costs of ITV, it encompasses systems with widely ranging services and great differences in the quality of services. Our comments on costs will be limited to an on-campus, ITV system at the university level commensurate with the one that participated in the effectiveness portion of this research.

One should keep in mind that low unit costs are achieved by a large-scale system and by operating facilities at capacity. To some extent, the results of a cost-effectiveness analysis can be predetermined by designing a large-scale system that spreads the overhead across a sufficient number of users.

Because we were unable to collect information on the precise facilities serving the accounting course examined, cost data were adapted from reported figures for another Colorado State University ITV project. A committee report by the American Society of Engineering Education on the Cost-Effectiveness of the ITV Continuing Engineering Studies[7] gives a full account of the costs associated with the off-campus graduate engineering program at CSU known by the acronym SURGE. Estimates compiled by Lionel Baldwin report the costs of videotaped production and distribution for the SURGE program in 1972-73 costs. The unit cost of ITV is calculated by first estimating the cost per programming hour, then deriving the cost per course credit, and finally the cost per student-credit.

Although both projects took place at the same institution, several important differences can be seen. The SURGE program taped live, on-campus lectures of graduate engineering for distribution to remote locations for viewing, later to be returned and reused. Television was used to transcend the geographical and time barriers to transmit more or less conventional classroom lectures. By contrast, the on-campus accounting sequence involved two 25-minute TV lectures per week for a three-hour course, with discussion led by a graduate assistant consuming the other two hours per week. Instead of taped classroom experiences, these were produced TV lectures, often by different professors lecturing on their specialties. The lectures were revised and updated at the rate of one-third per year.

The Morris Report describes a system capable of producing taped lectures at the rate of 30 hours per course for 200 courses for a total of 6,000 lecture hours or 9,450 delivered tapes per year in 315 sections.

---

[7]Hereafter referred to as the "Morris Report" after the Committee Chairman.

Allowing an hour for actual taping, and an hour for rehearsal and editing, we assume that 6,000 hours of taped classroom lectures translate to a minimum of 3,000 hours of videotaped lectures. Furthermore, for simplicity, we assume that the two 25-minute lectures per week (Monday and Wednesday) are together equivalent to an hour of produced programming per week. Although CSU is in a quarter system, most schools are not. Hence, calculations are performed for a semester system. This amounts to 15 hours total videotaped accounting lectures per course. If one third of the lectures are renewed and updated per year, five hours of new programming per year will be required for each course. Therefore, if the resources for SURGE can produce 3,000 hours of videotaped lectures per year for on-campus use, the system could theoretically support 600 courses per year, or 1,800 course-credits.[8]

Production costs for the on-campus ITV services are the sum of costs for production studio, recording facilities, tape, and operating costs. Under the assumptions above, the costs incurred for studio and operating costs in SURGE are capable of producing the required tape replacement for 600 courses per year. Since recording facility costs need only be proportional to the number of tapes actually made, those for the on-campus system were figured at one-third of what was purchased for the SURGE system. (SURGE taping facilities produce 9,450 tapes per year for 315 course sections.)

[8]This assumes that each of the 600 courses are only offered once a year. If offered more than once or to more than one section of students at the same time, ITV costs are dispersed over more users and the system becomes more cost effective.

In addition to production costs, the cost of the on-campus system includes playback facilities and classroom modifications. Together, these constitute the reception costs. Instruction costs are also included for the on-campus system. Although technically production costs, they are intentionally included at the end to stress their importance. The amount for instructional costs is somewhat arbitrary and alternate measures could be incorporated with drastic effects on the results.

The so-called production costs are summarized in Table 7.1. Each category is fully itemized in an appendix to this chapter. Notice that, while the total outlay for capital equipment is sizable ($117,750), its capitalized value is only $19,040, which is small compared with the annual cost of labor required (operating cost = $81,300). Furthermore, even if the cost for instruction was a modest $100 per recorded lecture hour, it would imply an investment of $300,000 per year. It is not unreasonable to imagine an instruction cost figure twice as high. The total investment in videotape alone would amount to $240,000. If the program began at the same level of expenditures budgeted for the tape replacement operation, only one-third of the required lessons could be taped in each of the first three years. This amounts to 3,000 tapes, or $60,000 purchased in each of the first four years. The bottom line of Table 7.1 shows that a one-hour, videotaped lecture could be produced for $33.70, excluding the cost of instruction and instructional support cost.

Table 7.2 presents the total costs of ITV per course in terms of costs associated with the five hours of required tape replacements and the costs of playback and reception of the 15 total hours of ITV per course.

TABLE 7.1

ITV PRODUCTION COSTS[1]

Capital Costs	Costs[2]	Depreciation	Annual Capital Charge
1. Studio classrooms and consoles (3), $30,000 each	$ 90,000		
2. Interconnect between classrooms and master control	$ 3,500		
Subtotal	$ 93,500	10 yrs at 6%	$12,697
Recording Facilities			
3. VTR	$ 12,380	3 yrs at 6%	$ 4,731
4. Other equipment, cable, switching, storage, but no video tape	$ 11,870	10 yrs at 6%	$ 1,612
Subtotal	$ 24,250		
Total Capital Cost	$117,750		$19,040
Investment Cost in Video Tape			
5. Tape, 15 on-line plus 5 replacement = 20 tapes/course x 600 courses = 12,000 tapes; $20/tape (60 minute). Total value of tape inventory (tape cost included below as $/hr)[3]	$240,000		
Total Capital Investment, Equipment and Inventory	$357,750		
Operating Costs			
6. Base operating costs per year			$81,300
Total System Production Costs for 3,000 hrs taped ITV			$100,340
Total System Costs per Taped Hour		$33.50	
Cost of Tape per Hour ($20/hr ÷ 100 uses)		0.20	
Total Production Costs per Hour		$33.70	

---

[1]Source: A.J. Morris, "Final Report on the Cost-Effectiveness of Continuing Engineering Studies by TV." (Mimeo, American Society of Engineering Education, 1974).

[2]For a breakdown of costs by item, see appendix to this chapter. Item 1, Table A7.1a; items 3 and 4, A7.1b, item 6, A7.1c. Costs pertinent to 1972-73 academic year.

TABLE 7.2

TOTAL COSTS OF ITV

	Cost Per Taped Hour	Total Cost of Tape Replacement per Course (5 hrs of new tape/course	Total Cost of Playback & Reception per Course (15 hrs of ITV/course)	Total Cost of ITV Per Course
Production Costs				
System Costs	$ 33.70			
Instruction Costs				
Instructional Cost per[1] hour taped ($15,000, 1/4 time = 1 lecture per week)	$125.00			
Instructional Support Cost per hour, including drafting, slides, art-work, etc.	10.00			
	$168.70	$843.50		
Reception Costs[2]				
Playback costs	$ 13.00			
Classroom costs	$ 0.75			
Cost of tape	$ 0.20			
	$ 13.95		$209.25	
				$1,053
			Cost of ITV per course credit	$351

[1]This is equivalent to substituting one TV lecture per week for a traditional three-hour course.

[2]For breakdown of reception costs, see appendix tables A7.2a, A7.2b. Cost pertinent to 1972-73 academic year.

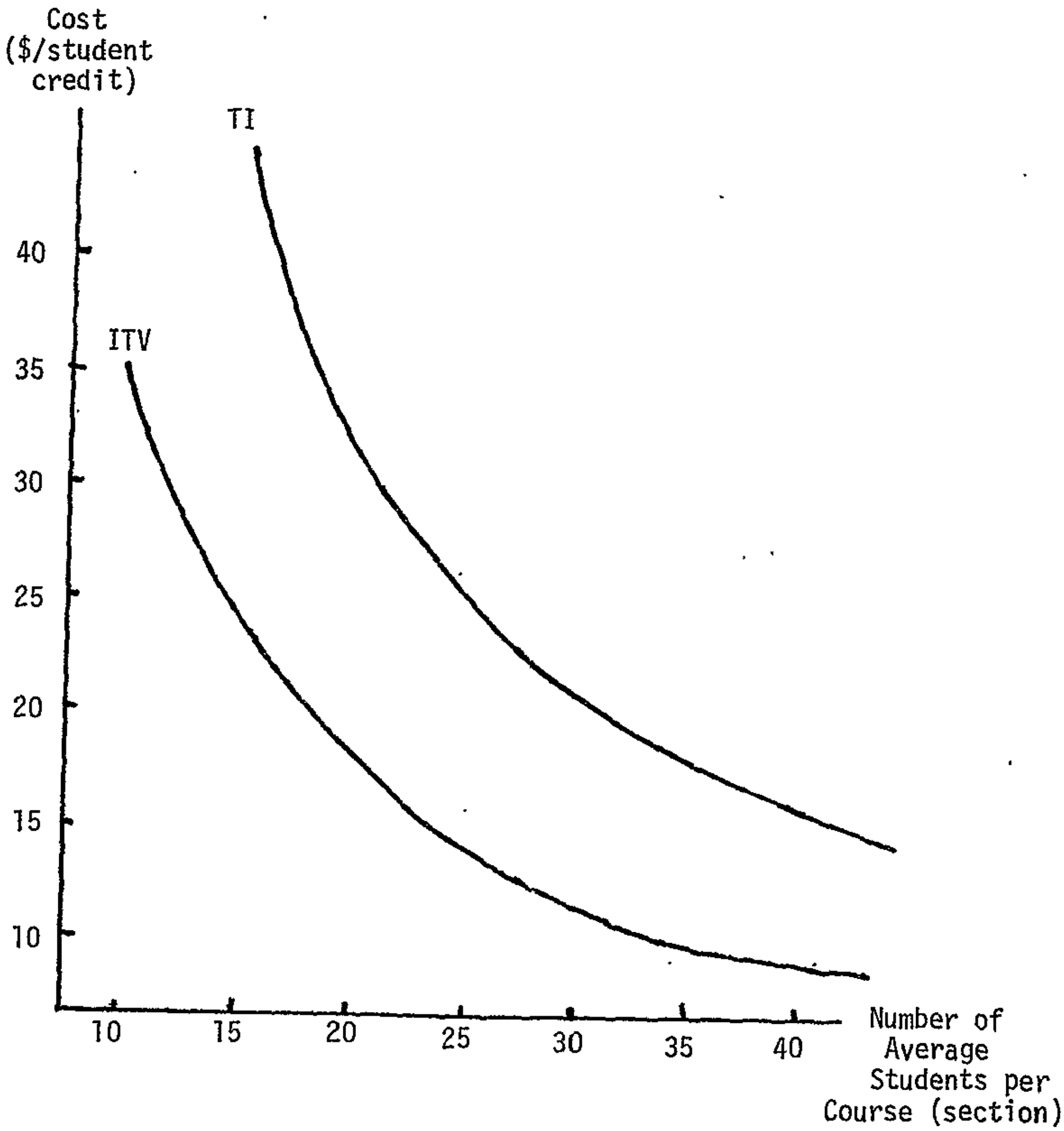
Since this study concerns alternative instruction methods, it is not permissible to treat ITV as an add-on cost and ignore the cost of the instruction itself. Obviously, the figures one elects to use for instructional costs are no more determinate here than they were earlier in the traditional cost analysis. Instructional cost per hour, in fact, does vary considerably, and the choice of a number is somewhat arbitrary. If we assume the same salary and teaching load as earlier, and allow a three to one substitution of the traditional lecture for videotaped lectures, one finished hour of taped programming per week is equivalent to one-fourth time, or $125 per hour ITV. Allowing $10 per hour for teaching aids results in a total of $135 for instruction costs per hour of videotaped lecture. Adding in the system production costs gives an average cost of $168.70 for each hour of videotaped lecture.

In addition to the cost of producing taped lectures, we must consider the costs of transmitting and receiving one hour of taped ITV per week to a classroom on campus. According to Table 7.2, this is $13.95 per hour. Summing the cost of production and delivery gives a total cost of $1,053 per course for ITV, or $351 per course credit.

Figure 7.1 illustrates the cost reduction associated with increased enrollment for ITV and TI. For any given enrollment, the cost of ITV per student credit is less than the cost of TI. Since findings of the previous chapter show ITV to be at least as effective as TI, it is clearly the more cost effective of the two. However, it would be impossible

Figure 7.1

THE RELATIONSHIP BETWEEN CLASS SIZE AND UNIT COST OF TI AND VIDEOTAPED ITV*



* The instructional cost component for both systems is based on a 12 hour course load per semester and $15,000 salary per nine months.

to overemphasize the role played by the many assumptions[9] in the cost analysis in shaping this conclusion.

Computer-Assisted Instruction Costs

Since student use of CAI can be expected to vary considerably, it is useful, at least initially, to compare the costs of CAI and traditional instruction in terms of student contact hours (SCH). The following three factors are the most important considerations to the unit cost of CAI: (1) The communication costs associated with the location and grouping of terminals; (2) The number of terminals and rate of utilization; and, (3) The cost of lesson preparation and the number of users across which these costs may be distributed.

Much of the problem in CAI system design and cost centers around the variable costs of program distribution. System costs strongly depend on the planned location of terminals--distance and density. In general, the greater the distance, the greater the cost, and where density can be increased by grouping, cost savings result. This is the case because engineering constraints require different equipment for different terminal configurations, and the expense can vary considerably.[10]

One of the advantages of a CAI system is its potential for distribution over a broad user base. Most of the cost analysis conducted thus

---

[9]Some of these assumptions are involved in estimating instructional costs: salary and course load, the rate of exchange between production of ITV and TI lectures, and frequency in which taped lectures are revised. Obviously, changing the system capacity or the average class size will have a drastic effect on the difference between costs of ITV and TI.

[10]For an interesting and readable discussion of the impact of engineering constraints on system configuration and costs, see Eastwood and Ballard (1975) and Morley and Eastwood (1975).

far has pertained to a broadly based CAI system.[11] However, since the particular research at hand concerns the cost of substituting CAI for traditional instruction on campus, we are concerned with an essentially local CAI system. The least cost distribution is achieved by cable when the terminal is in the vicinity of the central computer ($.01/SCH). At most, one might choose to link campuses in a metropolitan area, but for distances less than 15 miles, the technology is simple and the cost is less than $.13 per SCH.[12] Hence, the communication costs will be between $.01 and $.13 per SCH for a local CAI system.

In an essentially local system one must be concerned with system capacity and determine if sufficient users are available to reduce the cost per contact hour to feasible levels. At this point, we consider only the cost of central computer facilities. Terminal and lesson preparation costs are dealt with later. A $4.5 million investment for the central computer facility, including necessary hardware and $1.5 million in software plus some course development, totals $6.0 million. Over a 5 year period the annualized cost is $1.2 million.

Although the PLATO IV was designed for a system capacity of 4,000 terminals, the central computer has become fully utilized at 1,000 terminals. One may regard the 1,000 terminals as a lower bound on the capacity of an experimental system. There are two reasons for this. Course authorship requires five times more computer core than student use, and use by authors was greater than expected. Second, terminals have been

---

[11]Albert and Skaperdas (1973), Simonsen and Renshaw (1974), and Eastwood and Ballard (1975).

[12]Morley and Eastwood (1975).

used independently to study different courseware. If terminals were distributed in larger groups, course sharing by terminals may occur, and scheduling of author functions during non-peak times would restore the 4,000 capacity for an operational system.

Table 7.3 shows the costs and capacity of the system under three different assumptions concerning utilization rates. In order to calculate the unit cost of CAI, some assumption must be made regarding the utilization of terminals. Morley and Eastwood (1975) assume each terminal is used for 160 hours per month. Albert and Skaperdas (1973) assume 8 hours per day for 300 days per year. McKeown (1975) assumes 70 hours per week, 40 weeks per year. The system costs per SCH shown in the last column reflect these assumptions. The same rates multiplied by the number of terminals (1,000 and 4,000 respectively) give the total system hours utilized per week (column 2).

If we knew the average student course load and the average number of contact hours per course per week, we could estimate the number of users necessary to reach system capacity (as defined by the assumptions). The average CAI use per week in accounting was 2.1 hours and the average course load was 15 hours. If this were generally true, then the average student has approximately 10 hours (2.1 x 5 courses) of computer contact per week. The number of students at which the system reaches capacity are given in column three of Table 7.3. For example, a system consisting of 4,000 terminals in use for 56 hours per week would have a system cost of $.12/SCH only if the number of users approach 22,400. Similarly, a school with 5,600 students and a 1,000 terminal system would have a cost equal to $.48/SCH. It may be that the contact hours per course

TABLE 7.3

SYSTEM CAPACITY, USER BASE, AND UNIT COST
OF A CAI FACILITY

Assumption with Respect to Terminal Use*	Total System Hours Per Week		Number of Users at System Capacity**		Cost of Central Computer Facility Per SCH***	
	Terminal	Capacity	Terminal	Capacity	Terminal	Capacity
	1,000	4,000	1,000	4,000	1,000	4,000
1. 40 hours/week (Morley and Eastwood, 1975)	40,000	160,000	4,000	16,000	$.67	$.17
2. 56 hours/week (Albert and Skaperdas, 1973)	56,000	224,000	5,600	22,400	.48	.12
3. 70 hours/week (McKeown, 1974)	70,000	280,000	7,000	28,000	.38	.10

* Morley and Eastwood (1975) assumed 160 SCH/terminal/month. Albert and Skaperdas (1973) assumed 8 hours per day, 300 days per year. McKeown (1975) assumed 70 hours/week, 40 weeks/year.

** The number of users with average 15 credit course load necessary to fully utilize the system assuming 2 contact hours per week for a 3 credit course.

*** Cost per SCH of 6 million dollar investment in computer hardware and software amortized over 5 years.

per week are somewhat higher than the 2.1 assumed above. If so, the stipulated capacities will be reached for lower numbers of users. Nevertheless, Table 7.3 provides some insight on the size of the user base assumed in the different cost calculations. It appears that only very large universities would entail the necessary number of users to result in full utilization of 4,000 terminals within 15 miles and costs of $.17 and less. Smaller universities would have to support broader geographical dissemination of CAI than what is implied by a local system, or use a system with fewer terminals in which case unit computer costs could run as high as $.67/SCH.

The cost of lesson preparation is another major determinant of the unit cost of CAI. Estimates of these costs are largely arbitrary and their importance in the cost calculations are governed by the following factors: (1) Is the full cost of lesson preparation treated as a CAI cost, or is it charged off in another form? (2) Are the lessons shared with other systems? (3) What is the useful lifetime of lessons, and what is the effective user base?

The lower bound to lesson cost is not difficult to conceive. Researchers admit that unit costs can be reduced to several cents per SCH if lessons are widely distributed over thousands of users. Or, one may adopt the view that CAI course programming is roughly equivalent to preparing a good textbook, and the unit cost of lesson preparation becomes a royalty or rental cost. McKeown (1974) estimates author royalties and CPU charges at $.15 per SCH. We regard this as a reasonable lower bound.

On the other hand, higher, more realistic estimates of lesson preparation costs are difficult to reach. According to McKeown (1974), the first semester accounting course used as the test sample in this research required approximately 8,000 man-hours of programming. "Although the dollar cost related to these figures is hard to compute (most people involved worked many more hours than they were paid for), it is obvious that this level of cost can only be justified when viewed as a fixed cost to be spread over large numbers of students either at one school or a number of cooperating institutions."[13]

It is difficult to arrive at a specific estimate for lesson preparation costs. Moreover, the attempts to calculate these costs in the literature reach diverse results and usually do not indicate how estimates are obtained. Simonsen and Renshaw (1974) point out that various estimates range from 40 to 200 hours of instructor preparation per hour of presented lesson at an average cost of $1,000 per hour of CAI. The question becomes how reliable are these figures, what is the useful lifetime of the lessons, and what is their distribution over students? Compared with the volume of lesson preparation for the other two technologies, the amount of experience gained preparing CAI lessons is small. We may have to wait until more CAI lessons preparation has been completed before dependable cost information becomes available. Nevertheless, if the Simonsen and Renshaw estimate is the best available, then the unit cost of the lesson component is $1,000 divided by the cumulative number of course sections offered times the average section size. Obviously,

---

[13]McKeown (1974), p. 17.

the unit cost of lesson preparation will vary considerably with the course, but it could easily be in the $2-3 range--larger than all other cost components combined. Furthermore, CAI programming for specialized courses is uneconomical unless it would be possible to share the product with other systems.[14]

The cost of a complete student terminal is $5,300.[15] Having had several years' experience with the system first-hand, McKeown estimated the terminal usage at "minimum 70 hours per week for minimum 40 weeks per year."[16] This gives a cost per SCH of $.22. Assuming 160 student contact hours per month results in $.53 in Morley and Eastwood (1975). A summary of system costs is given in Table 7.4. Both a low and a high estimate are shown. The low is distinguished by the less expensive connection costs, relatively higher terminal usage rate, and by 4,000 terminal capacity as opposed to 1,000. Summing system components suggests a total cost range from $.50 per SCH to $1.33 plus lesson preparation costs per SCH. This is consistent with Simonsen and Renshaw (1974) who calculated the cost at $1.13.[16]

If we take for traditional instruction our previous assumption of a $15,000 salary and a 12 hour teaching load, 30 weeks of traditional instruction cost $2.08 per SCH for an average section of 20 students.

---

[14] Simonsen and Renshaw (1974) and Eastwood and Ballard (1975) feel that this is not presently feasible.

[15] In a recent conversation with Eastwood, it was noted that this price is now about $8,800.

[16] P. 18.

TABLE 7.4**

COSTS OF THE PLATO CAI SYSTEM

		Cost	Depreciation	Unit Cost (S.C.H.)* Low Estimate	Unit Cost (S.C.H.)* High Estimate
A. Production					
1. Central Computer	hardware	$4.5 million	5 years		
	software	$1.5 million	3 years		
		$6.0 million		0.12	0.67
2. Lesson Preparation (or Author Royalties)				0.15	$1,000/hr ÷ [(# course sections) x (ave. section size)]
B. Distribution					
3. Communications Transmission Costs (local only; less than 15 mi) low, wire cable; high, leased telephone line				0.01	0.13
C. Reception					
4. Student Terminal (plasma display panel with driver keyset, random-access slide selector)		$5,300	5 years	0.22	0.53
			TOTAL UNIT COST:	$0.50/SCH	$1.33 + $1,000 ÷ [(# sec) x (sec size)]

* SCH = student contact hour. The "low" estimate involves a system with 4,000 terminals while the "high" symbolizes an experimental 1,000 term facility. The low estimate assumes minimum transmission costs where terminals are connected via coaxial cable in the locality of the central computer. Terminal utilization assumed--70 hr/wk, 40 wks/yr. The "high" estimate assumes leased phone line linkage to central facility within 15 miles, and a usage rate of 160 SCH/month for student terminals.

** Source of data: #1, #3 high, #3 low, #4 high, Eastwood and Ballard (Spring 1975); #2 low, #4 low, McKeown (1974); #2 high, Simonsen and Renshaw (1974).

McKeown suggests a minimum cost for conventional instruction of $1.35 per SCH, while Baldwin claims graduate engineering instruction at CSU goes for $6.50 per SCH.

At first glance the $.50-1.33 cost range for CAI seems to compare favorably with costs of traditional instruction. A system serving 1,000 terminals and 4,000 full time students is probably well within reasonable limits for a local CAI system. Moreover, the $.67 per SCH for such a system does not make CAI prohibitively expensive. Other systems sizes depicted in Table 7.3 would involve lower costs. Therefore, distribution of system costs among users for a local system does not seem to the critical issue in the cost effectiveness determination.

The problem is the cost of lesson preparation. The low estimate of $.50 is achievable only if lessons are rented or shared at extremely low cost. The $1.33 figure includes no lesson costs. With no sharing of courseware, CAI lesson development could undoubtedly only be justified in a few cases such as large lecture courses at universities where 5,000 or more users could be exposed to the program over its 3 or 4 year effective lifetime. In this case, the cost of the lesson per SCH would reduce to $.20 or less, and the total CAI cost for the high estimate becomes $1.53/SCH, a reasonably competitive figure. However, until sharing courseware among CAI systems becomes possible at low cost, it would not be feasible to develop the range and variety of courseware necessary to replace all traditional instruction on campus with CAI.

Just as with ITV, the cost calculations for CAI are largely arbitrary in that they are so heavily influenced by the system characteristics and other assumptions in the analysis. However, after comparing the costs of CAI and TI, the cost of lesson preparation appears to deal a major blow to the feasibility of replacing traditional instruction on campus with CAI. On the other hand, recent technological developments in CAI are producing a new generation of CAI systems served by significantly smaller computers. This favors decentralized CAI systems which along with advancements in lesson programming may permit lessons to be jointly developed and shared. If CAI systems follow this trend, the obstacle of lesson preparation will eventually vanish and CAI will become a cost effective alternative to traditional instruction.

Before we proceed, a word should be said about the ratio of substituting program time for class time. As McKeown (1974) carefully points out, "It must be remembered that PLATO contact hours do not necessarily replace classroom contact hours on a one for one basis."[17] Some of the work programmed on PLATO involves tasks that the student would otherwise perform as part of his homework outside of class. Furthermore, McKeown found evidence that PLATO students could spend less time in study and review outside of class and attain the same achievement level as a control group. Although we have not been able to measure it in this study, the cost savings represented by the reduction in student (and possibly faculty) time is a gain in efficiency that undoubtedly overshadows whatever small gains may have been observed in the achievement tests. It also suggests the cost definitions should include indirect costs as discussed earlier.

---

[17]P. 19.

TABLE A7.1

PRODUCTION COSTS[1]

A7.1a: Studio Classroom and Master Control

Capital Costs

3 TV cameras at $1,000 each		$3,000
1 Sync generator		1,000
1 Pan tilt control unit		1,100
5 TV monitors at $160		800
2 Zoom lenses at $1,100		2,200
Instruction desk with control unit, split screen generator, and back pack playback recorder		4,000
Electronic control, amplifiers, cables and special room wiring		2,300
Master Control panel with TV monitors switching unit		5,600
Studio classroom air conditioning and necessary remodeling		5,000
Related labor		5,000
	Total Cost	$30,000

A7.1b: Recording Facilities Costs[2]

Video Tape Recorders			
1" VTRs (4) $995 each		$ 3,980	
1/2" VTRs (6) $700 each		4,200	
1/4" VTRs (3) $1,400 each		4,200	
	Subtotal		$12,380
Shelves and racks		$ 600	
TV monitors (9) $180 each		1,620	
Custom switcher		7,000	
Cabinets		250	
Cables and carts		100	
Labor		2,300	
	Subtotal		$11,870
	Total		$24,250

TABLE A7.1
(continued)

A7.1c: Base Operating Costs

Administrator, $24,000	1/10 time		$ 2,400
Coordinator, $16,000	3/4 time		12,000
TV Engineer, $15,000	1/5 time		3,000
TV Technicians, $10,800	3 full time		32,200
Secretary, $5,300	1 1/2 full time		8,000
Student labor at $2/hr	6,000 hours		12,000
Supplies and spare parts			11,700
		Total	$81,300

---

1. Source: A.J. Morris, "Final Report on the Cost-Effectiveness of Continuing Engineering Studies by TV." (Mimeo, American Society of Engineering Education, 1974). (Costs pertain to 1972-73.)

2. SURGE had a facility designed to accomplish 9,450 delivered tapes, 30 copies for each of 315 course sections. This analysis assumes 5 tapes/course for 600 courses or 3,000 delivered tapes. The scale of the recording facilities have been reduced by 2/3 in the appropriate categories.

[3] These were operating costs for the off-campus SURGE program in graduate engineering at Colorado State University. Figures are for 200 course capacity, at 30 taped hours per course or 6,000 hours of taped lecture per year.

TABLE A7.2

RECEPTION COSTS

A7.2a: Playback Costs - Closed Circuit[1]

	Cost	Depreciation	Annualized Cost
VTR	$ 1,000	3 yrs at 6%	$ 374
Coaxial cable	10,000	10 yrs at 6%	1,358
Cable maintenance			1,000
TV technicial, $10,800 1/2 time			5,400
Building space, rent			1,000
Replacement parts, etc.			3,000
		Total	$12,132
Cost per hour delivery assuming 30 hours use per week for 30 weeks per year			$13./hr.

A7.2b: Classroom Costs[1]

	Depreciation	Annualized Cost
Television Sets (2) $200 each	3 yrs at 6%	$149
Improved Speaker system $200	10 yrs at 6%	27
Installation $100	10 yrs	10
Maintenance		40
	Total	$226
Cost per hour delivery assuming 30 hours use per week for 30 weeks per year[2]		$.75/hr.

---

1. Source: D.W. Paden, "Teaching Economics Via Television at the College Level," in J.W. Meaney, Televised College Courses, (Washington, D.C.: Academy for Educational Development, 1962). These tables were updated with cost figures from 1972-73 in the Morris Report (May 1974).

2. With three credit courses, this amounts to 10 courses per classroom per week. One of the three hours per week is televised programming, the remainder is discussion section. [226 ÷ 10 x 30]

CHAPTER EIGHT

CONCLUSION

Chapters 1 and 2 apply production theory to analyze effectiveness in education and discuss the empirical problems of estimating the educational production function. Pitfalls await the researcher who ignores the multidimensional nature of educational production. Schools may have preferences for different amounts of outputs that are not part of the effectiveness analysis. Also, different capital endowments or managerial skill differences may give some schools a comparative advantage in producing the one cognitive output under study. Either factor can cause the biased estimates of the single output models that prevail in school effectiveness studies and could easily be responsible for the inconsistent results in the literature.

The crucial issues in estimating an educational production function are summarized below:

(1) What estimation technique is appropriate? Regression analysis and production frontier estimation were considered. If one is confident of the model specification and its ability to measure each variable, the frontier approach is preferred. Otherwise, we expect regression analysis to be more useful.

(2) Functional Form. The properties of different functional forms were discussed with respect to modeling the education process. Since no

single form seems to yield a superior statistical fit, the linear and Cobb-Douglas models were chosen for the empirical work.

(3) Selection and Measurement of Variables. This aspect of estimating educational production functions seems to cause the most problems for researchers. The kind of data needed are frequently not available, and the data that are available usually exhibit multicollinearity. Output measures are often not well enough mated to the programs that are being evaluated. Hence, they are unable to detect subtle differences in programs among schools.

This section resulted in the following major methodological suggestions. School effectiveness studies should adopt a micro or disaggregated approach. This means that experimental data will usually have to be collected for the project, the unit of analysis should be the individual student, and more attention must be given to developing an output measure appropriate to the programs being studied. The project should be defined, not for a general area like mathematics or verbal ability, but for a specific area.

Chapter 3 reviewed the literature and discussed the results in general on variable-by-variable basis. After adjusting for other factors, different schools made no significant impact on student performance. The most important determinant at the elementary and secondary levels was student background characteristics and, in some cases, teacher ability.[1] Evidence at the college level also indicated that schools

[1]As measured directly by verbal tests rather than such proxies as years of experience and salary level.

made no significant difference and student ability replaced student background as the most significant single determinant of achievement.

Chapter 4 outlined the use of ITV and CAI in higher education and reviewed the technology effectiveness literature. The most outstanding conclusion was that no significant difference could be found in student performance as a result of the different technologies. Once again, this could be caused by output measures suited for general subject areas and lacking the individual program validity needed to pick up the differences. Furthermore, the control-experimental group approach held all inputs constant except the delivery mode. Some of these inputs--such as teaching style, textbooks, and course content--are complementary to the technology in the production process. They may require modification and should not be held constant. Our goal was to compare effectiveness of different technology-based programs, not merely to compare the different modes of delivery.

Experimental data were collected and analyzed according to the methodological suggestions advocated above. The interesting findings from the effectiveness analysis are summarized below:

(1) After adjusting for other factors, a statistically significant difference was discovered for the schools and technologies only when the data were pooled. When the sample was partitioned to investigate the interaction among variables, a gain of about 2 percent in effectiveness for ITV and CAI was observed, but it was not statistically significant.

(2) Variables representing student ability were the only other consistently significant variables. However, the impact of SATM and SATV on student achievement differed by technology. Typically, a 100-point

increase in SATM or SATV resulted in a 4 or 3 percent gain with ITV, and in a 2 or 6 percent gain for TI, respectively. In CAI, a 100-point increase in SATM was associated with a 5 to 9 percent gain, while SATV was negative and insignificant. Moreover, investigation revealed that the achievement gain increased for higher SATM scores, while it decreased with higher SATV scores.

In light of these results, certain data limitations must be taken into account:

(1) Only three schools were used to test the effectiveness of three technologies. If one could locate an identical course taught at several schools for each of the three technologies, the effect of the school could be distinguished from that caused by technology.

(2) These results may not be generalized to other subject matter, nor do they apply to ITV and CAI technologies that differ markedly from those at Colorado State and the University of Illinois.

(3) Although output measure used in this study may be better than most, it is still not ideal. The experimental examination was not exactly the same at each school and allowing Colorado State to select the initial questions may have biased the results.

(4) Since much of the data was collected by a student questionnaire, some of the information, such as study time per week and number of credits accumulated in related course areas, may not be reliable. The variable STUDY was only significant in some cases at the 10 percent level, and CREDM, CREDB, and CREDE never seemed important.

In spite of the data limitations, the signs on the coefficients generally conformed to our expectations. Only those inputs that were expected to be important turned out that way. The fact that the data analysis held no unusual and significant surprises is the best indication that the data in this study are better than most. This supports the methodology recommendations described above; that is, technology and school evaluation at the course level.

System costs were examined on the assumption that the technology would supplant traditional teaching on the college campus. ITV costs were calculated assuming total capacity producing 3,000 hours of new videotaped lectures per year. CAI costs assumed that at least a 1,000 terminal capacity was fully used on campus. If we assume 20 students per section as a matter of illustration, traditional costs amount to $2.08 per student contact hour and ITV comes to $1.16. CAI was calculated to be $1.33 plus lesson preparation costs. Given the assumptions of the cost analysis, ITV was the least expensive, followed by TI and CAI.

Since ITV and CAI were found to be about equally effective, it will not matter to the cost-effectiveness conclusion whether we believe that their effectiveness is equal or 2 percent better than TI. Of the three systems, the one with the lowest unit cost will be the most cost effective. ITV proved to be more cost effective than TI, which in turn was more cost effective than CAI.

The qualification of this result is most important. First, although there was an observed difference in the effectiveness among the technologies, the differences in costs ultimately determined which was more cost effective. Second, embedded in the cost analysis are assumptions concerning teaching load, professor salaries, class size and other characteristics that make a significant impact on the cost calculations.

For example, in the ITV system, each of the 600 courses were assumed to be offered only once a year. If each course was offered both semesters or in more than one section, the unit cost of ITV would fall, making it even more cost effective. Numerous other assumptions are capable of influencing the order of the cost-effectiveness ratios.

Although our findings imply that CAI is not cost effective with respect to TI at the present time, low cost lesson sharing, technical innovations, and widely based systems may make it cost effective in the near future. It is clear that CAI has great potential for its effective individualized instruction and for possible time savings by students and teachers. However, depending on assumptions, either CAI or ITV could be cost effective compared with traditional instruction. With more research along the lines pursued here, the relative effectiveness of different technologies can be predicted. However, the costs of technologies must be calculated on a case-by-case basis as a result of the importance of enrollment, number of course credit hours offered, class size, and teacher salaries.

A researcher can calculate the cost effectiveness and make policy suggestions. But adoption of large-scale technology systems in higher education will involve more considerations than the cost effectiveness of the systems. Educators are generally apprehensive about technical innovation and tend to protect the status quo. Widespread adoption of ITV or CAI systems must be preceded by research on their economic impact and an extensive educational and public relations effort.

APPENDIX

STUDENT RESEARCH QUESTIONNAIRE

Please answer each item as fully and accurately as possible.

1. (Please circle) CLASS: Fr Soph Jr Sr SEX: M F

2. AGE: ________

3. Do you have a part time job while you're in school? _______

   What is it? ________________ On the average how many hours per week? __________

4. Racial Ethnic Group: Cauc ___ Amer. Oriental ___ Afro Amer. ___ Spanish-Amer. ___

5. High school attended: _____________ Location: _____________ (city, state)

6. Occupation of father: _____________ Total # yrs. schooling: _______

   Occupation of mother: _____________ Total # yrs. schooling: _______

7. LAST WEEK, how many hours did you spend on this course,

   A. On the computer terminal_____ B. In study outside of class_____ (include text, suppl. reading, homework problems)

8. What would be the AVERAGE NUMBER OF HOURS IN AN AVERAGE WEEK spent on this course:

   A. On the computer terminal _____ B. In study outside of class_____

9. What is your major? ___________ Minor? ____________

10. How many hours are you taking this term? ___________

11. Current college GPA ________ out of possible ______.

    Grade point last term alone ______

12. Check any courses you had in high school: Accounting____

    Bookeeping____ Algebra____ Geometry____ Calculus____ Trig____

    Economics____ Other Business____

STUDENT RESEARCH QUESTIONNAIRE
(continued)

13. How many college credits do you have in these areas up to but not including the current term:

Business______ Economics______ Mathematics______

14. Do you intent to take another accounting course? Yes____ No____

15. Circle the number along the continuum which you feel most closely describes your study habits with specific reference to the accounting course.

I always have to catch up and cram before exams	1	2	3	4	5	I always work regularly and keep up

16. Do the same thing but this time describe your study habits in general.

Always cram before exams	1	2	3	4	5	Always keep up

Thank you for your cooperation!!

STATEMENT OF COMPLIANCE

This is to certify that I am completing the questionnaire on background and ability characteristics of accounting students voluntarily as a part of the research project of Roger L. McClung.

Furthermore, I give Mr. McClung permission to record information pertaining to my academic performance such as test scores and class rankings on the following conditions:

1. As soon as the information stated above has been recorded on my questionnaire, my name and any other unique identification will be deleted from it. In this way, complete anonymity will be assured.

2. This information is to be available to only Mr. McClung and is to be used for research purposes only.

3. No information on an individual person-by-person basis will be published.

4. I realize that I may withdraw from this study at any time.

Signed ______________________________

Date ______________________________

OUTPUT INSTRUMENT

The exam which follows was administered by Colorado State University, the ITV school. The traditional school, the University of Missouri at St. Louis, modified it for their use by deleting eight multiple choice questions and the work study problem. These are indicated by a "TI" in the margin. University of Illinois, the CAI school, omitted two multiple choice questions indicated by a "CAI" in the margin.

BA 200
Final
Spring, 1974

Name ____________________

Section ____________________

Part I. (76 points)

Write the letter on the line that best answers the question:

_____ 1. The liability created by a business enterprise when it makes a purchase on account is termed:
a. Account receivable
b. Account payable
c. Depreciation
d. Dividend
e. None of the above

_____ 2. Land with an assessed value of $30,000 for property tax purposes is offered for sale at $70,000. The land is acquired by a business enterprise for $20,000 cash and a non-interest-bearing note payable of $45,000 due in 30 days. The amount used in the buyer's accounting records to record the acquisition of the land is:
a. $30,000 b. $70,000 c. $20,000 d. $65,000 e. None of the above

_____ 3. If total assets decreased by $5,000 during a period of time and capital increased by $15,000 during the same period, the amount and direction (increase or decrease) of the period's change in total liabilities is:
a. $10,000 increase b. $10,000 decrease c. $20,000 increase
d. $20,000 decrease e. None of the above

_____ 4. A business enterprise paid creditors on account, $1,000. The effect of the transaction on the accounting equation was:
a. Increase in one asset, decrease in another asset
b. Increase in an asset, increase in a liability
c. Increase in an asset, increase in capital
d. Decrease in an asset, decrease in capital
e. None of the above

_____ 5. The total assets and the total liabilities of a particular business enterprise at the beginning and at the end of the year are stated as follows: During the year the owner had withdrawn $18,000 for personal use and had made an additional investment in the enterprise of $5,000.

	ASSETS	LIABILITIES
Beginning of the year	$166,000	$72,000
End of the year	177,000	99,000

The amount of net income or net loss for the year was:

a. Net income of $11,000 b. Net income of $13,000 c. Net loss of $3,000 d. Net loss of $27,000 e. None of the above

C-3

BA 200
Final
Spring, 1974

Name ____________________

Section ____________________

_____ 6. The financial statement that presents the assets, liabilities, and capital of a business as of a specific date is termed:
a. Balance sheet b. Income statement c. Capital statement
d. Funds statement e. None of the above

_____ 7. A debit may signify:
a. Increase in asset accounts b. Decrease in liability accounts
c. Decrease in capital accounts d. All of the above e. None of the above

_____ 8. Which of the following applications of the rules of debit and credit is <u>false</u>?

	Recorded in Account as	Normal Bal. of Account
a. Increase in salary expense account	Debit	Debit
b. Increase in supplies account	Debit	Debit
c. Decrease in accounts payable account	Debit	Credit
d. Decrease in accounts receivable account	Credit	Debit
e. None of the above		

_____ 9. The verification that the debits and credits in the ledger are equal is called:
a. Balance sheet b. Retained earnings statement c. Trial Balance
d. Account e. None of the above

_____10. Which of the following errors, each considered individually, would cause the trial balance totals to be unequal?
a. A payment of $500 to a creditor was posted as a debit of $500 to Accounts Payable and a credit of $50 to Cash.
b. Cash received from customers on account was posted as a debit of $250 to Cash and a debit of $250 to Accounts Receivable.
c. A payment of $275 for equipment was posted as a debit of $275 to Equipment and a credit of $257 to Cash.
d. All of the above.
e. None of the above.

_____11. A series of entries journalized at the end of the accounting period to remove the balances from the temporary accounts so that they will be ready for use in accumulating data for the following accounting period are termed:
a. Adjusting entries b. Closing entries c. Correcting entries
d. All of the above e. None of the above

BA 200
Final
Spring, 1974

Name ____________________

Section ____________________

_____ 12. If the effect of the debit portion of an adjusting entry is to increase the balance of the expense account, which of the following describes the effect of the credit portion of the entry?
a. Decreases the balance of an asset account
b. Increases the balance of an asset account
c. Decreases the balance of a liability account
d. Increases the balance of a revenue account
e. None of the above

_____ 13. Which of the following accounts should be closed to Income Summary at the end of the fiscal year?
a. Depreciation expense b. Sales c. Supplies expense
d. All of the above e. None of the above

_____ 14. The balance in the prepaid insurance account before adjustment at the end of the year is $1,740, and the amount of insurance expired during the year is $620. The adjusting entry required is:
a. Debit insurance expense, $620; credit Prepaid insurance, $620
b. Debit PREPAID INSURANCE, $620; credit Insurance Expense, $620
c. Debit Insurance Expense, $1,120; credit PREPAID Insurance, $1,120
d. Debit Prepaid Insurance, $1,120; credit Insurance Expense, $1,120
e. None of the above.

_____ 15. A business enterprise pays weekly salaries of $5,500 on Friday for a five-day week ending on that day. The adjusting entry necessary at the end of the fiscal period ending on Tuesday is:
a. Debit Salaries Payable, $2,200; credit Salary Expense, $2,200
b. Debit Salary Expense, $2,200; credit Salaries Payable, $2,200
c. Debit Salary Expense, $2,200; credit Drawings, $2,200
d. Debit Drawings, $2,200; credit Salaries Payable, $2,200
e. None of the above

_____ 16. Cash of $520 received from a customer on account was recorded as as $250 debit to Accounts Receivable and credit to Cash. The necessary correcting entry is:
a. Debit Cash, $270; credit Accounts Receivable, $270
b. Debit Accounts Receivable, $270; credit Cash, $270
c. Debit Cash, $520; credit Accounts Receivable, $520
d. Debit Cash, $770; credit Accounts Receivable, $770
e. None of the above

_____ 17. The adjusting entry to record depreciation of equipment for the fiscal period is:
a. Debit Depreciation Expense, Credit Depreciation Payable
b. Debit Depreciation Payable, credit Depreciation Expense
c. Debit Depreciation Expense, credit Equipment
d. Debit Equipment, credit Depreciation Expense
e. None of the above

BA 200
Final
Spring 1974

Name ____________________

Section ____________________

_____ 18. At the end of the preceding fiscal year the usual adjusting entry for accrued salaries owed to employees was inadvertently omitted. The error was not corrected, but the accrued salaries were included in the first salary payment in the current fiscal year. Which of the following statements is true?
a. Salary expense was understated and net income was overstated for the preceding year.
b. Salary expense was overstated and net income was understated for current year.
c. Salaries payable was understated at end of preceding fiscal year
d. All of the above
e. None of the above

_____ 19. The receipt of cash from sales should be recorded in the:
TI
a. Sales journal b. Cash receipts journal c. Purchases journal d. General journal e. None of the above

_____ 20. Which of the following transactions would be recorded in the sales journal?
TI
a. Sale of merchandise on account b. Sale of merchandise for cash c. Sale of vacant land (plant asset) for cash d. Sale of vacant land (plant asset) in return for note receivable
e. None of the above

_____ 21. The journal entry to record the issuance of a credit memorandum to a customer, John Coe, for merchandise returned is:
a. Debit Sales Returns and Allowances, credit Accounts Receivable-John Coe b. Debit Accounts Receivable-John Coe, credit Sales Returns and Allowances c. Debit Accounts Receivable-John Coe, credit Sales d. Debit Accounts Payable-John Coe, credit Sales Returns and Allowances. e. None of the above

_____ 22. For each transaction recorded in the purchases journal, the credit is entered in the:
TI
a. Accounts Payable Cr. column b. Purchases Dr. column
c. Accounts Payable Dr. column d. Purchases Cr. column
e. None of the above

_____ 23. Which of the following would be recorded in a multi-column purchases journal?
TI
a. Merchandise purchased on account for resale to customers
b. Supplies purchased on account for use in the business
c. Equipment purchased on account for use in the business
d. All of the above
e. None of the above

BA 200
Final
Spring 1974

Name ____________________

Section ____________________

_____24. The controlling account in the general ledger that summarizes the debits and the credits to the individual accounts in the creditors ledger is entitled:
TI
a. Accounts receivable b. Accounts payable c. Purchases
d. Sales e. None of the above

_____25. If merchandise purchased on account for resale to customers is returned to the seller, the buyer may inform the creditor of the details by issuing a:
TI
CAI
a. Sales invoice b. Purchase invoice c. Debit memorandum
d. Credit memorandum e. None of the above

_____26. A purchase invoice included the following information: merchandise price, \$500; transportation, \$25; terms, FOB shipping point, 2/10, n/30. If the invoice is paid within the discount period, what is the correct amount of the payment?
a. \$500 b. \$525 c. \$490 d. \$515 e. None of the above

_____27. Transportation costs incurred on the purchase of office equipment for use in the business had been charged to Purchases. Assuming that the above error has been posted and that it is discovered in the same fiscal period in which it occurred, the necessary correcting entry is:
a. Debit office equipment; credit Purchases
b. Debit Purchases; credit Office Equipment
c. Debit Transportation In; credit Purchases
d. Debit Purchases; credit Transportation In
e. None of the above

_____28. On the basis of the following data, what is the proper adjusting entry at December 31, the end of the fiscal year: supplies account balance before adjustment, \$550; supplies physical inventory on December 31, \$170?
a. Debit Supplies, \$170; credit Supplies Expense, \$170
b. Debit Supplies Expense, \$170, credit Supplies, \$170
c. Debit Supplies Expense, \$380; credit Supplies, \$380
d. Debit Supplies, \$380; credit Supplies Expense, \$380
e. None of the above

_____29. On the basis of the following data, what is the proper adjusting entry at December 31, the end of the fiscal year: prepaid insurance account balance before adjustment, \$1,750; unexpired premiums per analysis of policies, \$950?
a. Debit Insurance Expense, \$950; credit Prepaid Insurance, \$950
b. Debit Prepaid Insurance, \$950; credit Insurance Expense, \$950
c. Debit Prepaid Insurance, \$800; credit Insurance Expense, \$800
d. Debit Insurance Expense, \$800; credit Prepaid Insurance, \$800
e. None of the above

BA 200
Final
Spring 1974

Name ____________________

Section ____________________

_____30. The closing entry to transfer the net loss for a single proprietorship to the appropriate account at the end of the fiscal year is:
a. Debit Income Summary, credit Retained Earnings
b. Debit Retained Earnings, credit Income Summary
c. Debit Income Summary, credit Capital
d. Debit Capital, credit Income Summary
e. None of the above

_____31. If the effect of the credit portion of the particular adjusting entry is to increase a liability, the effect of the debit portion of the entry would be to:
a. Decrease an expense b. Increase an asset c. Increase an expense d. Decrease an asset e. None of the above

_____32. The prepaid insurance account has a balance of $600 at the beginning of the year and was debited during the year for $750 representing the total of premiums on policies purchased during the year. If it is ascertained that $480 of insurance premiums have expired during the year, the amount of prepaid insurance to be reported on the balance sheet at the end of the year would be:
a. $270 b. $480 c. $1,230 d. $870 e. None of the above

_____33. Unearned rent would appear on the balance sheet as a:
a. Current asset b. Plant asset c. Current liability
d. Long-term liability e. None of the above

_____34. The general term employed to indicate an expense or a revenue that gradually increases with the passage of time but has not yet been recognized in the accounts by a routine entry is:
a. Deferral b. Accrual c. Depreciation d. Receivable
e. None of the above

_____35. The real estate tax is estimated at $9,000 for the fiscal year beginning January 1, which coincides with the fiscal year of the taxing authority. Appropriate accruals were made in January and February and the tax statement for $9,300 was received in March. The entry to record the accrual for March would debit Property Tax Expense for:
TI
CAI
a. $9,000 b. $775 c. $750 d. $825 e. None of the above

_____36. Prepaid expenses that are not initially recorded as expenses would be recorded as:
a. Revenues b. Assets c. Liabilities d. Capital
e. None of the above

BA 200
Final
Spring 1974

Name ______________________

Section ______________________

_____37. In the first year of operations, Cam Publishers received $190,000 from advertising contracts and $280,000 from magazine subscriptions, crediting the two amounts to revenue accounts. At the end of the year, the deferral of advertising revenues amounted to $35,000 and the deferral of magazine subscriptions amounted to $115,000. The total amount of revenue to appear on the income statement for the year would be:
a. $470,000 b. $245,000 c. $320,000 d. $150,000
e. None of the above

_____38. Which of the following accounts in the ledger of a partnership will ordinarily appear in the post-closing trial balance?
a. Depreciation expense b. Purchases c. Sales d. All of the above e. None of the above

BA 200
Final
Spring 1974

Name ______________________

Section ______________________

PART II. (30 points)
TI

Prepare an eight column work sheet for Ross Machine Shop for October. The balances in the ledger as of October 31, 19__, before adjustments, are as follows:

Cash	$ 6,250
Supplies	3,300
Prepaid Insurance	750
Equipment	11,000
Accumulated Depreciation	6,500
John Ross, Capital	10,200
John Ross, Drawing	450
Sales	10,500
Salary Expense	4,250
Miscellaneous Expense	1,200

Adjustment data for October 31, 19__: Supplies on hand, $1,200; insurance expired, $40; depreciation on equipment, $120; salaries accrued, $300

Ross Machine Shop
Work Sheet
For Month Ended October 31, 19__

Account Title	Trial Balance		Adjustments		Income Statement		Balance Sheet	
	Dr.	Cr.	Dr.	Cr.	Dr.	Cr.	Dr.	Cr.
Cash								
Supplies								
Prepaid Insurance								
Equipment								
Accum. Depreciation								
John Ross, Capital								
John Ross, Drawing								
Sales								
Salary Expense								
Misc. Expense								
Supplies Expense								
Insurance Expense								
Deprec. Expense								
Salaries Payable								
NET INCOME								

MEANS AND STANDARD DEVIATIONS OF VARIABLES
FOR FIVE SAMPLES AT THE THREE INSTITUTIONS
PLUS TWO COMPOSITE SAMPLES*

DSN # Cases	LECT1 26	LECT2** 27	TV 190	CAI1 66	CAI2** 29	POOL 246	ALL 338
OUTPUT	71.5 (13.1)	70.8 (12.3)	79.2 (11.8)	76.4 (11.4)	72.7 (15.0)	77.5 (12.6)	76.8 (12.5)
EDUCF	11.7 (3.5)	13.1 (2.6)	14.4 (3.1)	14.9 (2.8)	13.3 (3.4)	14.2 (3.1)	14.1 (3.2)
EDUCM	11.9 (2.1)	11.9 (1.9)	13.3 (2.5)	13.6 (2.2)	13.4 (2.7)	13.1 (2.5)	13.1 (2.4)
JOB	15.2 (11.2)	19.5 (14.0)	3.4 (8.2)	3.6 (6.7)	13.5 (15.2)	6.3 (11.4)	6.5 (11.0)
SATM	529 (120)	531 (95)	560 (85)	630 (77)	624 (66)	565 (87)	575 (92)
SATV	471 (104)	459 (89)	484 (84)	505 (88)	555 (92)	490 (89)	491 (90)
RANK	71.1 (20.1)	65.7 (21.7)	74.0 (20.0)	84.4 (13.2)	84.7 (11.5)	74.3 (19.8)	76.0 (19.1)
LOAD	14.6 (1.6)	11.0 (3.9)	15.5 (2.0)	15.4 (2.1)	12.5 (3.5)	14.7 (2.9)	14.8 (2.7)
STUDY	3.2 (1.48)	5.7 (3.8)	4.0 (2.2)	2.4 (1.7)	3.4 (2.9)	4.2 (2.6)	3.7 (2.5)
CREDB	5.1 (3.6)	3.5 (3.4)	4.3 (4.6)	1.6 (3.5)	1.8 (4.2)	4.0 (4.5)	3.6 (4.4)
CREDE	4.0 (3.1)	4.3 (3.5)	4.8 (4.5)	4.5 (2.3)	3.7 (4.8)	4.6 (4.5)	4.6 (4.0)
CREDM	8.5 (4.5)	8.0 (6.5)	8.2 (5.2)	7.2 (3.5)	8.3 (5.6)	8.2 (5.4)	8.0 (5.0)
CAITIM				32.0 (11.4)	25.5 (12.8)		30.0*** (12.2)
CAILSN				92.7 (12.0)	67.0 (25.9)		84.9*** (20.9)
CAIDAZ				19.7 (5.8)	14.6 (5.4)		18.1*** (6.1)

* Standard deviations appear in parentheses.

** Data collected during summer term. LOAD and STUDY are adjusted for comparability to fall term.

*** For the subset of 95 CAI students.

BIBLIOGRAPHY

1. Acland, H. (1972). "Social Determinants of Education Achievement: An Evaluation and Criticism of Research." PhD Dissertation, Oxford University.

2. Adams, E. (1969). "Field Evaluation of the German CAI Lab." In Atkinson and Wilson, ed., Computer Assisted Instruction: A Book of Readings. New York: Academic Press.

3. Adams, E. and H. Morrison. (1969). "Pilot Study of a CAI Laboratory in German." In Atkinson and Wilson, ed., Computer Instruction: A Book of Readings. New York: Academic Press.

4. Aigner, D.J. and S.F. Chu. (1968). "On Estimating the Industry Production Function." American Economic Review 58 (September 1968): 826-839.

5. Alpert, D.A., and D.L. Bitzer. (1970). "Advances in Computer Based Education." Science, 167 (March 20, 1970): 1582-90.

6. Alpert, D.A., and D. Skaperdas. (1973). "The Design of an Economically Viable Large-Scale Computer Based Education System." University of Illinois, CERL Report No. X-5, Education Research Laboratory.

7. Anderson, B. and E. Greenberg. (1972). "Educational Production Functions For Teacher-Technology Mixes: Problems and Possibilities." Memorandum 72/2, Center for Development Technology, Washington University (March 1972).

8. Astin, A.W. (1973). "Measurement and Determinants of the Outputs of Higher Education." In L. Solomon and P. Taubman, eds. (1973).

9. ________. (1970). "The Methodology of Research on College Impact." Sociology of Education, 43 (pt. 1): 223-254; 43 (pt. 2): 437-50.

10. ________. (1976). Predicting Academic Performance in College. New York: Free Press.

11. ________. (1962). "'Productivity' of Undergraduate Institutions." Science, 136 (April 13, 1962): 129-35.

12. ________. (1968). "Undergraduate Achievement and Institutional Excellence." Science, 161 (August 16, 1968): 661-668.

13. Atkinson, R.C. (1968). "Computerized Instruction and The Learning Process." American Psychologist, 23: 225-239.

14. Attiyeh, R. and K. Lumsden. (1972). The Production of Economic Understanding: Analysis of First Year University Economics Course, Economics Education Project, Heriot-Watt University.

15. Averch, H., et al. (1972). "How Effective is Schooling? A Critical Review and Synthesis of Research Findings," R-956-PCSF/RC, Santa Monica, California: Rand Corporation.

16. Averch, H., and H. Kiesling. (1970). "The Relationship of School and Environment to School Performance: Some Simultaneous Models for PROJ TALENT High Schools." Santa Monica, California: Rand Corporation (unpublished).

17. Axeen, M.F. (1967). "Teaching the Use of the Library to Undergraduates: An Experimental Comparison of Computer Based Instruction and the Conventional Lecture Method." Report No. R-361, University of Illinois, Coordinated Science Lab.

18. Baldwin, L.V. (1973). "Video Tape Applications in Engineering Education." In Harrison and Stolwrow (January 1974): 172-92.

19. Barnett, W. (1974). "Production Function For Education: An Empirical Study," PhD dissertation, Michigan State University.

20. Bassett, L.R. and R.L. Miller. (1970). "The Production of Education and Socio-economic Inputs." Institute for Economic Research Discussion Paper #127, University of Washington, Seattle, Washington.

21. Benson, C.S. et al. (1965). "State and Local Fiscal Relationship in Public Education in California." Report of Senate Committee on Revenue and Taxation.

22. Bieker, F. and K.R. Anschel. (1973) "Estimating Educational Production Functions for Rural High Schools: Some Findings." American Journal of Agricultural Economics 6 (August 1973): 515-519.

23. Bitzer, M. and M. Boudreax. (1969). "Using a Computer to Teach Nursing." Nursing Forum, 8.

24. Bowles, S. (1969). "Educational Production Functions." (Final Report) Washington, D.C.: HEW/OE.

25. ________ (1968a). "Towards an Educational Production Function." Working Paper.

26. ________ (1970). "Towards an Educational Production Function." In W.L. Hansen, ed. Education, Income, and Human Capital. New York: National Bureau of Economic Research, No. 35.

27. ________ (1972). "Unequal Education and the Reproduction of the Social Division of Labor." In M. Carnoy, ed. Schooling in a Corporate Society, New York: D. McKay (1972): 36-66.

28. Bowles, S., and H.M. Levin. (1968b). "The Determinants of Scholastic Achievement--An Appraisal of Some Recent Evidence." Journal of Human Resources 3 (Winter 1968): 3-24.

29. ________. (1968c). "More on Multicollinearity and the Effectiveness of Schools." Journal of Human Resources (Summer 1968): 393-400.

30. Brown, B. and D. Saks. (1975). "Production and Distribution of Cognitive Skills Within Schools." Journal of Political Economy 83 (June 1975): 571-594.

31. Buckles, S. and M. McMahon. (1971). "Further Evidence on the Value of Lectures in Elementary Economics." Journal of Economic Education 2: 139-141.

32. Burkhead, J. et al. (1967). Input and Output in Large City High Schools, Syracuse, New York: Syracuse University Press.

33. Cain, G. and H. Watts. (1970). "Problems in Making Policy Inferences from the Coleman Report." American Sociological Review 35 (April 1970): 228-242.

34. Carlson, S. (1939). A Study on the Pure Theory of Production, London: King and Son.

35. Carnegie Commission on Higher Education. (1972). The Fourth Revolution, New York: McGraw Hill.

36. Carnoy, M. (1971). "A Systems Approach to Evaluating Education." Stanford University, mimeo.

37. Carnoy, M., ed. (1972). Schooling in a Corporate Society, New York: D. McKay.

38. Carroll, J. (1963). "A Model of School Learning." Teachers College Record 64: 723-33.

39. Castleberry, S. and Lagowski. (1970). "Individualized Instruction Using Computer Techniques." Journal of Chemical Education 47: 91-96.

40. Chiswick, B. (1973). "Schooling, Screening and Income." In Solomon and Taubman, eds., Does College Matter? New York: Academic Press, 151-158.

41. Chu, G.C. and W. Schramm. (1967). Learning from Television: What the Research Says, Washington, D.C.: National Association of Educational Broadcasters.

42. Cohn, E. (1968). "Economies of Scale in Iowa High School Operations." Journal of Human Resources 3 (Fall 1968): 422-434.

43. ________. (1972). "Student Characteristics and Performance in Economic Statistics." Journal of Economic Education 3: 106-111.

44. Coleman, J.S., et al. (1966). Equality of Educational Opportunity, Washington, D.C.: Health, Education and Welfare, Office of Education (OE-38001).

45. ________. (1968). "Equality of Educational Opportunity: Reply to Bowles and Levin." Journal of Human Resources 3 (Spring 1968): 237-246.

46. ________. (1972). "The Evaluation of Equality of Educational Opportunity." In Mosteller and Moynihan, eds., On Equality of Educational Opportunity, New York: Random House.

47. ________. (1970). "Reply to Cain and Watts." American Sociological Review 35 (Spril 1970): 242-249.

48. Coleman, J.S. and N. Karweit. (1970). "Measure of School Performance." R-488-RC, Santa Monica, California: Rand Corporation.

49. Conlisk, J. (1971). "A Bit of Evidence on the Income Education Ability Interaction." Journal of Human Resources (Summer 1971): 358-362.

50. Comstock, G. (1972). "National Utilization of Computers." In R. Levin, et al., The Emerging Technology, New York: McGraw Hill.

51. Coombs, F.S. and J. Peters. (1971). "Student Explorations of Political Processes: The Use of Computer Based Education." Proceedings of the Tenth National Gaming Symposium.

52. Cronbach, L.J. (1969). "Validation of Education Measures." Proceedings of 1969 Invitational Conference on Testing Problems, Princeton, New Jersey: ETS.

53. Cronbach, L.J. and L. Furby. (1970). "How Should We Measure 'Change',--or Should We?" Psychology Bulletin 74: 68-60.

54. Dubin, R. and R.A. Hedley. (1969). The Medium May Be Related to the Message: College Instruction By TV, Eugene, Oregon: Oregon University Press.

55. Eastwood, L.F. and R.J. Ballard. "The PLATO IV CAI System: Where Is It Now? Where Can It Go?" Journal of Educational Technology Systems (Spring 1975).

56. Emery, E. and T. Enger. (1972). "Computer Gaming and Learning in an Introductory Economics Course." Journal of Economic Education 3: 77-85.

57. Farrell, M. (1957). "The Measurement of Productive Efficiency." Journal of the Royal Statistics Society 120: 253-81.

58. Feldman, K.A. and T.M. Newcomb. (1969). The Impact of College on Students, San Francisco, California: Jossey Bass.

59. Flanagan, J.C. and W. Cooley. (1966). Project Talent: One Year Follow Up Studies, Pittsburgh: University of Pittsburgh.

60. Fletcher, J. and R.C. Atkinson. (1972). "An Evaluation of the Stanford CAI Program in Initial Reading (Grade K Through 3)." Journal of Educational Psychology 63: 597-602.

61. Fox, T.G. (1969). "School System Resource Use in Production of Interdependent Educational Outputs." Joint National Meeting of American Astronautical Society and Operations Research Society.

62. Garner, W.T. (1973). "Identification of an Education Production Function by Experimental Means." PhD dissertation, University of Chicago.

63. Gintis, H. (1971). "Education and the Characteristics of Worker Productivity." American Economics Review 61 (May 1971): 266-279.

64. ________. (1972). "Toward a Political Economy of Education." Harvard Educational Review 42 (February 1972): 70-97.

65. Goodman, S.M. (1959). The Assessment of School Quality, Albany: New York State Department of Education.

66. Gourman, J. (1967). The Gourman Report, Phoenix: Continuing Education Institute.

67. Grandey, R.C. (1974). "The Use of Computers to Aid Instruction in Beginning Chemistry." Journal of Chemistry Education.

68. Guthrie, J.W. (1970). "A Survey of School Effectiveness Studies." In Mood, A.M., ed., Do Teachers Make A Difference? Health, Education, and Welfare, Office of Education.

69. Guthrie, J.W. et al. (1971). Schools and Inequality, Cambridge, Mass: MIT Press.

70. Hansen, D., W. Dick and H. Lippert. (1968). "Research and Implementation of Collegiate Instruction of Physics via CAI." Technical Report No. 3, Florida State University.

71. Hanushek, E. (1972). Education and Race: An Analysis of Education Production Process, Lexington, Mass: D.C. Health Lexington Books.

72. ________. (1968). "The Education of Negroes and Whites." PhD dissertation, Department of Economics, MIT.

73. ________. (1975). "Learning by Observing the Performance of Schools." In R.A. Wallhaus, ed., Measuring and Increasing Academic Productivity, San Francisco: Jossey-Bass.

74. ________. (1970a). "The Production of Education, Teacher Quality, and Efficiency." In Mood, A.M., ed., Do Teachers Make a Difference? Health, Education and Welfare, Office of Education.

75. ________. (1971). "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro-Data." American Economics Review 61 (May 1971): 280-288.

76. ________. (1970b). "The Value of Teachers in Teaching." RM-6362-CC/RC, Santa Monica, California: The Rand Corporation.

77. Hanushek, E. and J.F. Kain. (1972). "On the Value of Equality of Educational Opportunity as a Guide to Public Policy." In Mosteller and Moynihan, eds., On Equality of Educational Opportunity, New York: Random House.

78. Harris, C.W. (1963). Problems in Measuring Change, Madison, Wisc: University of Wisconsin Press.

79. Harrison, S. and L. Stolwrow, eds. (1974). "Productivity in Higher Education." Educational Technology Symposium, Washington, D.C.: Health, Education and Welfare, OE/NIE (January 1974).

80. Hartnett, R.T. (1972). Accountability in Higher Education, Princeton, New Jersey: E.T.S.

81. Havrilesky, T. (1971). "A Test of the Effectiveness of Teaching Money and Banking by Programmed Instruction." Journal of Economic Education 2: 151-154.

82. Hawkridge, D., et al. (1968). "A Study of Selected Exemplary Programs for the Education of the Disadvantaged Children," I & II. American Institute for Research, Palo Alto, California.

83. Heady, E.O. and J.L. Dillon. (1960). Agricultural Production Functions, Ames, Iowa: Iowa State University Press.

84. Henderson, J.M. and R. Quandt. (1958). Microeconomic Theory, New York: McGraw Hill.

85. Homeyer, F. (1970). "Development and Evaluation of an Automated Assembly Language Teacher." Technical Report No. 3, University of Texas at Austin, CAI Lab.

86. Jamison, D. et al. (1974). "The Effectiveness of Alternative Instructional Media: A Survey." Review of Educational Research 44: 1-68.

87. Jencks, C. (1972). "The Effects of High Schools on Their Students." Harvard Center for Educational Policy Research.

88. ________. (1972). "The Quality of the Data Collected by the Equality of Educational Opportunity Survey." In Mosteller and Moynihan, eds., On Equality of Educational Opportunity, New York: Random House.

89. Jencks, C. et al. (1972). Inequality, New York: Basic Books.

90. Jenkins, E. (1976). "The Potential of PLATO." CHANGE Special Report of Teaching: 1.

91. Johnson, M. and Scriven. (1967). "Class Size and Achievement Gains in Seventh and Eighth Grade English and Mathematics." School Review 75: 300-310.

92. Judd, W.A., C. Bunderson, and E.W. Bessent. (1970). "An Investigation of the Effects of Learner Control in CAI Prerequisite Mathematics." MATHS Technical Report No. 5, University of Texas, Austin, CAI Lab.

93. Katzman, M.T. (1967). "Distribution and Production in a Big City Elementary School System." PhD dissertation, Department of Economics, Yale University. Also in Yale Economic Essays (Spring 1968) 8 (1).

94. ________. (1971). The Political Economy of Urban Schools, Cambridge: Harvard University Press.

95. Keeler, E. and J. McCall. (1973). "A Note on the Effectiveness of Teacher Experience." (P-5010), Santa Monica, California: Rand Corporation.

96. Kiesling, H.J. (1968). "High School Size and Cost Factors." Washington, D.C.: HEW, Office of Education.

97. ________. (1971). "Input and Output in California Compensatory Education Projects." R-781-CC/RL, Santa Monica, California: Rand Corporation.

98. ________. (1967). "Measuring a Local Government Service: A Study of School Districts in New York State." Review of Economics and Statistics 49 (August 1967): 356-367.

99. ________. (1969). "The Relationship of School Inputs to Public School Performance in New York State." P-4211, Santa Monica, California: Rand Corporation.

100. ________. (1970). "A Study of Cost and Quality of New York School Districts." Washington, D.C.: HEW, Office of Education.

101. Klees, S. (1975). "Instructional Technology and Its Relationship to Quality and Equality in Education in a Developing Nation." PhD dissertation, Department of Economics, Stanford University.

102. Klitgaard, R. and G. Hall. (1973). "Are There Unusually Effective Schools?" P-4995, Santa Monica, California: Rand Corporation.

103. ________. (1973). "A Statistical Search for Unusually Effective Schools." R-1210-CC/RC, Santa Monica, California: Rand Corporation.

104. ________. (1974). "Achievement Scores and Educational Objectives." R-1217-NIE, Santa Monica, California: Rand Corporation.

105. Lawler, R.M. (1971). "An Investigation of Selected Instructional Strategies in an Undergraduate Computer Managed Instruction Course." Technical Report No. 19, Florida State University.

106. Levien, R.E., et al. (1972). The Emerging Technology, New York: McGraw Hill.

107. Levin, H. (1971). "Concepts of Economic Efficiency and Educational Production." Paper delivered at the National Bureau of Economic Research Conference on Education as an Industry (June 4-5, 1971).

108. ________. (1970b). "A Cost Effectiveness Analysis of Teacher Selection." Journal of Human Resources 5 (Winter 1970): 24-33.

109. ________. (1971). "Frontier Functions: An Econometric Approach to Evaluation of Education Effectiveness." Memo 80, Stanford University, Center for Research and Development in Teaching.

110. ________. (1974). "Measuring Efficiency in Educational Production." Public Finance Quarterly 2 (January 1974): 3-25.

111. ________. (1970a). "New Model of School Effectiveness." In Mood, A.M., ed., Do Teachers Make A Difference? Health, Education and Welfare, Office of Education (OG-58042).

112. ________. (1968). "Recruiting Teachers for Large City Schools." Washington, D.C.: Brookings Institute.

113. McKeown, J.C. (1974). "Computer Assisted Instruction for Elementary Accounting." Urbana Champaign: University of Illinois, College of Commerce and Business.

114. Marklund, S. (1963). "Scholastic Attainment as Related to Size and Homogeneity of Class." Review of Education Research 6 (November 1963): 63-67.

115. Melmed, A.S., Chairman. (1972). "Productivity and Efficiency in Education." Education Panel to the Committee on Automation Opportunities in the Service Areas, Federal Countil for Science and Technology.

116. Michelson, S. (1970). "The Association of Teacher Resourcefulness with Children's Characteristics." In Mood, A.M., ed., Do Teachers Make A Difference? Health, Education and Welfare, Office of Education (OG-58042).

117. ________. (1969). "Equal Protection and School Resources." Inequality in Education (December 5, 1969): 9-16.

118. Miner, J. (1963). Social and Economic Factors in Spending for Public Education. Syracuse, New York: Syracuse University Press.

119. Mollenkopf, W. and D.S. Melville. (1956). "A Study of Secondary School Characteristics as Related to Test Scores." Princeton, New Jersey: E.T.S. (RB-56-6).

120. Mood, A.M., ed. (1970). Do Teachers Make A Difference? Health, Education and Welfare, Office of Education (OG-58042).

121. Morley, R. and L.F. Eastwood. (1975). "Alternate Communication Network Designs for an Operational PLATO IV CAI System." Center for Development Technology Report, Washington University.

122. Morris, A.J. (1974). "Cost Effectiveness of Continuing Engineering Studies by TV." Final Report on Continuing Education Studies, American Society of Engineering Education.

123. Mosteller, F. and D.P. Moynihan. (1972). On Equality of Educational Opportunity, New York: Random House.

124. Moyer, M. and D. Paden. (1968). "On the Efficiency of the High School Economics Course." American Economics Review 59: 870-877.

125. Nachman, M. and S. Opochinsky. (1958). "The Effects of Different Teaching Methods: A Methodological Study." Journal of Educational Psychology 49 (October 1958): 171-90.

126. Oates, W. and R. Quandt. (1970). The Effectiveness of Graduate Students as Teachers of Principles of Economics." Journal of Economics Education 1: 130-138.

127. Paden, D.W. and M. Moyer. (1969). "The Relative Effectiveness of Three Methods of Teaching Principles of Economics." Journal of Economics Education 1: 33-45.

128. Paden, D.W. (1962). "Teaching Economics Via Television at the College Level." In Meaney, J.W., Televised College Courses Washington, D.C.: Academy for Educational Development.

129. Perl, L.J. (1973). "Family Background, Secondary School Expenditure, and Student Ability." Journal of Human Resources (Spring 1973): 156-180.

130. ________. (1971). "Graduation, Graduate School Attendance and Investments in College Training." New York: National Bureau of Economics Research.

131. ________. (1970). "The Role of Educational Investment and Family Background in the Production and Distribution of Educational Services." PhD dissertation, University of California, Berkeley.

132. Pfouts, R.W. (1961). "The Theory of Cost and Production in the Multiproduct Firm." Econometrica 29: 650-658.

133. Plowden Report. (1967). Children and Their Primary Schools, Great Britain.

134. Polachek, S. et al. (1975). "Educational Production Functions: A Microeconomic Analysis." Chapel Hill, North Carolina: University of North Carolina.

135. Raymond, R. (1968). "Determinants of the Quality of Primary and Secondary Public Education in West Virginia." Journal of Human Resources 3 (Fall 1968): 450-470.

136. Ribich, T.I. (1968). Education and Poverty, Washington, D.C.: Brookings Institute.

137. Saettler, P.H. (1968). A History of Instructional Technology, New York: McGraw Hill.

138. Sewell, W. and V.P. Shah. (1967). "Socioeconomic Status, Intelligence, and the Attainment of Higher Education." Sociology of Education 40 (Winter 1967): 1-23.

139. Shaycoft, M.F. (1967). The High School Years: Growth in Cognitive Skills, Pittsburgh: American Institute for Research and University of Pittsburgh, Project Talent.

140. Skornia, A.J. (1968). "Television, Radio, and Other New Media in Education." IEEE International Conference on Communications Conference Record 4: 854-58.

141. Simonsen, R. and K.S. Renshaw. (1974). "CAI-Boon or Boondoggle?" Datamation 20 (March 1974): 90-102.

142. Smith, M.S. (1971). "The Brookline Study." Harvard Center for Educational Policy Research.

143. ________. (1972). "Equality of Educational Opportunity: The Basic Findings Reconsidered." In Mosteller and Moynihan, On Equality of Educational Opportunity, New York: Random House.

144. ________. (1968). "Equality of Educational Opportunity: Comments on Bowles and Levin." Journal of Human Resources 3 (Summer 1968): 384-389.

145. Solomon, L. (1973). "Schooling and Subsequent Success." And "The Definition and Impact of College Quality." In Solomon and Taubman, eds. Does College Matter? New York: Academic Press.

146. Solomon, L.C. and P.J. Taubman, eds. (1973). Does College Matter? New York: Academic Press.

147. Summers, A. and B. Wolfe. (1975). "Equality of Educational Opportunity Quantified: A Production Function Approach." Presented at the Winter meetings of the Econometric Society (December 27, 1974).

148. ________. (1975) "Which School Resources Help Learning? Efficiency and Equity in Philadelphia Public Schools." Federal Reserve Bank of Philadelphia Business Review (February 1975).

149. Suppes, P. and M. Morningstar. (1969). "Computer Assisted Instruction." Science 166: 343-350.

150. Thomas, J.A. (1962). "Efficiency in Education: A Study of the Relationships Between Selected Inputs and Mean Test Scores in "A Sample of Senior High Schools." PhD dissertation, Stanford University.

151. ________. (1962). "Efficiency in Education: An Empirical Study." Administrator's Notebook (October 1962).

152. Timmer, C.P. (1969). "On Measuring Technical Efficiency." PhD dissertation, Harvard University.

153. Tuckman, H.P. (1971). "High School Inputs and Their Contribution to School Performance." Journal of Human Resources 6 (Fall 1971): 490-509.

154. Wallhaus, R.A. (1975). "The Many Dimensions of Productivity." in Wallhaus, ed. Measuring and Increasing Academic Productivity, San Francisco: Jossey Bass.

155. Wells, S.J. (1974). "Technology Efficiency and Educational Production." PhD dissertation, Department of Economics, Stanford University.

156. Wells, S., B. Whelchel, and D. Jamison. (1974). "The Impact of CAI on the Academic Performance of Disadvantaged Students." Princeton, New Jersey: Educational Testing Service.

157. Wilkinson, G.L. (1972). "Needed: Information for Cost Analysis." Educational Technology (July 1972).

158. Winkler, D.R. (1975). "Educational Achievement and School Peer Group Composition." Journal of Human Resources 10 (Spring 1975): 189-204.

159. ________. (1972). "Production of Human Capital: A Study of Minority Achievement." PhD dissertation, Berkeley, California.

160. Wong, M. (1974). "The Role of Technology in Non-Traditional Higher Education." Center for Development Technology, Washington University, St. Louis, Missouri.