CONCEPTUAL AND EMPIRICAL ISSUES IN THE ESTIMATION OF EDUCATIONAL PRODUCTION FUNCTIONS*

ERIC A. HANUSHEK

ABSTRACT

Measuring educational performance and understanding its determinants are important for designing policies with respect to such varying issues as teacher accountability, educational finance systems, and school integration. Unfortunately, past analyses of student achievement and educational production relationships have been plagued by both a lack of conceptual clarity and a number of potentially severe analytical problems. As a result, there is considerable confusion not only about what has been learned, but also about how such studies should be conducted and what can be learned. This review considers each of these issues and also relates knowledge from these studies to research about areas other than just school operations and performance.

Despite a substantial and growing volume of research into educational production relationships and the determination of student achievement, considerable confusion remains about how such studies should be conducted, how past analyses should be interpreted, and what has been and can be learned from such studies. These studies are interesting and important from a number of perspectives. First, they exemplify the difficulty frequently encountered in the empirical application of some basic economic models, and the lessons there apply to a wider class of problems than just understanding schools. Second, the results of these studies have ramifications for a variety of analyses that do not focus on schools, such as wage determination, status achievement, the financing of schools, and the impacts of quality of education on urban location and housing choice.

The author is Professor of Economics and Political Science, and Director, Public Policy Analysis Program, University of Rochester.

* This paper benefited from helpful comments on previous versions by Charles Bidwell, Mary Curran, Henry Levin, Richard Murnane, Edward Pauly, John Quigley, James Rosenbaum, and especially Richard Nelson. The research was largely done while the author was at Yale University. [Manuscript received May 1978; accepted November 1978].

However, perhaps most important—and a distinguishing feature of this research—is that the concluding sections of these papers on "policy implications" do not have the more common hollow ring because the results of this research have frequently entered into judicial proceedings, legislative debate, and executive branch policy deliberations. Unfortunately, this research has been unusually difficult to follow, in part because its development has frequently crossed disciplinary boundaries.

I. SOME BACKGROUND

The first, and perhaps still the most influential, study is *Equality of Educational Opportunity,* or the "Coleman Report" [37]. That study, mandated by the Civil Rights Act of 1964, was startling in a number of ways. First, the survey information for over half a million students, containing data not only about students and the characteristics of their more than 3,000 schools, but also about their achievement in school, provided the most complete description of elementary and secondary schools ever produced in this country. Second, and most relevant for this discussion, it directed attention to the importance of the relationship between school inputs and student achievement.[1] Finally, it introduced into the public policy arena a bewildering array of technical and esoteric issues such as statistical significance, analysis of covariance, production efficiency, multicollinearity, residual variation, estimation bias, and simultaneous equations.

The attention paid to the input-output analysis in the Coleman Report clearly reflects the direct policy importance of the analysis. Such information is critical not only to "school management," but also to such diverse policy issues as school integration, accountability in schools, and the finance of elementary and secondary schools. The policy relevance of input-output studies has led to both a rapid growth in number of analyses and a concerted effort to interpret the many different, and apparently contradictory, results (e.g., Hanushek and Kain [66], Bowles [21], Bowles and Levin [23, 24], Cain and Watts [28], Averch and others [11], Levin [91]).

As economists entered this area, the relationships estimated became known as "educational production functions" instead of simply input-output analyses. This was more than a semantic change, however; the term *production function* has special connotations that alter the interpretation of the results. In fact, one part of the discussion of these analyses is whether or not the models estimated are production functions in the usual technical meaning of the term.

1 This perspective on the analysis of schooling, as pointed out in the response of Coleman [38] to his critics, was perhaps the most important aspect of the "Coleman Report."

II. TEXTBOOK ANALYSES

All sophomore economics classes develop the concept of a production function and its uses in analyzing the decisions of firms. In its usual abstract presentation, a firm's production possibilities are assumed to be governed by certain technical relationships, and the production function simply describes the maximum output feasible with different sets of inputs. The key distinction between a "production function" and any number of alternative descriptions of input and output relationships is the notion that it represents the maximum achievable output for given inputs. Firms are conceptualized as attempting to maximize profits through decisions about level of production and mix of inputs (given the production function, product demand, and input prices). With some embellishments, this represents the motivation for and theoretical backdrop to production function studies.[2]

The production function, along with the related theoretical apparatus of optimal firm decisions, is a powerful pedagogical tool, since it provides a basis for describing efficient production, the appropriate response of firms to changes in technology or input costs, and so forth. Further, the basic analytical constructs seem applicable to a wide variety of applications—there is a priori no indication that this structure applies to, say, the steel industry and not the education industry.

These theoretical concepts are, however, deceptively simple. When taken out of the classroom, they often require substantial modification. At one level, the concept of a technologically based production function may, in reality, not be particularly applicable in a variety of actual situations.[3] At another level, even if appropriate, the actual production function is generally

---

2 This analysis can be extended to consider multiple outputs or wider ranges of inputs. Additionally, certain assumptions about the form of the production relationships can be included. For example, declining marginal products for inputs are often assumed; that is, the additional output per unit from increasing a given input might be expected to decline as larger quantities of the input are used (holding constant other inputs). One might also believe that certain inputs are complements in production; that is, the value of a given input might increase as more of another (complementary) input is used. Finally, the possible effects of different scales of operation can be incorporated.

3 The standard "textbook" view is that production functions are derived from known engineering relationships that reflect exogenously given technological processes. The firm decides upon a mix of inputs, and the best process for combining these inputs is indicated by the production function and, therefore, does not have to be explicitly considered. A recurring theme, elaborated below, is that this may not be a good characterization of many production processes and firm decisions where: (1) individuals involved in production actually have considerable discretion in choice of process; (2) the "best" process might not be generally known and uncertainty is important; or (3) dynamics are important and the production technology is changing. For a general statement of these issues, see Nelson and Winter [109, 110, 111].

not known a priori and must be estimated based upon the observed operations of firms. However, such estimation raises a series of conceptual and statistical problems.

Although some significant differences exist in the application of production functions to education and to other industries, the biggest differences probably arise from the potential uses of the analysis. Few people would expect manufacturing firms to change their behavior given estimated production functions for manufacturing industries (see, for example, Hildebrand and Liu [71] or Griliches [54]), and there is very little temptation to prescribe any public policies based upon the results. The same cannot be said for education. Congress holds hearings on the size of estimated coefficients (see, for example, [141]), commissions include results in supporting their policies (for example, [47] or [116]); and courts receive testimony about regression equations.[4] Because the findings and interpretations of educational production functions go far beyond answering a set of esoteric questions of economists, educational production functions have been discussed more widely, and the confusion surrounding them seems somewhat greater, than production function estimation in general.

III. CONCEPTUAL PRODUCTION FUNCTIONS AND EDUCATIONAL REALITIES

Studies included under the rubric educational production functions are generally statistical analyses relating observed student outcomes to characteristics of the students, their families, and other students in the school, as well as characteristics of schools. Most frequently, student outcomes are measured by various standardized test scores, although attitudes, college continuation, and attendance patterns have also been analyzed. These studies also diverge considerably in terms of the actual measured inputs; in terms of the level of aggregation of both dependent and independent variables (e.g., individual student, school average, or district average observations); and in terms of the precise statistical methods. Not surprisingly, given such differences, the conclusions of the various studies appear to be very different—and often apparently contradictory.

This paper considers the major conceptual and empirical issues in such analyses with an emphasis upon the implicit assumptions and alternative interpretations of past models and results. While there is no attempt to review systematically individual studies, a later section summarizes the

4 For example, *Keyes* v. *School District No. 1,* 413 U.S. 189 (1973); *Serrano* v. *Priest,* 5 Cal.3d 584 (1971); *Hobson* v. *Hansen,* 269 F.Supp. 401 (D.D.C.1967).

major findings and ambiguities along with indicating some areas of profitable future research.[5]

A. Measurement of Output

The vast majority of production function studies measure output by standardized achievement test scores. A few have also considered other measures such as student attitudes (Levin [89*f*, Michelson [101], Boardman, Davis, and Sanday [18]), attendance rates (Katzman [80]), and college continuation or dropout rates (Katzman [80], Burkhead, Fox, and Holland [26]). Are any, all, or none of these sensible measures of educational output?

While standard production theory concentrates upon varying quantities of a homogeneous output, this is not easily translated into an educational equivalent. Education is a service which transforms fixed quantities of inputs (i.e., individuals) into individuals with different quality attributes. Educational studies rightfully concentrate upon "quality" differences. However, simply because individuals can be ordinally ranked in terms of cognitive test scores does not imply that such a measure is necessarily appropriate.

Perhaps the most important concern with standardized tests is the lack of external validation. These tests do discriminate among individuals; that is, they can divide the population into different groups. However, questions are generally selected by criteria internal to tests: (a) their ability to divide students (so that questions that can be answered by all or none of the relevant population aren't useful); and (b) their consistency with other questions (i.e., whether individuals getting a given question right tend to get other questions on the test right). Further, a given text should produce the same score if taken at different times by the same individual, and slightly different wordings of questions covering the same concept should yield the same results. None of these relates directly to whether or not tests cover material, knowledge, or skills valued by society.

Clearly, much of the observed interest in school system performance relates to the perceived importance of schooling to future capabilities of students—the ability of students to cope with and perform in society *after* they have left school. To be sure, there is some value to knowledge for its own sake, other things being equal, and more knowledgeable individuals

5 Perhaps the most thorough review of the findings per se is still Averch and others [11]. However, this is now somewhat out of date. A number of noteworthy studies do not appear in that volume: Central Advisory Council [32], Levy [92], Boardman, Davis, and Sanday [18], Garner [50], Perl [115], Heim and Perl [69], Summers and Wolfe [133, 134], Murnane [106], Winkler [149], Jencks and Brown [77], Henderson, Mieszkowski, and Sauvegeau [70], Armor and others [6], Ritzen and Winkler [120], Winkler [150], Maynard and Crawford [100], Link and Ratledge [96].

may be more interesting, happier, or whatever. However, if schools were perceived to perform a simple monastic role, it is inconceivable that they would receive the attention and interest that they do. Here we consider two dimensions of school effects: the effect on labor market performance and the effect on socialization—that is, political awareness, citizenship, moral values, etc.

Economists have analyzed the influence of education on earnings and labor market performance (see reviews by Mincer [103] and Rosen [123]). Sociologists have explored the effects of schooling on occupational choice, mobility, earnings, and the relationship between schooling and personal and family characteristics (Sewell and Hauser [126], Boudon [20], Jencks and others [76], Alwin [4], Duncan, Featherman, and Duncan [44], Jencks and Brown [77], Blau and Duncan [17]). These studies direct attention to the critical question of what role formal education plays in influencing later lives of citizens—a focus frequently lost in research into school operations.

However, a recurring problem with such studies is the inadequate measure of the education individuals receive. Most commonly, years of schooling is used to measure education. (This is even the case in models of "human capital production functions"; see Ben-Porath [13].) Few measures of the quality of education have been incorporated in such studies. Since the most pressing school policy questions concern how to improve the quality of education, this is a particularly unfortunate limitation.

Some attempts have been made to incorporate qualitative measures, such as information about cognitive abilities of individuals or about school expenditure levels into labor market studies.[6] Such studies have been severely limited by data availability, the necessity to use fairly peculiar samples, and stringent assumptions about school operations. (For example, expenditure studies *assume* expenditure differences index quality differences.) Futher, the results with respect to the effects of quality differences have been quite inconclusive. Thus, while these studies offer an important perspective on how to observe educational outcomes, they do not currently provide much guidance to studies focusing on the operations of schools.[7]

6 See Welch [145, 147], Weisbrod and Karpoff [143], Ashenfelter and Mooney [10], Rogers [122], Weiss [144], Hansen, Weisbrod and Scanlon [60], Hanushek [62, 64], Johnson and Stafford [78], Morgenstern [104], Taubman and Wales [136], Solmon [129], Link and Ratledge [95], Jencks and Brown [77], Ribich and Murphy [117], Lee [87], Wachtel [142], and Akin and Garfinkel [2]. Only the Welch studies and the Jencks and Brown study attempt to consider the operations of schools.

Research on "ability" and earnings (e.g., Griliches and Mason [56] or the review in Griliches [55]) is also related if ability is considered endogenous.

7 Related research into evaluating the effects of various manpower and job-training programs on subsequent labor market performance (e.g., Cain and Hollister [27], Ashenfelter [9], Kerachsky and Mallar [81], and Kiefer [82]) offer a similar perspective.

Although the relationship of schooling and labor market performance is central to many policy questions, it is not the only area of interest. Hence, studies have also examined the role of education in increasing job satisfaction (Duncan [43], Black [16]), in maintaining personal health (Grossman [57], Manheim [97]), and in increasing the productivity of mothers engaged in household production, as well as the effects of the mother's education on the learning of young children (Hill and Stafford [72], Leibowitz [93], Lindert [94], Inman [73]). Further, political scientists have considered the effect of education on political socialization and voting behavior (Campbell and others [29], or the review by Niemi and Sobieszek [112]), and sociologists have considered the relationship between education and criminality.

While these studies have suggested some gross effects of quantity of schooling on other life outcomes, they virtually have never addressed the question here—how such outcomes vary in response to differences in school programs and operations. Analysis of nonlabor market areas—if refocused toward the performance of the educational system and qualitative differences in schooling—does have the potential for providing a more balanced perspective on educational productivity. Nevertheless, existing studies have yielded inconclusive results about the effects of even quantity of schooling, let alone the more detailed information.

A more fundamental shortcoming is the superficiality of the conceptual notions of the mechanisms by which education affects skills and later experiences. Cognitive skills, the chief measure of educational quality, may not be the only, let alone the most important, outcome of schooling in determining individuals' future success. One might think that more educated individuals can accomplish given tasks better or more swiftly, but surely this holds for only certain types of jobs. Less education may even be better in jobs requiring manual skills or jobs that are very repetitive. One rather commonly held presumption is that better educated individuals are able to perform more complicated tasks or are able to adapt to changing conditions and tasks (see Welch [146] and Nelson and Phelps [108]). This hypothesis has important implications for studying the productivity and outputs of schools through understanding of the mechanisms by which school interacts with the work place. Such understanding could provide considerable insight into how to measure the outcomes of schooling (or at least where to look) and how these outcomes might change with the character of the economy. The lack of conceptual clarity holds equally for other potential outcomes of the educational system.

While these studies have in general not considered programmatic differences in any detail, such consideration could provide additional insights into the characteristics of educational programs that are important for future success.

The uncertainty about the source of schooling-earnings relationships is also highlighted by the recent attention to "screening" aspects of schooling. Schools may produce more qualified individuals or may simply identify the more able. The latter view has been the subject of both theoretical and empirical treatment by economists and sociologists (Spence [130], Taubman and Wales [136], Berg [14], Thurow [138], Thurow and Lucas [139], Riley [119], Arrow [7], Stiglitz [131], Wolpin [152], Layard and Psacharopoulos [86]). Most of the attention paid to screening models arises from the implication that the social value of schooling may be considerably less than the private value (that is observed in earnings relationships) if schools are merely identifying the more able instead of actually changing their skills. Further, the screening model suggests both possible reinterpretation of the historical contribution of education to economic growth (see Denison [41]) and revisions of expectations about future returns to schooling. (These revisions depend upon the "quality" of the screening function as schooling distributions change and the response of firms to any such changes.) However, there are also direct implications of the screening model for the measurement of educational outcomes and the analysis of educational production relationships. In a screening model, the output of schools is information about the relative abilities of students, and this would suggest that more attention should be directed toward the distribution of observed educational outcomes (instead of simply the mean outcomes) and their relationship to the distribution of underlying abilities. Further, the interpretation of some studies, such as those of school dropout rates discussed below, might be radically altered, since schools with a higher dropout rate might actually be providing better information (higher output) than those with lower rates—an interpretation that is very different from that of the authors of these studies. Unfortunately, no persuasive test has been devised to distinguish between a screening model and the more standard "production" model.

These two views—production and screening—are also not the only models explaining subsequent performance. For example. Jencks and others [76] argue that luck and personal characteristics (that are unrelated to schooling) are the most important determinants of earnings differences. Bowles and Gintis [22] believe that earnings differences arise chiefly from the existing social structure and that schools adjust to instead of determine subsequent outcomes. While these latter two views are not completely convincing, available evidence does not conclusively differentiate among these four divergent views.[8]

8 The Jencks and others [76] conclusions rest upon the finding that regression analyses of earnings include sizable unexplained variation which they label "luck." However, since these analyses are based upon a small number of crudely measured characteristics (years of

Referring back to the original question, we find simply a large degree of uncertainty about the appropriateness of test scores as outcome measures. While the various studies of lifetime outcomes are conceptually very relevant to measuring school outputs, they have not been particularly illuminating for the study of school production functions. While it would not be particularly surprising if standardized test performance was not highly correlated with future success (since the standard test construction methodology makes little pretense of relating test performance to any external criteria), available empirical evidence is inconclusive about whether or not there is some fortuitous linkage between test scores and subsequent achievement.

Nevertheless, performance on tests is being used to evaluate educational programs, and even to allocate funds, and there are some pragmatic arguments for the use of test scores as output measures. Besides their common availability, one argument is that test scores appear to be valued in and of themselves. To a large extent, educators tend to believe that they are important, albeit incomplete, measures of education. Further, parents and decision-makers appear to value higher test scores—at least in the absence of evidence that they are unimportant. (Note the continued pressures to make scores more publicly available.)

A more persuasive argument for the use of test scores relates to continuation in schooling. Almost all studies of earnings which include both quantity of schooling and achievement differences find significant impacts of quantity that are independent of achievement differences.[9] This implies that achievement differences do not adequately measure all skill differences. However, at the same time, test scores appear to have an increasing use in selecting individuals for further schooling. Thus, they may relate directly to

---

schooling, age, and perhaps measured achievement-test scores or family background), it should not be surprising that much is left to be explained. Moreover, there is little basis for labeling our ignorance (the residuals from a regression analysis) in any particular manner. Direct analysis of individual earnings over time (Hanushek and Quigley [67]) indicates that about two-thirds of the unexplained variation in earnings models represents stable, but unmeasured, individual factors.

The Bowles and Gintis conclusions [22] in part rest on similar evidence—measures of IQ or cognitive ability differences do not explain much of the variation in individual earnings. This, combined with an analysis of the historical development of U.S. schools, is used to support their "social structure" view of earnings determination. However, their analysis relates solely to the U.S. economy and the U.S. schooling system. There is no independent analysis of how differences in social structure affect earnings possibilities, even though this is presumably the relevant evidence.

9 An exception is Hansen, Weisbrod, and Scanlon [60], but also see the comments on this study by Chiswick [33] and Masters and Ribich [98]. See also Gintis [51] for a review of some of the literature on this topic up to about 1970.

relate directly to the "real" outputs through the selection mechanism.[10]

Finally, a few miscellaneous issues about output measurement should be added. First, if one does use test score measurements, there are a number of choices, related simply to the scaling of scores. Tests are often available in "grade level" equivalent, percentile ranking, or raw score forms, all of which provide the same ordinal ranking (except for the possibility of some compression of the rankings). Yet, for most statistical work, one wants a scale which indicates how different individuals are rather than one that simply ranks them. The choice really depends upon the relationship of these estimates of output to the subsequent outcomes, which are not known.[11] Second, there is some movement toward criterion-references tests—tests that relate to some set of educational goals. The crucial issue is the development of goals. The previous discussion argues for goals that relate to performance outside of schools, but it is not obvious that these goals guide much of the current development work.

B. Multiple Outputs

Most educational production function studies have analyzed a single output or, alternatively, a series of output measures without consideration of their interactions. (Exceptions include Levin [89], Michelson [101], Boardman, Davis, and Sanday [18], and Brown and Saks [25].) If indeed the educational process is best characterized as producing a set of outcomes (say, cognitive skills and political awareness) and if there are important interactions among them in production, then interpretation of commonly estimated models for a single outcome becomes complicated.

Ordinary least squares (OLS) regression analysis, which is commonly used in analyzing production functions for single outcomes, is generally inappropriate when there are multiple outcomes that are simultaneously produced. In the simplest case, assume that there are two outputs, one of

---

10 This argument is found in Dugan [42]. The use for predicting future school performance and for selection is also central in Wirtz and others [151].

The interpretation of tests and their use in selection may be changing, however. On the one hand, at least by newspaper accounts, there is a growing concern about the information contained in test scores. On the other hand, courts are increasingly becoming concerned with the use of tests for selection, particularly when they might have discriminatory outcomes. For example, in *Griggs* v. *Duke Power Co.*, 401 U.S. 424 (1971), and a host of similar cases, a central issue is whether test performance relates to job performance.

11 This is actually just a special case of more general questions about the functional form of production functions (discussed below). Coleman and Karweit [39] argue against use of grade-level equivalents on the basis of their peculiar properties even for answering qualitative questions. Whether or not tests measure accurately the activities and learning that take place in schools has not been considered here. For the purposes here, such concern simply deserves little attention if unrelated to subsequent outcomes.

which is an "intermediate" outcome (such as attitudes toward school) that is not valued itself, and one of which is a "final" outcome (such as achievement) that is valued by decision-makers; also, assume that the underlying production relationships (the structural equations) are such that, in addition to a series of exogenous factors (such as family background and schools), attitudes affect achievement and, simultaneously, achievement affects attitudes. In such a situation, both structural equations can generally be estimated simultaneously (although not with OLS methods). Alternatively, it is possible to estimate the "reduced-form" equation for the separate outcomes (where the reduced form is the relationship between one of the outcomes and the exogenous variables and is found by substituting one structural equation into the other). The reduced-form equation, which can generally be analyzed with OLS techniques, indicates both the direct and indirect impacts (through the other outcomes) of the exogenous variables. The reduced-form estimates do not indicate the process by which an exogenous variable affects an outcome and may be misleading if one undertakes policies that change the structural relationships. However, there are no quarrels with the underlying statistical methods or the judicious interpretation of the results.[12] Many single-equation analyses can be interpreted simply as attempts to estimate reduced-form relationships.

However, alternative multiple-outcome models are more complicated. Consider again the case of two outcomes except now let both be "final" outcomes that are independently valued by decision-makers. With information about the alternative outputs, inputs, and decision-makers' valuation of outputs, the structural equations can again be estimated directly.[13] However, with information about only one output, estimation of the reduced form might be quite misleading.[14] The estimated effects of various inputs will reflect both the production technology (the effect of each input

---

12 Underlying this is a series of complicated statistical arguments and assumptions. For example, there are the issues of identification, the distribution of the error terms, etc. There are also specialized forms of the structural models—such as recursive models—which can be estimated by OLS. Discussion of these issues along with discussion of the desirability and methods of estimating the structural equations can be found in Hanushek and Jackson [65, Chs. 8 and 9].

Reduced-form estimation still requires some specification and measurement of the other equations in the system, which may be difficult. For example, with attitude formation, little is known about the determinants of attitudes, and clearly experiences outside schools (for which data are often lacking) are important.

13 There is, in reality, little information about decision-makers' choices among outcomes. This makes estimation of the structural equations very difficult, even when data are available about the relevant outputs and inputs.

14 In the general case, the errors in the reduced-form equation will be correlated with the exogenous factors through the decision function about different outputs.

on the single output) *and* the choice between outputs, not simply the production technology.

The empirical importance of this issue is generally unknown. It depends importantly upon the degree of "jointness" of production, the form of the production function, the variance of choices, the underlying decision rules for determining choices, and the accuracy of measuring inputs. It is possible to construct models where joint production appears extremely important (e.g., Brown and Saks [25] develop a model where both mean and variance of achievement are valued and where simple reduced-form estimates appear quite misleading). However, there is a wide variety of circumstances where such issues are inconsequential.[15] Without more information about both the range of outputs (and their measures) and the potential decision rules, there is little that can be said about this problem.

Given measures of alternative outputs, it would be appealing to look at production functions for "total" output. This is essentially what is done for production function estimates in other sectors, where market prices are used to aggregate outputs. However, these prices are not available for education. Even if available, they may be inappropriate since the weights in the decision function for outputs may differ from the market prices.[16]

Consideration of multiple outputs does suggest that production functions estimated with test-score measures might be more appropriate in earlier grades, where the emphasis tends to be more on basic cognitive skills —reading and arithmetic—than in later grades. In other words, these outputs appear to be much more heavily weighted than others at earlier grades, and therefore the potential problems of multiple outputs are less than in later grades.[17]

C. Inputs to the Production Process

The usual prescription for developing the relevant set of inputs to a produc-

15 Take a simple example where one is concerned with alternative outputs which are independently produced; say, one is concerned with reading ability and sex education and they do not interact (i.e., each does not appear in the other structural equation). As long as we have accurate measures of all exogenous variables (discussed below), estimation of a single equation may not be affected by the decision process which weights the two outputs. Alternatively, if the relative weights placed upon the two outputs vary dramatically—say, cognitive skills are emphasized much more than other skills—the problems may be empirically insignificant.

16 This is not an issue in studying competitive industries where it is assumed that managers maximize profits; thus, the weights in the output decision function are simply the market prices.

17 Note that all production function studies have been conducted for elementary and secondary schools. In postsecondary education, few people believe that test scores adequately measure outputs.

tion process is to find an engineer who will describe the technical characteristics and specifications of the process. When considering education, the "engineers" are usually thought to be learning theorists. Nevertheless, almost all educational analyses begin with laments about how we do not have any learning theory that is suitable for guiding input-output analyses. In reality, engineers give little guidance to the development of production functions developed for any sector. Typical production function studies, say, for manufacturing industries, incorporate two or possibly three inputs (capital, labor, and possibly education level of labor), and few would argue that engineers had much to do with this specification. This set of inputs does, however, match the pedagogical models and, while there are some minor debates about such issues as the measurement of capital, generally receives limited attention.

In education, the relatively fixed input of labor and capital (i.e., one teacher per classroom with a relatively small variance in class size) implies that this simple description of inputs could explain little. Somewhat ironically, because educational studies have attempted to provide much more detail about input differences, they have been faced with much more criticism about the specification of the inputs.[18]

Part of this criticism is explained by the fact that input specification has not received much attention in many past analyses. There is little conceptual clarity, and the choice of inputs seems, sometimes explicitly, to be guided more by data availability rather than any notions of conceptual desirability. For example, nowhere in the Coleman Report can one find a statement of an underlying conceptual model. At times such a model seems implied, but the statistical analyses do not seem to relate to the implied model (see Hanushek and Kain [66]).

Conceptually, a model such as equation (1) seems generally acceptable:

$$A_{it} = f(B_i^{(t)}, P_i^{(t)}, S_i^{(t)}, I_i) \tag{1}$$

where, for the ith student, A_{it} = achievement at time t; $B_i^{(t)}$ = vector of family background influences cumulative to time t; $P_i^{(t)}$ = vector of influences of peers cumulative to time t; $S_i^{(t)}$ = vector of school inputs cumulative to time t; and I_i = vector of innate abilities. The development of this model and background analyses entering it are discussed elsewhere (Hanushek [61]), and this discussion will only highlight important issues.

In the abstract, it is difficult to quarrel with this specification; controversy enters only when more detail about the definition and measurement of variables and the form of the functional relationship are introduced. The first important point is that the inputs are those that are relevant to the

18 Again, part of the attention to model specification relates to the very different purposes behind educational analyses and analyses of other sectors (see above).

individual student. Additionally, the model portrays the educational production relationship as cumulative—past inputs have some lasting effect, although the value in explaining output may diminish with more distant inputs. A corollary to this last point is that, without fairly strong assumptions about the dynamics of education—that is, the time paths of adjustment to change—the data requirements are huge.

In part to circumvent some of the data requirements (and in part because of other considerations discussed below), an alternative version of this model has sometimes been analyzed. If equation (1) holds at different points in time, say, a past period t^*, we can consider the change in achievement between t and t^* as in equation (2):

$$(2) \qquad A_{it} = f^*(B_i^{(t-t^*)}, P_i^{(t-t^*)}, S_i^{(t-t^*)}, I_i, A_{it^*})$$

where the inputs are measured over the period t^* to t.[19] This formulation, sometimes referred to as "value added" specification, apparently lessens the data requirements, but it does so at the expense of some additional assumptions about the relationships (discussed below).

Consider now the empirical implementation of these models. Most analyses are purely cross-sectional and include only contemporaneous measures of the inputs. No studies have adequate measures of initial endowments (or "learning capacity"). Many educational inputs (e.g., family educational inputs) are not measured directly, but instead are proxied by other observable attributes (such as socioeconomic background of the family). Little attention is given to the dynamic structure, that is, how the effects of different inputs cumulate. The relevant inputs (e.g., school factors) are often measured with considerable error.

The divergence of the conceptual model and the empirical models which have been estimated means that interpretation of the empirical results often requires a series of implicit assumptions, some of which are very dubious. The remainder of this section attempts to make explicit the most important assumptions underlying the empirical analyses.

The most consistent and obvious divergence of the empirical models from the conceptual models is the lack of measurement for innate abilities. In fact, there is little clarity about what should be measured in this term (I_i).

19 This equation results from simply subtracting equation (1) for time t^* from equation (1) for t. However, instead of analyzing $A_{it} - A_{it^*}$ as the dependent variable, A_{it^*} is put on the right-hand side. There are three reasons for doing this: (1) empirically, A_{it} and A_{it^*} may well be different tests with different scaling, etc.; (2) levels of starting achievement (A_{it^*}) may influence achievement gain; and (3) correlated errors in achievement measurement may suggest such a formulation (Cronbach and Furby [40]). However, the latter argument suggests that further corrections for errors in the exogenous variables—probably based upon test-reliability measures—are also needed since such errors, even if they have zero means, will yield inconsistent estimates.

Presumably, it includes "learning capacity," but this is not well defined.

In a regression framework, the effect of omitting an important variable is bias in the estimated regression coefficients. The importance (size of bias) is related both to the strength of the variable on achievement and the correlation of the omitted variable with other included variables in the model. If innate abilities were uncorrelated with all of the included variables, the only effect would be to increase the residual variance, and there would be no bias of the other coefficients estimated. However, there is some evidence that these correlations are not zero. If I_i is related to IQ, we know that, in particular, IQ is correlated with family background (either through genetics or environment). For example, Scarr and Weinberg [125] find that 10 to 30 percent of the variation in IQs of adolescents is explained by measured family characteristics.[20] Further, the correlations for younger children are higher. This implies that the omission of innate abilities probably biases upwards the estimated impact of family background on achievement.[21] At the same time, it is plausible to assume that biases in other parts of the model will be considerably less, particularly in the case of school inputs. The correlations between innate abilities and school attributes, after allowing for family background factors, is likely to be small.[22] Likewise, the importance of these omitted factors is lessened if the estimated model is equation (2), since "level" effect would be included in A_{it^*} and only "growth" effects of innate abilities would be omitted. (See Boardman and Murnane [19] for a discussion of potential biases in alternative specifications.)

The next major category of empirical problems is the accuracy of variable measurement, a problem which occurs in several different forms. Frequently, only contemporaneous measures of the exogenous variables are

---

20 The reported numbers are R^2s for regression equations which include family attributes (e.g., mother's and father's education, etc.) on children's IQs in biological families. The range in R^2s basically arises from the inclusion of parents' IQs. For adoptive families, the R^2s range between .02 and .16.

21 The bias is complicated in models with many exogenous variables; in that case it depends upon the sample partial correlations of the omitted variables on all of the exogenous variables. Plausible assumptions about the partial correlations, however, indicate an upward bias in family background coefficients. For details of biases, see Hanushek and Jackson [65, Ch. 4].

22 Innate abilities are probably positively correlated with school attributes and peers because higher SES families generally live in relatively homogeneous neighborhoods and because they also select, or demand, higher quality schools. However, the concern is not the simple correlations, but instead the correlations after controlling for family background differences; these are likely to be considerably smaller. On the other hand, the existence of ability-tracking might increase the correlations of innate abilities and school inputs if tracked students systematically receive different school inputs (see Rosenbaum [124] and Alexander and McDill [3] on tracking and school inputs).

available, implying that the cumulative variables are generally measured with considerable error. Even if the errors of measurement have a mean of zero, the coefficients will be biased; the amount of bias is roughly proportional to the variance of the measurement error relative to the variance of the true variable (see Hanushek and Jackson [65, Ch. 10]). In this case, however, the errors of measurement for background factors (and the biases in these coefficients) are probably less than for other factors, since current measures of backgrounds give a better picture of historical factors than either current measures of peers (because of migration or changing of schools) or current measures of school inputs. These measurement error problems are undoubtedly more severe in models such as equation (1) than for those in equation (2) where the relevant history is back to t^* rather than to birth.

The biases from "historical" measurement errors are probably most severe for schooling inputs and, to a lesser extent, for peer influences.[23] Common "contemporaneous" measurement errors probably also impact most severely on school inputs. Much analysis has tried to capitalize on readily available school data—data produced routinely for administrative purposes. Typically, these data provide measures of average teacher or school characteristics, but are not linked to individual students, as called for in the conceptual model. In fact, schools are often very heterogeneous institutions with considerable intraschool variance in staff and programs. This problem becomes more acute at later grade levels where average characteristics may give very misleading indications of the actual inputs to any given student.

Frequently, educational production functions are interpreted as if the included variables are conceptually and accurately measured, when in fact this is not the case. However, the severity of such problems differs significantly across studies and clearly explains part of the apparent inconsistency in findings. Moreover, within most studies, measurement errors are probably most important in the case of school inputs, leading in general to underestimates of the importance of school inputs.[24]

---

23 At the elementary school level, current measures of peers may not in general be too bad, i.e., may not have particularly large measurement errors. However, the problem worsens significantly at later grades where it is common to collect students from a variety of schools at a given junior high or high school. Nevertheless, for some analyses such as the effects of racial composition of peers, the situation is more complicated, even at the elementary school level. Commonly, schools observed to be well integrated (those with, say, 20 to 60 percent black students) at any point in time are actually quite unstable—they are going through a transition period. This implies that the current racial composition, particularly in the midranges, may not accurately reflect historic racial composition. Similarly, for schools where the racial composition has recently changed as with redistricting or busing, the current situation may be a poor indicator of the historic situation.

24 Background characteristics are generally measured by a variety of socioeconomic variables. Conceptually, the variables should measure the direct learning provided in the home along

Typically, the school inputs used in empirical analyses include objective measures of teacher characteristics and schools such as education levels, teacher experience, and age of school buildings. Some also include more detailed aspects of teachers such as undergraduate majors or teacher verbal ability, which can be interpreted as simply attempting to measure homogeneous "quantities" of inputs. While these fit clearly in the conceptual framework, consideration of another set of perhaps important variables (such as measures of the organizational context of schools and the educational process itself) introduces a set of conceptual problems with the production function terminology and framework. Typically (outside of education), measures of organization and process are seen as irrelevant in estimation. Production functions are interpreted as the relationship between inputs and outputs *mutatis mutandis.* Information about production possibilities is essentially viewed as being publicly available in the form of scientific and engineering knowledge, and production processes are reproducible through blueprints and machinery. The possibility of dynamic choices in process made by the actors in production is not considered, and the choice of "best" process is assumed automatically made after selection of inputs. While the appropriateness of this framework is open to question in a wide number of instances, it is particularly questionable in the case of education.[25]

In the education context, there are two separable classes of issues. First, there are observable "macro" organizational and process characteristics of the school (such as class organization, curricula, departmentalization, length of the school day, etc.) which represent clearly defined and reproducible educational practices. Second, there are aspects of the process which are difficult to disentangle from the characteristics of individual teachers (such as classroom management, methods of presenting abstract ideas, communication skills, etc.).

---

with attitudes, etc. Past research suggests that the learning environment in the home is highly correlated with SES. Empirically, the models do not appear to be sensitive to the precise measure of SES used. For policy considerations, however, it is important to note that the SES measures are only proxies for some more fundamental characteristics, and it is unlikely that changing the measured characteristic (say, current income) will have much of a short-run effect on achievement. At times we might be interested in determining whether there are any immediate effects of changing the measured SES of a family (say, through a negative income tax). However, it should be recognized that this is a reduced-form relationship (where there is another structural equation which relates learning environment in the home to current attributes) and that policies aimed at SES simply might alter the structural relationships.

25 The importance of "process" choice is apparent in any activities which involve individual "skill"; e.g., the differences between chefs is probably not just a difference in cookbooks, or blueprints. Organizational issues have been treated to some extent such as models of learning-by-doing, but in general have not been very well developed. See Nelson and Winter [111] for a more general critique of the shortcomings of the engineering view of production functions.

The first set of factors can readily be accommodated in the conceptual framework (although the actual empirical implementation may be more difficult). While decision-makers may not accurately perceive the impact of various macro organization and process choices and thus may not make the best choices, production functions can be estimated conditional upon these factors. In fact, there has been some, although not extensive, investigation along these lines.[26]

However, the second type of process effect creates more serious problems—both for the application of the general conceptual model and for the interpretation of any estimated effects. Many educational decisions are "micro" ones made by the actors themselves—mainly teachers. These are both difficult to observe and measure and, quite possibly, not easily reproduced. As a shorthand description, these factors will be referred to simply as "skill" differences. Once the possibility of skill differences, or embodied process in individuals, is introduced, the language—if not the conceptual framework—of production functions begins to fail. It is even difficult to define just what "maximum possible output" might mean since it is difficult to specify what the "homogeneous" inputs are.

There is some indication that these latter individual differences are quite important. The explanation of the apparent insignificance of macro process variables in Armor and others [6] is the great variation in implementation of overall process decisions at the classroom level. This is also supported by detailed analysis of the implementation of innovative techniques at the classroom level (see Berman and McLaughlin [15]). Finally, more direct analysis indicates that roughly only half of total teacher performance can be explained by measured teacher and classroom attributes.[27]

---

26 For example, Armor and others [6] test a variety of macro organization and process variables including techniques of reading instruction, time spent on reading, team-teaching, open classrooms, and variety of materials used, but found no significant impacts on achievement. In empirical work, organizational forms or process which represent simple, well-defined choices (such as the use of a given standard curriculum) are easily included. However, more complicated or less well-defined factors (such as departmentalization which depends not only upon the organization, but also the particular teachers) present more formidable problems that are related to the second category of process effects (below).

There have also been a large number of direct investigations of alternative processes or organizational forms, generally following experimental approaches and thus having a more narrow focus. See, for example, Jamison and others [74], Carpenter and Hall [31], Garfinkel and Gramlich [49], Gramlich and Koshel [52], Cicirelli and others [35], Armor [5], Barnow and Cain [12], Kiesling [84], Fox [48], and Rivlin and Timpane [121]. These analyses have fairly uniformly shown few achievement effects.

27 This evidence comes from an analysis which first estimates the "value added" of individual teachers through individual teacher dummy variables (Hanushek [61]) and then attempts to explain these estimated differences by measured teacher and classroom variables.

Recognition of skill differences has implications for discussions of "efficiency in production" (discussed below). It also alters our interpretation of teacher and school inputs. It is still reasonable to consider the impact of measured attributes of teachers, since many school decisions such as hiring and salary are based upon a set of these characteristics. However, the estimated impact of these measured attributes, following the above discussion, indicates the ability either to predict or to develop more skilled teachers. For example, the almost universal finding that more education of teachers has no impact on achievement can be interpreted as indicating that teacher training institutions do not, on average, change the skills of teachers. Or, alternatively, the frequent finding that class size doesn't affect achievement may arise from complicated (and unobserved) interactions with teacher process choices; therefore, while it is possible that smaller classes could be beneficial in specific circumstances, it is also true that, in the context of typical school and teacher operations, there is no apparent gain.

One implication of this discussion is that more effort should be devoted to understanding and measuring both the macro and micro organization and process characteristics of schools. This represents a distinct break from the tradition of production function analysis. There is no presumption that schools systematically choose the best process given the inputs; thus, estimates of education "technology" must be made conditional upon the chosen macro organization and process characteristics. At the individual teacher level, the estimated impact of teacher characteristics can be thought of as reduced from coefficients which include both direct effects (say, of teacher experience) and indirect effects through systematic choice of micro process.

D. Efficiency in Production

One important issue is whether or not schools are efficient in production. This has important policy implications since inefficiency indicates the possibility of increasing school outputs with no additional inputs. However, there is a prior statistical and interpretive issue: Since estimation is based upon the observed behavior of schools, the estimated relationships may not trace out the production frontier if schools are not producing the maximum output for given inputs. In such cases, the relationships will describe average behavior which may not be particularly useful in predicting how changes in inputs would affect outputs. Past discussions of efficiency have nevertheless been confused because both the concepts of efficiency being applied and the appropriate ones in this case have not been clear.

Traditionally, two concepts of efficiency are considered. Economic efficiency refers to the correct choice of input mix given the prices of inputs (and the production function). Technical efficiency refers to operating on the production frontier, that is, maximizing output for a given set of inputs. Past efficiency discussions have blurred these two concepts and, more importantly, have neglected consideration of how expanding the usual concept of production functions to recognize both macro organization and process choice and skill differences of inputs affects efficiency discussions.

Two arguments have been used to support the assertion that schools are technically inefficient.[28] First, educational decision-makers are apparently not guided by incentives to maximize profits or to conserve on costs. Second, they might not understand the production process and therefore can't be expected to be on the production frontier. The first argument, while raising the possibility of economic inefficiency, does not necessarily imply being off the production frontier unless resources are also wantonly squandered.[29] The focus of the second argument generally appears related to the importance of macro organizational and process choices. The relevance of these can be analyzed, and, importantly, their presence does not significantly alter the interpretation of empirical analyses as production functions.[30] Direct analyses of these factors, while not completely conclusive, do not indicate their overwhelming importance (see above).

The possibility of skill (or "embodied process") differences among inputs to schooling introduces a new dimension to the efficiency discussion. The standard conceptual framework indicates that, if two production processes are using the same inputs, any systematic difference in outputs reflects inefficiency. However, the concept of skill differences simply recognizes that individuals with the same measured characteristics make a series of important production decisions (reflected in behavior, process choices, etc.) that are difficult to identify, measure, and model. Therefore, it is not surprising that the same measured inputs yield variations in output, but at the same time it is difficult to label such observed variation as efficiency differences.

Introduction of skill differences does not, however, eliminate the usefulness of a general production framework. For many purposes, the

---

28 In most production function estimation outside of education, it is assumed that profit motivation dictates efficiency. An exception is Leibenstein [88] who argues that production inefficiency is more common than generally assumed.

29 In fact, if all schools were economically efficient and input prices were the same everywhere, it would not be possible to estimate production functions. The production function is, in education and in other areas, identified either by input price variations or by economic inefficiency.

30 As suggested above, production functions can be estimated conditional upon macro organization and process choice. Comparison of the "best" technology with alternative ones then provides estimates of this source of inefficiency. Clearly, if these factors are important and correlated with observed input usage, estimation not considering them would be misleading (i.e., would describe average relationships).

desired information is what aspects of teaching can be replicated (or predicted) in different situations. Most research has concentrated upon systematic, measurable characteristics—the reduced-form models of teacher effects—and these estimates do indicate what can be replicated in the absence of shifts in the underlying structure.

Some research has in fact estimated the total effects of individual teachers without regard to actual measurement of underlying attributes and confirms that important dimensions of teacher quality are not captured by measured teacher attributes.[31] An important sidelight of such investigations is that decision-makers might be able to identify underlying skill differences among teachers with fair accuracy. Murnane [106] found that principals' evaluations of teachers were highly correlated with estimates of total effectiveness. For many purposes, this is almost as good as the ability to identify differences ex ante.[32]

Finally, concern about technical inefficiency has led to some, basically nonstatistical, estimation of the production frontier.[33] Besides assuming accurate measures of both inputs and outputs, this analysis appears internally inconsistent: it is motivated by the perceived uncertainty about the production process, yet assumes that the researcher knows and measures all of the inputs to the production process. Further, the possibility of nonreproducible skill differences is totally neglected.

---

31 Several studies have been conducted without resort to just measured teacher characteristics such as experience, education, etc., but have used general covariance analysis schemes to estimate the "total" effect of teachers—in terms of both measured and unmeasured characteristics (see Hanushek [61], Murnane [106], or Armor and others [6]). Similar analysis (Hanushek [63]) for "skill" differences among principals, while somewhat limited, show no differences.

32 This indicates that accountability of individual teachers is not impossible. There is some concern, however, about just how this information might be used for educational policy. The principal observations used by Murnane took place in a situation where they were not actually used for any decisions. (For interpretation of these results, it should be pointed out that the evaluations were correlated with the value added of teachers and not simply the overall level of student achievement. Thus, it is not the case that principals' evaluations just reflected observed total performance of students.)

33 One proposed technique is linear programming analysis (Aigner and Chu [1], Carlson [30], Levin [91]). This technique effectively locates the "most productive" schools and places the production plane through just these observations. In fact, the estimation is based upon a small set of observations (equal to the dimensionality of the inputs).

A related inquiry by Klitgaard and Hall [85] attempts to identify "unusually effective" schools, i.e., schools which perform significantly better than would be expected on the basis of SES composition of students and community characteristics. Such schools are identified on the basis of residuals from achievement regressed on SES of students. However, if school attributes are important and correlated with SES, this approach can be very misleading—particularly when important "school variables" are subsequently identified. The residuals simply cannot be interpreted in this manner because they will be biased and inconsistent.

E. Miscellaneous Issues and Nonissues

A number of more detailed criticisms of estimated educational production functions have also surfaced. This section briefly discusses the major ones (functional form, level of aggregation, selection effects, multicollinearity among inputs, and general statistical methodology), along with their effect on the interpretation and conduct of production analyses.

1. *Functional Form.* Educational production functions have been estimated in a variety of forms, although most frequently variants of linear or logarithmic models.[34] Conceptually, there is little guidance about functional form.[35] Empirically, the issue of functional form appears to be a second order problem, since distinguishing among alternative functional forms is often impossible. The point is a simple one: within a limited range of variation, many functional forms look very similar. An important corollary, however, is that over wider ranges of variation, different functional forms may yield very different results, implying that predictions based upon changes that are far from current observations may be perilous. For example, while variations in class size, within the limited ranges observed, have little apparent effect on achievement, this is not necessarily the case for radically different class sizes.

2. *Level of Aggregation.* While the conceptual model is at the individual student level, much analysis—relying upon data collected for other purposes—is actually conducted at a more aggregate level, say, the school or district level. The effects on the estimates of such aggregation depend crucially upon the nature of educational relationships.

In the simplest case, when the production process is approximately linear in the same parameters for all students, OLS estimates on the aggregate data will be unbiased, although probably less precise than if individual data were available.[36] In more complicated situaions, aggregation

---

34 The variety is actually fairly large. Various authors have considered various stratifications (such as by race or socioeconomic background); e.g., Hanushek [61], Coleman and others [37], Smith [128]. Others have conducted general covariance analyses which allow unconstrained functional forms in terms of underlying descriptors of teachers (Hanushek [61], Murnane [106]). Finally, a variety of interactions among variables have been introduced (Winkler [149], Summers and Wolfe [133, 134]).

35 Linear models imply independence of the various inputs and constant marginal products, while logarithmic models allow declining marginal products but constrain the form of interactions of variables. Much production function analysis outside of education has centered on the properties and usefulness of alternative functional forms. See, for example, the work on Cobb-Douglas forms in Hildebrand and Liu [71]; constant elasticity of substitution models in Arrow and others [8]; transcendental log functions in Christensen, Jorgenson, and Lau [34]; and generalized production functions in Hanoch [59]. However, as explained previously, these analyses are difficult to translate for education.

36 More generally, even with a distribution of parameters across students, OLS estimates of the mean parameters are unbiased as long as the parameters are uncorrelated with the

has less innocuous effects. For example, if two groups of students—say, blacks and whites—have different production relationships where the differences are not easily parameterized, estimation with aggregate data yields "average" coefficients which depend upon the weighting of the two groups in the sampled observations and which are difficult to interpret. (See Stodolsky and Lesser [132] on the possibility of racial differences.)

Nevertheless, probably the most serious "aggregation" problem is really one of errors of measurement. The researcher frequently has individual data about students (say, achievement and family background), but only aggregate data about schools. The temptation is to use all available data by mixing individual characteristics with aggregate school data. However, the school factors relevant to any individual may differ significantly from the average. (Consider, for example, the situation in a large comprehensive high school.) Here, aggregation generally helps; the errors in measurement for a model of average achievement and average characteristics are almost certainly less than with individual achievement and average school characteristics.[37]

Aggregation as an errors-in-variables problem may be quite pervasive. Even with data on individual classrooms, the internal allocation of time and attention of students implies that each student might receive different inputs (see, for example, Garner [50], Wiley and Harnischfeger [148], and Karweit [79]). We don't have a good understanding of the importance of such variations. Analysis of classroom composition effects can be interpreted as attempts to model this, but the results about the importance of peer compositions are mixed.[38] As discussed previously, the accuracy of measurement of inputs is an extremely important topic, and probably much more important than just aggregation.

3. *Selection Effects and Causation.* For policy purposes, information about causal relationships between school factors and achievement is needed. However, such information (about the direction of causation) cannot come directly from the observed data and correlations, but must be introduced from a priori information about the structure of the overall

exogenous variables and have the same mean across students; see Swamy [135] for discussion of such "random" coefficient models. With both random coefficients and simple aggregation, the efficiency of estimation can generally be improved with techniques other than OLS (see Hanushek and Jackson [65]). Such results do not, however, hold when there are important nonlinearities in the production process.

37 This errors-in-variables argument is a major criticism of the original Coleman work (see Hanushek and Kain [66] and Hanushek [61]). It is particularly damaging in the analysis of variance framework of Coleman and others [37].

38 Compare Hanushek [61] with Henderson and others [70] for peer estimates at the classroom level. See also Murnane [106].

model. The primary concern in the production function setting is the effects of teacher selection and assignment mechanisms.

Consider the simple case of an observed positive relationship between teacher experience and student achievement (holding other factors constant). Depending upon the mechanism by which teachers are assigned to schools, this need not imply that increasing average experience levels in a school will increase achievement (i.e., that there is a causal relationship running from experience to achievement). If, for example, more senior teachers were allowed to choose their schools and teachers had a preference for teaching higher achieving students, then achievement would, at least in part, "cause" experience; and a policy change that increased experience would not yield the (full) effect on achievement expected from the estimated relationship. Other, and more subtle, selection effects might also occur; more educated or more intelligent teachers may, through their own efforts or the direct assignments of principals, be placed in "faster" classes.

The situation is really another case of simultaneous equation bias. The importance of these effects depends upon the importance of achievement in determining assignments of different types of teachers, and there has been little direct analysis of this. The appropriate solution is estimation of the simultaneous system (see fn. 12), but this has not been done.

Greenberg and McCall [53] analyzed a single urban school system in the early 1970s and concluded that race and socioeconomic background of students were systematically related to the selection and transfer of teachers with different education and experience levels. However, Murnane [107] suggests, from analysis of a different school system, that declining enrollments and the subsequent surplus of teachers have led to a much greater reliance on institutional rules and much less on individual teacher preferences (which was the hypothesized mechanism in Greenberg and McCall [53]).

Nevertheless, the potential problems arise from achievement affecting selection, and not from family background, race, or other factors that are included on the right-hand side of the estimated model affecting selection. In the latter instance (which would be a recursive structure), even though some correlation among the right-hand side variables may be induced by this mechanism, there are generally not serious problems; without other such selection effects, the estimated relationships with achievement can plausibly be interpreted as causal relationships. Clearly the severity of the problem is related to the structure of the model estimated and in many instances is only serious in the presence of fairly subtle selection mechanisms (particularly in a "value-added" specification).

4. *Multicollinearity.* Since discussion of multicollinearity in educational research by Bowles and Levin [24], it is taken as an almost ever-present, but lamentable, fact of life in most estimation. In fact, it is the first item

discussed under "Analytical Problems" in the review of production function research by Averch and others [11]. Fortunately or unfortunately, multicollinearity does not appear to be the villain it has been made out to be, although it may partially explain some of the apparent inconsistencies in existing research.

The statistical story is that disentangling the separate effects of exogenous variables which are very highly intercorrelated can be difficult. Further, in the usual case of positive intercorrelations, the parameter estimates themselves will tend to be negatively correlated so that quite commonly a coefficient has the "wrong sign" because of the correlations of variables.

Nevertheless, the importance of multicollinearity is probably overrated. All correlations of exogenous variables do not have serious consequences, and all low *t*-statistics and wrong signs are not the result of multicollinearity. Right-hand side variables are often called independent variables, but this does not imply that they cannot be correlated. In fact, multiple regression analysis is used *because* there are correlations among the "independent" variables.

The importance of multicollinearity depends crucially upon the statistical methods used.[39] Further, the diagnosis of problems is sometimes difficult.[40] In micro or individual level data, such as those frequently used, the real estimation problems caused by multicollinearity are almost certainly not as severe as citations would indicate. This problem could, however, explain some of the variation in findings across studies since, in the absence of an agreed-upon theory of what variables should be included in the models and how they should be measured, researchers frequently determine model specification on the basis of coefficient significance tests. Variations in the sample intercorrelations will yield variations in model specifications under such criteria.

---

39 Bowles and Levin [24] were correct in the importance of multicollinearity for the Coleman analysis. However, the important feature in that criticism, that is often overlooked or misunderstood, is the interaction between the choice of statistical analysis (analysis of variance) and the correlation of exogenous variables. In the analysis of variance framework, *any* correlation causes trouble. See also Hanushek and Kain [66] and the next section of this paper.

40 Multivariate regression analysis is designed to take into account correlations among the exogenous variables. If the exogenous variables are uncorrelated, bivariate regression (or simple correlations) will suffice. However, with very high levels of correlation among the independent variables, the coefficient estimates become imprecise; in the extreme, with perfect linear relationships among exogenous variables, estimation of the independent effects of the variables is simply impossible. Diagnosis is often difficult because multicollinearity often causes high estimated variances of the coefficient, but such high variance can also result from including variables which are unimportant in determining achievement. For further discussion and diagnostic aids, see Hanushek and Jackson [65].

5. *Statistical Methods.* Throughout the previous discussion, the focus has been on the estimation of the parameters of the production process. An alternative focus, and a close statistical relative, is contained in analysis of variance (see Coleman and others [37]). In this methodology, the observed achievement variance is decomposed into that arising from different sources. Suffice it to say, this methodology is often inappropriate for the questions under consideration (see Cain and Watts [28]), and simply raises further, added interpretive questions with no apparent gain.[41]

IV. AN ASSESSMENT OF WHAT WE KNOW AND WHAT WE SHOULD DO

The discussion to this point has indicated a wide range of problems—from conceptual problems to technical and esoteric interpretive issues. However, the overall analytical power of the production function framework, which integrates observations about various aspects of schools, should not be lost. Clearly, more detailed theoretical and empirical analyses focusing on specific aspects of the production process (such as the mechanisms of peer influences or the organizational and decision-making framework of schools) have been conducted outside of the context of production function analysis. However, they generally suffer from one of two problems: Either they concentrate exclusively upon a given attribute of schools or the learning process, or they consider the relationship between a particular attribute and student outcomes to the exclusion of other attributes that simultaneously affect outcomes. While these studies are useful in clarifying the important attributes of schools and in describing what goes on in and around schools, they are considerably less useful in considering alternative policies with respect to schooling. The strength of the production function studies lies in their policy relevance through investigation of the independent influences of various factors—student characteristics, teacher and school inputs, and other environmental attributes—on performance of the schooling system.

Further, this discussion should not be interpreted as implying that we have not learned anything from past research. In fact, there are some startlingly consistent findings that are quite robust to some of the problems mentioned, and many apparent inconsistencies are significantly reduced

41 The results of these techniques are sample specific; that is, they depend importantly upon the observed sample variations in the dependent and independent variables. Further, some of the variation, maybe a significant portion of it, can be accounted for jointly by the various independent variables (see Hanushek and Kain [66]), which leaves the choice of arbitrarily allocating the joint variation (as in Coleman and others [37]) or simply identifying its importance (as in Mayeske and others [99]). In neither case can one indicate the expected effect of changing given inputs.

when the most obvious specification, estimation, and measurement errors are taken into account. Not only are these problems not uniformly distributed across studies, but also it is often possible to determine the direction, if not the plausible magnitude, of many such biases.

In terms of consistent findings, differences in family socioeconomic background without question lead to significant achievement differences. Socioeconomic status is interpreted as a proxy for quality of the home learning environment. How much arises from factors malleable in the short run (current income and consumption, physical surroundings, current attitudes, etc.) and how much arises in the longer run, less malleable attributes (such as parental education, patterns of family rearing, etc.) is unknown, although longer run attributes are probably more important.[42]

Second, there is conclusive evidence that differences among schools and teachers are important in achievement. Schools simply do not have homogeneous impacts on students. Yet, the identification and measurement of specific teacher or school attributes which are important is much less certain. The variation across studies in specific characteristics that appear important in part reflects incomparabilities in underlying samples and data. For example, measures of "general intelligence" of teachers appear consistently important when considered (see Hanushek [61]), but are most frequently unavailable. Part of the variation undoubtedly also reflects teacher "skill" differences that are difficult to identify, measure, and model. Nevertheless, there is some evidence that school officials can identify more productive teachers. While the inability to disentangle the attributes of effective teachers indicates difficulty in selecting "good" teachers ex ante or in improving teacher productivity, the fact that good teachers can be identified ex post indicates that schools can be improved by appropriate promotion and allocation decisions.

Third, there is quite conclusive evidence that schools are economically inefficient; that is, they do not employ the best mixes of inputs, given input prices and their apparent effectiveness.[43] The possibility of inefficiency in

42 Analysis of this issue, which deserves added attention, requires longitudinal data on individuals, preferably where current family characteristics change significantly. (Data generated from the negative income tax experiments in Gary, Seattle, and Denver, for example, seem appropriate). Further, it may be even more useful to actually model behavior within households. Nevertheless, simply because a current background measure, say, family income, appears important, one cannot conclude that changing this will increase achievement in the short run since it may proxy other attributes (that aren't changed) and may be contaminated by individuals' ability effects (see above).

43 The evidence on economic efficiency comes from two, almost universal, findings of no consistent or significant relationship: (1) between achievement and expenditures per pupil (either instructional expenditure or total expenditures); and (2) between achievement and specific purchased inputs (teacher experience, teacher education levels, class size, and administrative/supervisory expenditures). Teacher experience appears somewhat produc-

school choices of organization and process, while less conclusive, does not however appear overwhelming.[44]

Fourth, there are significant differences in production functions by race and, perhaps, family background; that is, school resources interact importantly with the background characteristics of individuals.[45] On the research side, this implies exercising considerable care in modeling efforts, particularly when forced to use aggregate data. On the policy side, it indicates that more attention should be given to the internal allocation of resources. If inputs were equally effective for all students, shuffling teachers (without changing the pool) is a zero-sum game—winners balance against losers; this is not the case when input effectiveness varies across students.

In addition to these significant substantive results, there are, however, a number of significant gaps in our knowledge. Several of these have come out in the previous discussion and require only summarizing here.

The primary gap in understanding—at both a conceptual and empirical level—is an inadequate picture of the relationship between school quality and subsequent performance and therefore of how to measure school quality. The previous findings on the operations of the school system relate almost exclusively to test-score measures of achievement, even though validity of this measure is quite uncertain. It is difficult to overemphasize the importance of pursuing this line of inquiry.[46]

Beyond this, there are also a series of uncertainties which are amenable to research. For example, the influence of peer compositions—which is central to such important questions as integration and ability-tracking policies—remains murky.[47] Also, the dynamics of the educational process

---

tive, but clearly less productive at the margin than its cost based upon typical salary schedules. The others do not appear to show any positive relationships, let alone relationships strong enough to justify their costs.

44 Relatedly, little information about the decision process which dictates organizational and process determination is available. While not essential for understanding the production process per se, this might be very relevant for developing appropriate strategies to institute change. Available work on introduction and implementation of "innovations" (really process changes) includes Berman and McLaughlin [15] and Silkman [127].

45 Previous work indicates that stratification by race or SES or inclusion of interaction terms among school resources and backgrounds consistently shows significant achievement relationship differences: for example, Hanushek [61] or Summers and Wolfe [134].

46 While data availability is a clear problem, more can be done along these lines even with existing data. Considerable individual data with information about qualitative aspects of schooling are becoming available, but have not been fully exploited for linkages with education analyses. For example, Sewell and Hauser [126] have data about specific schools attended and lifetime outcomes, as does part of the Survey of Income and Education. The Michigan Panel Study of Income Dynamics and the National Longitudinal Surveys have a variety of test measures along with outcome data. Each could potentially provide more information about school quality and subsequent outcomes.

47 In many studies, data limitations (either because of aggregation problems or incomplete

are imperfectly understood. Most studies have been cross-sectional and rely upon specialized samples with little information about the impacts of resources for different age–school year cohorts. Therefore, given other incommensurates, it has been very difficult to analyze such issues as the differential impact of early school programs or varying achievement patterns.[48] As a final example, there has been little investigation of the stability of teacher skills, or individual teacher effects, over time.

Many of the analytical problems are ones where the conceptual problems are minimal, but data problems are severe.[49] In particular, the current methodology is probably quite adequate for many further analyses of elementary education where, among other things, there are fewer measurement problems and generally simpler organizational structures. There simply seem to be large, immediate payoffs to collecting new data for a variety of school situations where the most significant measurement questions indicated above are avoided and where longitudinal information can be obtained.[50]

Another set of problems—ones where conceptual inadequacies appear paramount—is nevertheless very significant. They have occurred in the previous discussion and include measures of alternative outputs, investigation of decision processes with regard to alternative output mixes, identification and measurement of process and organizational variables for both schools and classrooms, and expansion of our models for more complicated realities such as high schools.

V. CONCLUDING REMARKS

While the primary motivation for analyzing educational production relationships is derived from the sector's important resource usage, from its

---

information) preclude analysis of peer influences. The studies directly considering peer influences (Hanushek [61] and Henderson and others [70]) yield conflicting results. The research into racial composition has generally been flawed by limited historical information.

48 Several studies have included measures of preschool programs (e.g., Hanushek [61], Ritzen and Winkler [120]). These generally show more impact than is found in direct analyses such as that of Headstart (Cicirelli and others [35]).

49 One clear problem has been the use of specialized samples which relate to limited situations and which (because generally collected for other purposes) are not always appropriate. Historically, however, there seems to be a bias against data collection. Considerable money was spent on reanalyzing the data from the Coleman Report, even though it was well recognized that there were many, near-fatal flaws in those data, while none was spent on simply gathering better data.

50 A panel study design could, in addition to allowing dynamic analyses, be used to minimize remaining problems such as missing data on individual abilities; see Boardman and Murnane [19] or Hanushek and Jackson [65].

importance as a policy instrument, and from concerns about efficiency given the lack of market incentives, such analyses also have direct linkages to other research and policies. For example, the entire focus of school finance system discussions has been the distribution of expenditures per student under alternative financing mechanisms, even though the apparent levels of economic inefficiency in education indicate that this has little to do with the distribution of educational services. Similarly, many studies (found in public finance or urban economics) control for the quality of governmental services by including school expenditures per student with little consideration of what this is actually measuring. While part of school integration discussions relate to the effect of racial composition on achievement, ambiguities in this work have led to considerable confusion in this area (see Clune [36]). As a final example, even though years of schooling is a clearly inadequate measure of individual skill and ability differences, most contemporary labor economics and sociological research deals at only this level.

While other examples are easy to compile, the message should be clear: Understanding the educational sector has important ramifications for understanding many other areas, but the treatment of education has, for the most part, been quite superficial.

REFERENCES

1. Dennis J. Aigner and S. F. Chu. "On Estimating the Industry Production Function." *American Economic Review* 58 (September 1968).
2. John S. Akin and Irwin Garfinkel. "School Expenditures and the Economic Returns to Schooling." *Journal of Human Resources* 12 (Fall 1977): 460–81.
3. Karl L. Alexander and Edward L. McDill. "Selection and Allocation Within Schools: Some Causes and Consequences of Curriculum Placement." *American Sociological Review* (December 1976).
4. Duane Alwin. "College Effects on Educational and Occupational Attainments." *American Sociological Review* (April 1974).
5. David Armor. "The Evidence on Busing." *Public Interest* (Summer 1972).
6. David Armor and others. *Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools.* R-2007-LAUSD. Santa Monica, Calif.: Rand Corporation, 1976.
7. Kenneth J. Arrow. "Higher Education as a Filter." *Journal of Public Economics* 2 (1973).
8. Kenneth J. Arrow and others. "Capital-Labor Substitution and Economic Efficiency." *Review of Economics and Statistics* (August 1961).
9. Orley Ashenfelter. "Estimating the Effects of Training on Earnings." *Review of Economics and Statistics* (February 1978).
10. Orley Ashenfelter and Joseph D. Mooney. "Graduate Education, Ability, and Earnings." *Review of Economics and Statistics* (February 1968): 78–86.

11. Harvey A. Averch and others. *How Effective Is Schooling?* Englewood Cliffs, N.J.: Educational Technology Publications, 1974.
12. Burt S. Barnow and Glen G. Cain. "A Reanalysis of the Effect of Head Start on Cognitive Development: Methodology and Empirical Findings." *Journal of Human Resources* 12 (Spring 1977): 177–97.
13. Yoram Ben-Porath. "The Production of Human Capital Over Time." In *Education, Income, and Human Capital,* ed. W. Lee Hansen. New York: National Bureau of Economic Research, 1970.
14. Ivar E. Berg. *Education and Jobs: The Great Training Robbery.* New York: Praeger Publishers, 1970.
15. Paul Berman and Milbrey W. McLaughlin. *Federal Programs Supporting Educational Change, Vol. IV: The Findings in Review.* R-1589/4-HEW. Santa Monica, Calif.: Rand Corporation, 1975.
16. Matthew Black. "The Determinants and Implications of Job Search and Quit Activity." Ph.D. dissertation, University of Michigan, 1976.
17. Peter Blau and Otis D. Duncan. *The American Occupational Structure.* New York: John Wiley & Sons, 1967.
18. Anthony E. Boardman, Otto A. Davis, and Peggy R. Sanday. "A Simultaneous Equations Model of the Educational Process." Pittsburgh: School of Urban and Public Affairs, Carnegie-Mellon University, 1973, mimeo.
19. Anthony E. Boardman and Richard J. Murnane. "The Use of Panel Data in Education Research." Fels Discussion Paper No. 110. Philadelphia: University of Pennsylvania, 1977.
20. Raymond Boudon. *Education, Opportunity, and Social Inequality.* New York: John Wiley & Sons, 1974.
21. Samuel Bowles. "Towards an Education Production Function." In *Education, Income, and Human Capital,* ed. W. Lee Hansen. New York: National Bureau of Economic Research, 1970.
22. Samuel Bowles and Herbert Gintis. *Schooling in Capitalist America.* New York: Basic Books, 1976.
23. Samuel Bowles and Henry M. Levin. "The Determinants of Scholastic Achievement—A Critical Appraisal of Some Recent Evidence." *Journal of Human Resources* 3 (Winter 1968): 3–24.
24. ————. "More on Multicollinearity and the Effectiveness of Schools." *Journal of Human Resources* 3 (Summer 1968): 393–400.
25. Byron W. Brown and Daniel H. Saks. "The Production and Distribution of Cognitive Skills Within Schools." *Journal of Political Economy* 83 (April 1975).
26. Jesse Burkhead, with Thomas G. Fox and John W. Holland. *Input and Output in Large-City Schools.* Syracuse, N.Y.: Syracuse University Press, 1967.
27. Glen G. Cain and Robinson G. Hollister. "The Methodology for Evaluating Social Action Programs." In *An Anthology of Labor Economics: Readings and Commentary,* eds. Ray Marshall and Richard Perlman. New York: John Wiley & Sons, 1972.
28. Glen G. Cain and Harold W. Watts. "Problems in Making Policy Inferences from the Coleman Report." *American Sociological Review* (April 1970).
29. Angus Campbell and others. *The American Voter.* Chicago: University of Chicago Press, 1976.

30. Daryl Carlson. "Examining Efficient Joint Production Processes." In *Measuring and Increasing Academic Productivity,* ed. Robert A. Wallhaus. New Directions for Institutional Research No. 8, Winter 1975.
31. Polly Carpenter and George R. Hall. *Case Studies in Educational Performance Contracting: Conclusions and Implications.* R-900/1-HEW. Santa Monica, Calif.: Rand Corporation, 1971.
32. Central Advisory Council for Education (England). *Children and Their Primary Schools.* London: Her Majesty's Stationery Office, 1967.
33. Barry R. Chiswick. "Schools and Earnings of Low Achievers: Comment." *American Economic Review* 62 (September 1972): 752–54.
34. Laurits R. Christensen, Dale W. Jorgenson, and Laurence J. Lau. "Transcendental Logarithmic Production Frontiers." *Review of Economics and Statistics* (February 1973).
35. Victor G. Cicirelli and others. *The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development.* Westinghouse Learning Corporation, 1969.
36. William H. Clune. "Law and Economics in *Hobson* v. *Hansen*—An Introductory Note." *Journal of Human Resources* 7 (Summer 1972): 275–82.
37. James S. Coleman and others. *Equality of Educational Opportunity.* Washington: U.S. Government Printing Office, 1966.
38. James S. Coleman. "The Evaluation of 'Equality of Educational Opportunity.'" In *On Equality of Educational Opportunity,* eds. F. Mosteller and D. P. Moynihan. New York: Random House, 1972.
39. James S. Coleman and Nancy L. Karweit. *Measures of School Performance.* R-488-RC. Santa Monica, Calif.: Rand Corporation, 1970.
40. Lee J. Cronbach and Lita Furby. "How Should We Measure 'Change'—or Should We?" *Psychology Bulletin* (1970).
41. Edward F. Denison. *Accounting for United States Economic Growth 1929–69.* Washington: Brookings Institution, 1974.
42. Dennis J. Dugan. "Scholastic Achievement: Its Determinants and Effects in the Education Industry." In *Education as an Industry,* eds. Joseph T. Froomkin, Dean T. Jamison, and Roy Radner. Cambridge, Mass.: Ballinger, 1976.
43. Greg J. Duncan. "Nonpecuniary Work Rewards." In *Five Thousand American Families: Patterns of Economic Progess,* Vol. II, ed. James N. Morgan. Ann Arbor: Institute for Social Research, University of Michigan, 1974.
44. Otis D. Duncan, David Featherman, and Beverly Duncan. *Socioeconomic Background and Achievement.* New York: Seminar Press, 1972.
45. Rupert N. Evans and Joel D. Galloway. "Verbal Ability and Socioeconomic Status of 9th and 12th Grade College Preparatory, General, and Vocational Students." *Journal of Human Resources* 8 (Winter 1973): 24–36.
46. Franklin M. Fisher. *The Identification Problem in Econometrics.* New York: McGraw-Hill, 1966.
47. Manley Fleischmann and others. *Report of the New York State Commission on the Quality, Cost, and Financing of Elementary and Secondary Education.* 1978.

48. David J. Fox. "Expansion of the More Effective School Program." New York: Center for Urban Education, 1967, mimeo.
49. Irwin Garfinkel and Edward M. Gramlich. "A Statistical Analysis of the OEO Experiment in Educational Performance Contracting." *Journal of Human Resources* 8 (Summer 1973): 275–305.
50. William T. Garner. "The Identification of an Education Production Function by Experimental Means." Ph.D. dissertation, University of Chicago, 1973.
51. Herbert J. Gintis. "Education, Technology, and the Characteristics of Worker Productivity." *American Economic Review* 61 (May 1971).
52. Edward M. Gramlich and Patricia P. Koshel. *Educational Performance Contracting.* Washington: Brookings Institution, 1975.
53. David Greenberg and John McCall. "Teacher Mobility and Allocation." *Journal of Human Resources* 9 (Fall 1974): 480–502.
54. Zvi Griliches. "Production Functions in Manufacturing: Some Preliminary Results." In *The Theory and Empirical Analysis of Production,* ed. Murray Brown. New York: National Bureau of Economic Research, 1967.
55. ————. "Estimating the Returns to Schooling: Some Econometric Problems." *Econometrica* (January 1977).
56. Zvi Griliches and William Mason. "Education, Income, and Ability." *Journal of Political Economy* 80 (May/June 1972, Pt. 2).
57. Michael Grossman. *The Demand for Health: A Theoretical and Empirical Analysis.* National Bureau of Economic Research Occasional Paper 119. New York: Columbia University Press, 1972.
58. James W. Guthrie and others. *Schools and Inequality.* Cambridge, Mass.: MIT Press, 1971.
59. Giora Hanoch. "Production and Demand Models with Direct or Indirect Implicit Additivity." *Econometrica* (May 1975).
60. W. Lee Hansen, Burton Weisbrod, and William J. Scanlon. "Schooling and Earnings of Low Achievers." *American Economic Review* 60 (June 1970).
61. Eric A. Hanushek. *Education and Race.* Lexington, Mass.: D.C. Heath, 1972.
62. ————. "Regional Differences in the Structure of Earnings." *Review of Economics and Statistics* (May 1973).
63. ————. "Comment." In *Education as an Industry,* eds. Joseph T. Froomkin and others. Cambridge, Mass.: Ballinger, 1976.
64. ————. "Ethnic Income Variations: Magnitudes and Explanations." In *American Ethnic Groups,* ed. Thomas Sowell. Washington: Urban Institute, 1978.
65. Eric A. Hanushek and John E. Jackson. *Statistical Methods for Social Scientists.* New York: Academic Press, 1977.
66. Eric A. Hanushek and John F. Kain. "On the Value of 'Equality of Educational Opportunity' as a Guide to Public Policy." In *On Equality of Educational Opportunity,* eds. F. Mosteller and D. P. Moynihan. New York: Random House, 1972.
67. Eric A. Hanushek and John M. Quigley. "Implicit Investment Profiles and Intertemporal Adjustments of Relative Wages." *American Economic Review* 68 (March 1978).

68. Donald P. Hayes and Judith Grether. "The School Year and Vacations: When Do Students Learn?" Washington: Urban Institute, 1969.
69. John Heim and Lewis Perl. *The Educational Production Frontier.* Ithaca, N.Y.: Cornell University, 1974.
70. Vernon Henderson, Peter Mieszkowski, and Yvon Sauvageau. *Peer Group Effects and Educational Production Functions.* Ottawa: Economic Council of Canada, 1976.
71. George H. Hildebrand and T. C. Liu. *Manufacturing Production Functions in the U.S., 1957.* Ithaca, N.Y.: Cornell University Studies in Industrial and Labor Relations, 1965.
72. Russell Hill and Frank Stafford. "Allocation of Time to Preschool Children and Educational Opportunity." *Journal of Human Resources* 9 (Summer 1974): 323–41.
73. Robert P. Inman. "The Fiscal Reform of Metropolitan Schools: Some Simulation Results." *American Economic Review* 68 (March 1978).
74. Dean T. Jamison and others. "Cost and Performance of Computer-Assisted Instruction for Education of Disadvantaged Students." In *Education as an Industry,* eds. Joseph T. Froomkin and others. Cambridge, Mass.: Ballinger, 1976.
75. Christopher S. Jencks. "The Quality of Data Collected by the 'Equality of Educational Opportunity' Survey." In *On Equality of Educational Opportunity,* eds. F. Mosteller and D. P. Moynihan. New York: Random House, 1972.
76. Christopher S. Jencks and others. *Inequality: A Reassessment of the Effects of Family and Schooling in America.* New York: Basic Books, 1972.
77. Christopher S. Jencks and Marsha D. Brown. "Effects of High Schools on Their Students." *Harvard Educational Review* (August 1975).
78. George E. Johnson and Frank P. Stafford. "Social Returns to Quantity and Quality of Schooling." *Journal of Human Resources* 8 (Spring 1973): 139–55.
79. Nancy Karweit. "The Organization of Time in Schools: Time Scales and Learning." Baltimore: Johns Hopkins University, 1978, mimeo.
80. Martin Katzman. *The Political Economy of Urban Schools.* Cambridge, Mass.: Harvard University Press, 1971.
81. Stuart Kerachsky and Charles D. Mallar. "Design of an Evaluation of the Job Corps." Working Paper No. C-12. Princeton, N.J.: Mathematica Policy Research, February 1977.
82. Nicholas Kiefer. "Federally Subsidized Occupational Training and the Employment and Earnings of Male Trainees." *Journal of Econometrics* (forthcoming).
83. Herbert J. Kiesling. "Measuring a Local Government Service: A Study of School Districts in New York State." *Review of Economics and Statistics* (August 1967).
84. ————. "A Study of the Relationship of Instruction Process and Program Organization to the Success of Compensatory Education Projects in California." In *Education as an Industry,* eds. Joseph T. Froomkin and others. Cambridge, Mass.: Ballinger, 1976.
85. Robert E. Klitgaard and George R. Hall. "Are There Unusually Effective Schools?" *Journal of Human Resources* 10 (Winter 1975): 90–106.

86. Richard Layard and George Psacharopoulos. "The Screening Hypothesis and the Returns to Education." *Journal of Political Economy* 82 (September/ October 1974).
87. A. James Lee. "The Economics of Compulsory Schooling—A Theoretic and Empirical Investigation." Ph.D. dissertation, University of Wisconsin, 1976.
88. Harvey Leibenstein. "Allocative Efficiency vs. 'X-Efficiency.'" *American Economic Review* 56 (June 1966).
89. Henry M. Levin. "A New Model of School Effectiveness." In *Do Teachers Make a Difference?* U.S. Office of Education. Washington: U.S. Department of Health, Education, and Welfare, 1970.
90. ————, ed. *Community Control of Schools.* Washington: Brookings Institution, 1970.
91. ————. "Concepts of Economic Efficiency and Educational Production." In *Education as an Industry,* eds. Joseph T. Froomkin and others. Cambridge, Mass.: Ballinger, 1976.
92. Frank S. Levy. "The Racial Imbalance Act of Massachusetts." Ph.D. dissertation, Yale University, 1969.
93. Arleen Leibowitz. "Home Investments in Children." *Journal of Political Economy* 82 (March/April 1974, Pt. II).
94. Peter Lindert. "Family Inputs and Inequality Among Children." Discussion Paper No. 218-74. Madison: Institute for Research on Poverty, University of Wisconsin, 1974.
95. Charles R. Link and Edward C. Ratledge. "Social Returns to Quantity and Quality of Education: A Further Statement." *Journal of Human Resources* 10 (Winter 1975): 78–89.
96. ————. "Student Perceptions, I.Q., and Achievement." *Journal of Human Resources* 14 (Winter 1979): 98–111.
97. Larry M. Manheim. "Health, Health Practices, and Socioeconomic Status: The Role of Education." Ph.D. dissertation, University of California, Berkeley, 1975.
98. Stanley Masters and Thomas Ribich. "Schooling and Low Achievers: Comment." *American Economic Review* 62 (September 1972): 755–59.
99. George Mayeske and others. *A Study of the Achievement of Our Nation's Students.* Washington: U.S. Department of Health, Education, and Welfare, 1973.
100. Rebecca Maynard and David L. Crawford. "School Performance." In *Rural Income Maintenance Experiment: Final Report,* Vol. VI, Ch. 12. Madison: Institute for Research on Poverty, University of Wisconsin, 1976.
101. Stephan Michelson. "The Association of Teacher Resources with Children's Characteristics." In *Do Teachers Make a Difference?* U.S. Office of Education. Washington: U.S. Department of Health, Education, and Welfare, 1970.
102. ————. "For the Plaintiffs—Equal School Resource Allocation." *Journal of Human Resources* 7 (Summer 1972): 283–306.
103. Jacob Mincer. "The Distribution of Labor Income: A Survey with Special Reference to the Human Capital Approach." *Journal of Economic Literature* 8 (March 1970).
104. Richard D. Morgenstern. "Direct and Indirect Effects on Earnings of

Schooling and Socio-Economic Background." *Review of Economics and Statistics* 55 (May 1973): 225–33.

105. Frederick Mosteller and Daniel P. Moynihan, eds. *On Equality of Educational Opportunity.* New York: Random House, 1972.
106. Richard Murnane. *The Impact of School Resources on the Learning of Inner City Children.* Cambridge, Mass.: Ballinger, 1975.
107. ———. "Labor Mobility in a Declining Industry: The Case of Teachers." Paper presented at Econometric Society meetings, August 1978.
108. Richard R. Nelson and Edmund Phelps. "Investment in Humans, Technology, Diffusion, and Economic Growth." *American Economic Review* 56 (May 1966).
109. Richard R. Nelson and Sidney G. Winter. "Neoclassical vs. Evolutionary Theories of Economic Growth: Critique and Prospectus." *Economic Journal* (December 1974).
110. ———. "Factor Price Changes and Factor Substitution in an Evolutionary Model." *Bell Journal of Economics* (Autumn 1975).
111. ———. "Organizational Capability in a Dynamic World: A Behaviorist Reconstruction of the Production Set Idea." New Haven, Conn.: Yale University, October 1976, mimeo.
112. Richard Niemi and Barbara I. Sobieszek. "Political Socialization." *Annual Review of Sociology* (1977).
113. Dave M. O'Neill, Burton Gray, and Stanley Horowitz. "For the Defendants —Educational Equality and Expenditure Equalization Orders." *Journal of Human Resources* 7 (Summer 1972): 307–25.
114. John D. Owen. "The Distribution of Educational Resources in Large American Cities." *Journal of Human Resources* 7 (Winter 1972): 26–38.
115. Lewis J. Perl. "Family Background, Secondary School Expenditure, and Student Ability." *Journal of Human Resources* 8 (Spring 1973): 156–80.
116. President's Commission on School Finance. *Schools, People, and Money.* Washington: U.S. Government Printing Office, 1972.
117. Thomas I. Ribich and James L. Murphy. "The Economic Returns to Increased Educational Spending." *Journal of Human Resources* 10 (Winter 1975): 56–77.
118. John Riew. "Economies of Scale in High School Operations." *Review of Economics and Statistics* (August 1966).
119. John G. Riley. "Competitive Signalling." *Journal of Economic Theory* 10 (1975).
120. Jozef M. Ritzen and Donald R. Winkler. "The Production of Human Capital Over Time." *Review of Economics and Statistics* (November 1977).
121. Alice Rivlin and P. Michael Timpane, eds. *Planned Variation in Education.* Washington: Brookings Institution, 1975.
122. Daniel P. Rogers. "Private Rates of Return to Education in the United States: A Case Study." *Yale Economic Essays* (Spring 1969).
123. Sherwin Rosen. "Human Capital: A Survey of Empirical Research." In *Research in Labor Economics,* Vol. I, ed. Ronald G. Ehrenberg. Greenwich, Conn.: JAI Press, 1977.
124. James E. Rosenbaum. *Making Inequality.* New York: John Wiley & Sons, 1976.
125. Sandra Scarr and Richard A. Weinberg. "The Influence of 'Family Background' on Intellectual Attainment: The Unique Contribution of Adoptive

Studies for Estimating Environmental Effects." New Haven, Conn.: Yale University, September 1977, mimeo.

126. William Sewell and Robert Hauser. *Education, Occupation, and Earnings.* New York: Academic Press, 1975.
127. Richard Silkman. "The Diffusion of Educational Innovation." Ph.D. dissertation, Yale University, in progress.
128. Marshall Smith. "Equality of Educational Opportunity: The Basic Findings Reconsidered." In *On Equality of Educational Opportunity,* eds. F. Mosteller and D. P. Moynihan. New York: Random House, 1972.
129. Lewis Solmon. "The Definition and Impact of College Quality." In *Explorations in Economic Research.* New York: National Bureau of Economic Research, Fall 1975.
130. Michael Spence. "Job Market Signalling." *Quarterly Journal of Economics* (August 1973).
131. Joseph E. Stiglitz. "The Theory of Screening, Education, and the Distribution of Income." *American Economic Review* 65 (June 1975).
132. Susan B. Stodolsky and Gerald Lesser. "Learnings Patterns of the Disadvantaged." *Harvard Educational Review* (Fall 1967).
133. Anita Summers and Barbara Wolfe. "Equality of Educational Opportunity Quantified: A Production Function Approach." *Philadelphia Federal Reserve Papers,* 1974.
134. ———. "Do Schools Make a Difference?" *American Economic Review* 67 (September 1977).
135. P. A. V. B. Swamy. "Criteria, Constraints, and Multicollinearity in Random Coefficient Models." *Annals of Economic and Social Measurement* (October 1973).
136. Paul Taubman and Terence Wales. *Higher Education and Earnings.* New York: McGraw-Hill, 1974.
137. Michael K. Taussig. "An Economic Analysis of Vocational Education in the New York City High Schools." *Journal of Human Resources* 3 Supp. (1968): 59–87.
138. Lester Thurow. *Generating Inequality.* New York: Basic Books, 1975.
139. Lester Thurow and R. E. B. Lucas. *American Distribution of Income: A Structural Problem.* Joint Economic Committee, 92nd Congress, March 1972.
140. U.S. Commission on Civil Rights. *Racial Isolation in the Public Schools,* Vols. I and II. Washington: U.S. Government Printing Office, 1967.
141. U.S. Senate. *Hearings Before the Select Committee on Equal Educational Opportunity,* Pts. 1–21. 92nd Congress, 1971.
142. Paul Wachtel. "The Effect on Earnings of School and College Investment Expenditures." *Review of Economics and Statistics* (August 1976).
143. Burton A. Weisbrod and Peter Karpoff. "Monetary Returns to College Education, Student Ability, and College Quality." *Review of Economics and Statistics* (1968).
144. Randall D. Weiss. "The Effect of Education on the Earnings of Blacks and Whites." *Review of Economics and Statistics* (May 1970).
145. Finis Welch. "Measurement of the Quality of Schooling." *American Economic Review* 56 (May 1966).

146. ———. "Education in Production." *Journal of Political Economy* 78 (January/February 1970).
147. ———. "Black-White Differences in the Returns to Schooling." *American Economic Review* 63 (December 1973).
148. David Wiley and Annegret Harnischfeger. "Explosion of a Myth: Quantity of Schooling and Exposure to Instruction, Major Educational Vehicles." *Education Researcher* (March 1974).
149. Donald R. Winkler. "Educational Achievement and School Peer Composition." *Journal of Human Resources* 10 (Spring 1975): 189–204.
150. ———. *The Production of Human Capital.* New York: Arno Press, 1977.
151. Willard Wirtz and others. *On Further Examination: Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline.* New York: College Entrance Examination Board, 1977.
152. Kenneth I. Wolpin. "Education and Screening." *American Economic Review* 67 (December 1977): 949–58.