From Fiscal Year 1965 Through 1970." Palo Alto, Calif.: American Institutes for Research, March 1972.
Weiler, Daniel, et al. *A Public School Voucher Demonstration: The First Year at Alum Rock*. Santa Monica, Calif.: Rand Corporation, June 1974.
Williams, Andrew T. "Education and the Joint Product Production Function." Paper prepared for the Seminar in Economics of Education, Stanford University, 1970, processed.
Winkler, Donald R. *The Production of Human Capital: A Study of Minority Achievement*. Ph.D. diss., Department of Economics, University of California at Berkeley, 1972.
Wynne, Edward. *The Politics of School Accountability*. Berkeley, Calif.: McCutchan Publishing Co., 1972.

4 COMMENTS

Eric A. Hanushek
Yale University

It always helps to take stock periodically of what we are doing and where we are going. Educational research is certainly no exception. It is simply too easy to continue doing what we know best. However, after being forced by Levin into a reappraisal of the directions of educational research, I remain unpersuaded that drastic modification is called for. I have some serious reservations about the conclusions and implications both for research and public policy drawn from Levin's analysis, particularly as it pertains to technical inefficiency in the schools.

In many ways, Levin's viewpoint does not diverge significantly from my own. We start from the same data; there is no disagreement about the constraints on the system or about the amount of knowledge and information available to the participants in the educational process. We also agree on many of the results that develop from constraints on information and possible actions. Our major points of disagreement arise from nomenclature of observed output differences and the subsequent implications for future research and public policy.

In his taxonomy of types of inefficiency, Levin argues that there are reasons to believe that schools are not operating on the production frontier (technical inefficiency), are not operating with the best input mix (allocative inefficiency), and are not providing the desired output mix (social welfare inefficiency). The heart of his analysis is directed toward the evidence concerning technical inefficiency and its implications for research and public policy. That is also the central issue in my discussion.

Before entering that debate, however, I wish to make two points relating to allocative and social welfare efficiency. These are not points of disagreement with Levin; they are simply added to emphasize certain aspects of the discussion. First, allocate inefficiency has not only been the central concern of economists but is also almost a necessary condition for analysis. In the absence of large differences in the relative prices of inputs, allocative inefficiency is needed to analyze educational production functions. Otherwise we would observe one point on the production function, and our statistical techniques are noticeably weak at drawing multidimensional planes through one point. Second, the whole issue of social welfare efficiency, or producing the best mixture of outputs, has the same elusive character as choosing the right quantity and mixture of general public goods. The optimum marginal conditions on the social welfare function are easy enough to derive, but the operational questions have generally been beyond the economist's ability to answer. Nevertheless, in my subjective evaluation, this is probably the most important area of concern in education today. The question of whether or not schools are producing the outputs desired and needed in society remains important but unresolved. We are not sure what the outputs of schools are, how to measure the outputs, how to produce each, or what tradeoffs exist among outputs. Not only space limitations on this discussion but also the difficulty of the issue preclude my going into more detail on this.

The main message delivered by Levin is that there are compelling reasons to believe that what have paraded under the banner of educational production functions are not really production functions in the economist's usage of the term, because they do not describe the frontier of possible production. Instead they are a weighted average of the practices of efficient, or "frontier," schools and inefficient, or "nonfrontier," schools. With these average relationships, blind application of well-known optimization rules could even degrade the production by a school system which is almost universally cited as inefficient.

A crucial facet of the debate is how one should define technical inefficiency. Past discussions, for example, Leibenstein's development of X-efficiency [2] and Levin's presentation, rest heavily upon a microeconomic textbook treatment of production, where output is a function of a quantity of homogeneous capital and homogeneous labor. Then, noting that these inputs are really not homogeneous, firms with poorer "homogeneous" inputs are observed to produce less output than firms with identical quantities but better quality "homogeneous" inputs. The availability of better or worse "homogeneous" inputs can be related to incomplete labor contracts, lack of knowledge of the production function, motivational differences, or simply general managerial ability. Differences in output for "equal" inputs are used as a measure of technical inefficiency.

A real problem remains in specifically defining technical inefficiency. In reality, technical inefficiency or X-inefficiency is a measure of the strength of variables omitted from a model of the production process. These omitted variables may take the form of education, motivation, laziness, or what have

you. They are just the explanations given for efficiency differences. In these terms, defining technical inefficiency becomes very difficult. Before any analysis of technical inefficiency can be developed, one must define what the base model of the production process should look like. Amounts of inefficiency then become difficult to measure, since they are a function of the chosen degree of misspecification in the base model. A common way to set the frame of reference appears to be using the model which can be developed from available data. This, of course, creates problems, because the amount of inefficiency can change over time simply due to better data becoming available.

A different way of looking at this "inefficiency," however, is to use a well-specified model as the standard and to view observed production differences in terms of model misspecification. The traditional production function that pictures output as a function of the quantity of man-hours or man-hours within given human capital classifications provides an incomplete view of the labor input to production. There are more attributes to labor than are represented in these functions. These omitted attributes often tend to be correlated with management ability or firm size or nationality in Leibenstein's international examples. Estimated production functions can then give a distorted view of the production potential. There is, however, no reason why the analyst cannot specify or attempt to specify all of the attributes that go into production. He need not be bound to specifying just those inputs as seen in a microeconomic text or those explicitly purchased by the firm.

In point of fact, this extension of the list of inputs has been the order of the day in educational research. Educational production functions of the past decade have not looked at schools as providing a given number of homogeneous teachers; nor have they looked at schools as providing only the set of purchased inputs (class size, experience, and graduate education). Instead they have looked at schools as providing a set of attributes, such as teacher verbal ability. The attributes explicitly measured may well be proxies for other attributes which have direct causal relationships with achievement. However, to the extent that a set of stable proxies which represent a fair proportion of the real teacher inputs to education have been analyzed, the importance of the technical inefficiency argument is considerably diminished.

If we map achievement outputs against only those inputs explicitly purchased by schools (class size, teacher experience, and teacher graduate education), we will certainly find the picture indicated by Levin's Figure 1. This will happen because, according to past analyses, the purchased factors have a small or nonexistent effect on output, but other nonpurchased characteristics of teachers do have an important effect. Since these other factors are not randomly distributed by schools—as shown in Levin [3], schools with apparently the same input levels will show different outputs. Yet, within the context of educational production functions, the real question is: Do schools have different outputs after the relevant teacher inputs are held constant?

It is reasonable that past discussions in fields other than education have

centered upon technical inefficiency in production. This arises largely from having poorer data sources for, say, aggregate manufacturing firms than for educational firms. Research in education has been aided by having detailed measures of relevant inputs. Further, the emphasis within educational research has been on refining the measures of inputs. This is not to say that we now have perfectly specified models of educational production. We have a long way to go in that regard. It does imply that attention has been placed where I think it properly should be—on model specification and, to a certain extent, on experimental design.

The case by Levin for technical inefficiency derives chiefly from the observations that school managers do not know what the production function for education looks like and that these managers are severely constrained in their operating and hiring practices. Other factors relating to technical inefficiencies are the general lack of competition in education and lack of both incentives and clear-cut signals of success or failure.

From a specification point of view, the implications to be drawn from Levin's observations of current school operations change considerably. First, I am uncertain how the school principal, whether he knows the production function or not, affects technical efficiency. If, as past research would suggest, the main school inputs to education under the current technology are attributes of the teacher, it is hard to see how the principal affects the relationship between these attributes and achievement by very much. In terms of managing teachers, the principal may assign his best reading teacher to teach physical education; this is an allocatively poor decision that would reduce total achievement in a school for his expenditures, but not necessarily one that falls off the production frontier for education. It indicates that the analyst must be careful to separate the characteristics of the physical education teacher and the reading teacher. But, given this, there seems to be no reason to require the principal to know that he is making a mistake.

The fact that there are constraints on the manager's actions does not seem to destroy the usefulness of estimated production functions either. Constraints imply that he can only operate on a limited portion of possible input mixes. For example, a principal probably does not have the option to install a Computer-Assisted Instruction (CAI) program on his own. Nevertheless, he can attempt to suboptimize within the portion of the production frontier available to him. There is no reason to suspect that any such suboptimization attempts lead to technical inefficiency.

The other conceptual reasons for concluding that technical inefficiency is probably large produce a similar discussion. Such reasons seem to imply that schools could be allocatively very inefficient but not technically inefficient.

There is an empirical question about the importance of variables relating to facilities, curriculum, and management which may be systematically related to achievement and not generally included in production models. I have made a modest attempt to answer this question with a sample of 515 students from blue-collar families within one school system. The data sam-

ple and estimated educational models are reported elsewhere [1]. After standardizing for different teacher inputs, I attempted to find out whether there were characteristics of schools and principals which systematically affected output. For this analysis, each of twenty-three schools in the sample was allowed to have its own intercept value, and statistical tests were performed to ascertain whether these intercepts differed by school. The intercept dummy variables provide estimates of the systematic school effects, regardless of whether the components of these effects can be adequately specified or measured. These effects would be equivalent to a measure of technical efficiency.

Within this sample, only one school out of twenty-three (comprising two per cent of the students) produced significantly higher outputs after standardizing for teacher inputs.[1] This appears to be very weak evidence for the existence of important technical inefficiencies. Matched against this is the finding that the total wage bill could be reduced by approximately 22 per cent with no decrease in achievement by not hiring individuals possessing superior experience or graduate education, or by not paying for such experience and graduate education, which were shown to have no impact on achievement. (In other words, by improving allocative efficiency, a savings of 22 per cent could be realized.)

Finally, we know that there is a large random component associated with individual achievement. There is no reason to suspect that we get more or better information about educational production by looking at a smaller sample, whether by linear programming or least squares. Also, even in the context of viewing "efficient" production with linear programming, there is no reason to believe that specification problems are any less severe. If we wish to make decisions about educational production from considering "efficient" schools, we are still left with trying to decide why such schools are efficient. In other words, we are left with the same specification problems.

CONCLUSIONS

It is not evident to me that technical inefficiency is a particularly large problem, unless we use obviously misspecified models as the standard. Within the context of well-specified models, similar to those developed within the past few years, emphasis upon allocative efficiency appears warranted. I do not wish to indicate that we know all there is to know about educational production. Yet, both conceptually and empirically, allocative inefficiency seems more important than technical inefficiency.

The difference in my approach and Levin's is more than a question of semantics. First, use of the term inefficiency tends to imply that there is a free lunch, that some organizational changes within the school will bring about significant changes in outputs at little or no cost. On the other hand, when viewed in terms of omitted variables, it is immediately obvious that bringing "inefficient" schools up to the level of "efficient" schools may not

be free. Second, the term technical inefficiency seems to imply that the observed differences in outputs are related almost exclusively to management differences. However, my work has led me to suspect that the real efforts should be directed toward better specifying teachers and their inputs to education. Third, the concept of technical efficiency appears vacuous from a public policy viewpoint. Even if some consensus could be arrived at as to how this inefficiency should be measured, we are at best led to trying to explain these differences in order to reduce the differentials involved.

If the problem is looked upon as one of specification problems, it leads to intensifying data collection efforts and broadening the scope of our measurement of teacher attributes. It also calls for experimentation in order to observe other parts of the production frontier. If instead, one concludes that school management in terms of approaching the production frontier is the key issue, a different course of action is called for. In this case, much more effort should be directed toward analyzing organizational behavior and the relationship between management, teachers, and facilities. In my judgment, the former course of action will have much higher payoffs.

On the other hand, Levin's observations about the definition and measurement of educational outputs cannot be disregarded. Even though cognitive ability, as measured by test scores, is undoubtedly an important aspect of elementary and secondary schools, this is not the sole output of schools. While the joint product problem is not completely developed by Levin, it represents a very important issue for future research. Unfortunately, the methodology for handling joint production when there are no prices (or weights) to combine the different dimensions of output is an underdeveloped area of economics.

NOTE

1. Another significant aspect of this estimation was the finding that the dummy variable for this school had a very low correlation with each of the included school variables. (The simple correlation was always less than .1.) This implies that even if we were to believe that the dummy variable represented some omitted management aspects for this school, its effect on the included coefficient estimates is small; that is, the amount of specification bias would be small.

REFERENCES

1. Hanushek, Eric A. *Education and Race*, Lexington, Mass.: D.C. Heath, 1972.
2. Leibenstein, Harvey. "Allocative Efficiency vs. 'X-Efficiency'," *American Economic Review* 56 (June 1966): 392–415.
3. Levin, Henry M. "Recruiting Teachers for Large-City Schools." Washington, D.C.: Brookings Institution, 1968, processed.

Harold W. Watts
University of Wisconsin

The primary conclusion Levin arrives at in this paper is that valid prescriptions for improving efficiency of the educational process (or that part of it which takes place in public schools) cannot be derived from existing estimated production functions. He argues that the world is much more complicated than the available econometric models of educational production and that naive attempts to draw normative conclusions from such models could be counter-productive. These conclusions seem appropriate enough as warnings or expressions of humility, but they also seem quite anticlimactic at the end of so many pages of analytic threshing about.

Levin opens by attempting to motivate our interest in efficiency by observing the rising costs of education combined with dissatisfaction over the quality of schools and loss of voter support for bond issues. It is quite clear that improvements in efficiency, if any are available, would offset for a while some of the increase in costs and might produce a more popular product. But the existence of inefficiency, in any of its varieties, does not imply either rising costs or consumer dissatisfaction. If inefficiency were getting worse at some steady rate, we might expect the consequences of rising costs and/or falling quality, but Levin does not provide evidence of progressive inefficiency. In no sense does the hypothesis of inefficiency provide an alternative to the "pessimistic" view of Baumol that "unprogressive" sectors will suffer cost increases as other sectors enjoy productivity gains and consequent wage increases.

But there is probably plenty of interest in efficiency as a property of the educational system and further motivation is unnecessary. Levin proceeds to use production isoquants and output transformation loci to illustrate various kinds of inefficiency and also to introduce various ways that model misspecification can foul up econometric estimates of production functions. Here one principal point is that productive units that are not using efficient techniques will not lie on the production frontier and will result in the estimation of a subfrontier production function. A second one is that output measurements may be incorrectly specified (either in one or many dimensions) and that spurious inefficiencies or "second bests" may be perceived as a consequence of that misspecification.

The next section of Levin's paper presents a long a priori argument in support of the proposition that schools must be inefficient! The main premise seems to be that they are unlike private competitive firms in a number of critical respects, and without those characteristics educational "firms" have no basis for achieving efficiency. I find myself quite convinced that the education "industry" is poles apart from the straw-man industry which has all the perfect properties of the competitive model. However, it seems that most of the real world of productive enterprises share enough of those imperfections to invalidate for them the simpleminded empirical analysis that Levin criticizes for education. Again, I can easily accept the argument that our public schools leave something to be desired in terms of efficiency, but I

cannot agree that the contrast with other private or public production is particularly unfavorable.

The other part of the section spells out the limitations of standardized achievement tests as measures of school output. I fully share all Levin's reservations here and would welcome any progress toward satisfactory measures of neglected aspects, but again, this is a problem of measuring the outputs of a human-service industry and that is an unsolved problem everywhere.

The next section explores the GIGO[1] production function as applied to econometric research and derived policy prescriptions in education. Levin is quite persuasive about the various kinds of mischief that can result from a zealous application of intermediate theory to the estimated production functions for education which have appeared in respectable journals. He is motivated in this analysis by a belief that there is a real danger of these half-baked conclusions being promulgated by ukase, and even worse, that they will affect school practice.

My own appreciation of how hard it is to get any real change in the way schools and teachers behave, combined with Levin's own sense about how varied schools are, both on and off the efficiency frontier, make the threat of lockstep imitation of the latest econometric optimality formula pretty remote. Consequently, I can accept his analysis of what-if-everyone-acted-silly without agreeing on the likelihood of the premise.

In the end, Levin pleads for better models—always a popular plea—and suggests that a "behavioral theory of schools" may be under construction by Bowles and Gintis. Clearly one can begin to be relevant once a reasonably comprehensive concept of the objectives or outputs of schools has been specified; and maximum standardized achievement test scores do not fill the bill. Better models also include more attention to how observations are generated and to the implications for econometric estimation. The use of programming techniques to form "envelope" estimates is one possible improvement and Levin's numerical example shows that it may be of some importance. Clearly our economic and econometric analyses of the education industry in general and its production function in particular are very crude and are not strong enough to support policy recommendations. I am more optimistic than Levin that the work to date has been harmless and may even have been helpful in moving toward more useful models. I am quite pessimistic about the chance of an estimated second derivative ever becoming the basis of a universally followed command which will halt or reverse the upward trend of educational costs.

NOTE

1. Garbage In Garbage Out.