Carnegie-Rochester Conference Series on Public Policy 37 (1992) 205-238
North-Holland

Improving educational outcomes while controlling costs

Eric A. Hanushek
University of Rochester
Steven G. Rivkin
UCLA
and
Dean T. Jamison*
World Bank and *UCLA*

Abstract

An analysis of educational outcomes and costs in U.S. schools shows rapidly increasing expenditures per student but little in the way of increased student performance. A decomposition of costs in the 20th century shows the powerful effects of decreased pupil-teacher ratios and increased costs of teachers. In the postwar period, however, teacher salaries have fallen relative to other college graduates. Other analyses of education production functions indicate that pupil-teacher ratios and currently structured salaries are not directly related to student learning. The obvious implication is that output based incentives are required to improve performance at acceptable costs.

Educational reform is a fashionable industry these days. Presidential candidates have positions on elementary and secondary education even in the absence of any significant federal role in education. Business leaders band together to issue policy statements on education. Parents demand better and safer schools. And citizen groups are increasingly sponsoring propositions designed to implement school reform through the ballot box.

*This work incorporates preliminary discussions of the Panel on the Economics of Educational Reform, a group funded by the Pew Charitable Trusts to analyze school reform recommendations. This paper, however, does not necessarily represent the views of other Panel members.

0167-2231/92/$05.00 © 1992 – Elsevier Science Publishers B.V. All rights reserved.

Why are all of these different groups demanding school reform? The answer depends upon the group in question, but three factors seem most important. The first, and most widely discussed, is quality. There is a widely-held opinion that the performance of students exiting the elementary and secondary system has been slipping. Business and industry, while complaining about the state of education since the founding of the republic, appear to have intensified their concerns about qualifications of young workers. These opinions are reinforced by virtually every report that is released about performance on standardized tests. Second, others are concerned about the costs of education. Within the context of fiscal pressures at all levels of government, educational expenditure has marched steadily upwards. While there have been periodic barriers erected to control costs—Proposition 13 in California or Proposition $2\frac{1}{2}$ in Massachusetts being just the most obvious—the general situation has been rising costs within an environment of increasingly larger announced demands for future expenditure. Third, leaving aside the problems of aggregate performance, disparities across racial and ethnic groups reinforce concerns about equity. Blacks and Hispanics, when contrasted with majority whites, remain quite far behind in both years of school completed and measured cognitive performance, despite the general move toward convergence over the past two decades.

The quality concerns, of course, dovetail with a variety of macroeconomic phenomena. The decline in aggregate productivity growth and the national concern about "competitiveness" are two factors that are frequently linked to education and quality of the labor force, at least in the public discussion. People worried about those issues invariably argue for improving the performance of the educational system.

This paper reviews what we know about the value of education to the individual and society and about how the costs and quality of schooling have changed over time. Based on this, it assesses a variety of policy options. As such, it presents little in the way of new empirical analyses. Instead, it attempts to bring together several different strands of work that bear on the overall problem.

1. Returns to education

Our interest in schooling derives almost exclusively from schooling's impact on subsequent performance by students, both in the labor force and out of the labor force. This interest relates both to individual effects and to aggregate effects on the economy. This section quickly identifies major conclusions about each of these issues that are derived from previous analyses.

Labor economists have thoroughly documented the relationship between income and schooling at the individual level. Virtually every study of income

and quantity of schooling (i.e., years completed) demonstrates that incomes rise systematically with schooling. Moreover, recent evidence suggests that these returns have been increasing. The gap between college graduate and high school graduate earnings appears to have expanded dramatically since the mid-1970s. While the question in the 1970s was whether or not people were getting too much schooling, few ask this question today.[1]

The estimates of Kevin Murphy and Finis Welch (1989, 1991) suggest that the ratio of college to high school earnings for all males went from about 1.5 in the mid-1970s to 1.7 in the mid-1980s. The changes for new entrants into the labor market were much more dramatic, going from roughly 1.3 to 1.7. When translated into rates of return for schooling, these estimates confirm that college education is a good investment. The increase, which is generally interpreted as an increasing demand for skills in the labor market, is dramatic by any account.

The estimates of earnings differences to years of schooling tell only part of the story, however. First, they do not say anything directly about the pattern or importance of variations in school quality, the central issue here. Second, they do not provide direct information about how education affects the macroeconomy.

Returns to quality. There has been considerable concern about individual differences in qualitative dimensions that bear on labor-market productivity. Much of this is directed at obtaining more precise information about the structure and determinants of individual earnings differences. This work does, however, have a direct relationship to the understanding of school quality and school policy, because differences in labor-force quality must relate in part to differences in school outcomes. Specifically, a branch of the work has investigated the labor-market returns to differences in cognitive test performance. This is a useful starting point for other considerations of school quality, even though a variety of adjustments (discussed below) are required.

. Over an extended period of time, studies of the labor market have been concerned about how individual differences in cognitive ability affect earnings (and modify the estimated returns to quantity). The early work was subsumed under the general topic of "ability bias" in the returns to schooling. In that, the simple question was whether the tendency of more able individuals to continue into college led to an upward bias in the estimated returns to school (because of a straightforward omitted variables problem).[2] The correction most commonly employed was the inclusion of a cognitive ability measure in the earnings function estimates. Most of the early work concen-

[1]The situation is best captured by publication in 1976 of Richard Freeman's book, *The Overeducated American*, which expanded on a number of articles about the falling rate of return to college education.

[2]See, for example, Griliches (1974).

trated on how the estimated returns to schooling were altered by inclusion of cognitive ability. The estimated direct effects of cognitive achievement on earnings, however, generally indicated relatively modest impacts of variations in cognitive ability after holding constant years of schooling.[3] In this work, there was no real discussion of what led to any observed differences in cognitive ability, although much of the work implicitly treated it as innate, and not very related to variations in schooling.[4] Further, all of this work relied on nonrepresentative samples of the population.

The most recent direct investigations of cognitive achievement, however, have suggested generally larger labor-market returns to individual differences. For example, Bishop (1989, 1991), O'Neill (1990), and Murnane, Willett, and Levy (1991) each find that the earnings advantage to higher achievement on standardized tests is very substantial. These results are derived from quite different approaches. Bishop (1989) worries about the measurement errors that are inherent in most testing situations and demonstrates that careful treatment of that problem has a dramatic effect on the estimated importance of test differences. O'Neill, Murnane, Willett, and Levy, and Bishop (1991), on the other hand, simply rely upon more recent labor-market data along with more representative sampling and suggest that the earnings advantage to measured skill differences is larger than that found in earlier studies (even without correcting for test reliability). There is some question of whether this new evidence results from an actual widening of the returns of differential skills or whether it is something to do with the sampling, testing, or general analytical methods. The direct tests of time interactions in Bishop (1991) do not show a widening, but the evidence is not conclusive.

The NAS/NRC study on employment tests (Hartigan and Wigdor 1989) also supports the view of a significant relationship of tests and employment outcomes, although the strength of the relationship appears somewhat less strong than that in the direct earnings investigations. It considers the relationship between the General Aptitude Test Battery (GATB), the standard employment test of the Department of Labor, and job performance. Their synthesis of a wide number of studies suggests a systematic but somewhat modest relationship with correlations to performance on the order of .2 to

[3]An exception to the generally modest relationship of cognitive performance and income is the work of Young and Jamison (1974). Using a national sample of data on reading competence, they find a strong influence of test scores on income for whites (but not blacks). This held in both recursive and simultaneous equations models of the joint determination of achievement and income.

[4]Manski (forthcoming) represents more recent work with this same general thrust. He recasts the issue as a selection problem and considers how ability or quality interacts with earnings expectations to determine selection into schools. Currently, however, no empirical work along these lines identifies the quantitative importance of selection or the interaction of school quality and earnings in such models.

.4. The analysis also finds that the validity of these tests in predicting performance has gone down over time. These results, being at somewhat odds with the recent studies, may simply reflect the specialized nature of GATB.[5]

A related look at school quality is found in Juhn, Murphy, and Pierce (1991). They consider the reversal of the trend in black-white earnings differences that occurred during the 1980s. While lacking direct evidence on school quality, they produce evidence that is very consistent with increasing skill demands in the economy and with an increased role for school quality.

An additional part of the return to school quality comes through continuation in school. There is substantial evidence that students who do better in school, either through grades or scores on standardized achievement tests, tend to go farther in school (see, for example, Dugan 1976 and Manski and Wise 1983). Rivkin (1991) finds that variations in test scores capture a considerable proportion of the systematic variation in high school completion and in college continuation. Indeed, Rivkin (1991) finds that test-score differences fully explain black-white differences in schooling.

Economic growth and productivity. While much of the motivation and concern about the educational system is directly linked to perceived problems of economic growth and labor productivity improvement, the empirical basis for this is thin. Considerable attention has come from the new growth literature which includes aggregate human capital in various fashions (e.g., Lucas 1988; Romer 1989). The majority of this growth work has remained theoretical. On the empirical side, a variety of growth accounting and other analyses are available, but none suggest that productivity growth is closely related to school quality. Instead they tend to point to other aspects of the economy; see, for example, Congressional Budget Office (1987a, Chapter 3).

Going from individual productivity and earnings to the pace of aggregate productivity improvement is tenuous. Bishop (1989) applies his individual earnings results to a simulation model. This analysis, however, follows a discussion of how previous changes in school quality could *not* be the cause of the already observed declines in productivity growth (because of the pure timing of the events). It also requires a wide range of assumptions about the determinants of productivity growth.

The message from consideration of growth and productivity is that, even though a linkage with school quality seems plausible, the case is not fully made. In particular, there is little guidance about how to measure quality differences, and there is little to suggest one way or another about the use of common measures such as test scores.

[5]The GATB is a very old test that may not reflect changes in the economy. It also suffers from some psychometric problems (see Hartigan and Wigdor 1989). The central purpose of the study was assessment of the Department of Labor practice of providing test information normed to racial groups.

Other benefits to education. Education has also been linked to a wide variety of nonmarket results (see, for example, Haveman and Wolfe 1984 or Michael 1982). A variety of studies have pinpointed the importance of schooling in determining health status and health expenditure (e.g., Grossman 1975), in affecting consumption (e.g., Michael 1972), in relating to child upbringing (e.g., Leibowitz 1974), and in social phenomena such as crime or voting (e.g., Ehrlich 1975 or Niemi and Sobieszek 1977). There is substantial evidence that more schooling appears to be positively related with better outcomes in each of these areas. With few exceptions, though, there is no direct investigation of how varying quality of schooling relates to such outcomes.

2. The record of performance

Student learning[6]

The view we adopt here is that a significant portion of economic growth comes from growth in the size of the effective labor force. By effective labor force we rely on some notion of worker quality which includes additions to the human capital of each worker. As a simple view, we think of human-capital additions as coming from either increases in the average years of schooling or increases in the knowledge, cognitive skills, etc. per year of schooling. The simple evidence suggests rapid increases in both sources of human-capital additions through the 1950s and into the 1960s. After the mid-1960s, however, it appears that the rate of increase in years of schooling slowed and that what additions have occurred have been at least partly offset by decreases in the cognitive achievement of students. We present evidence below on the overall changes in cognitive ability as measured by a number of different test instruments.[7] (In the next section, we consider the influence of schooling per se on measured test score differences.)

Tests provide a convenient quantitative measure of differences in performance. By being able to provide a consistent instrument, it is possible to compare performance of students in different learning environments. The key is objective assessment versus subjective assessment from such things as grades or other descriptions of performance. These measures also provide the possibility of analyzing the importance of schooling in economic growth and so forth.

[6]The information about changes in U.S. test scores relies heavily on the work of Daniel Koretz that is found in Congressional Budget Office (1986, 1987b).

[7]Bishop (1991) takes a similar view, but attempts to put the data together much more formally. He arrives at very similar conclusions based on a long series of Iowa tests and on years of schooling completed.

Perhaps the strongest point about tests, however, is the ability to evaluate changes in the organization and character of the school environment. While such things as lifetime incomes might be a conceptually preferable measure of student success and achievement, no data can be available on such things until well beyond the point when analysis would be relevant or useful.

The weaknesses of tests are, however, important. First, every available test evaluates only a very limited range of skills. Relatedly, there is no reason to be sure that mastery of skills tested is highly correlated with other skills which may also be important in the labor market.[8] Second, testing skills, not just cognitive skills, may influence performance. Closely related to this is the possible "cultural non-neutrality" of the tests. (If there is significant "teaching to the tests," both the predictive ability of test scores might fall and the real skills in other areas not tested may fall.)

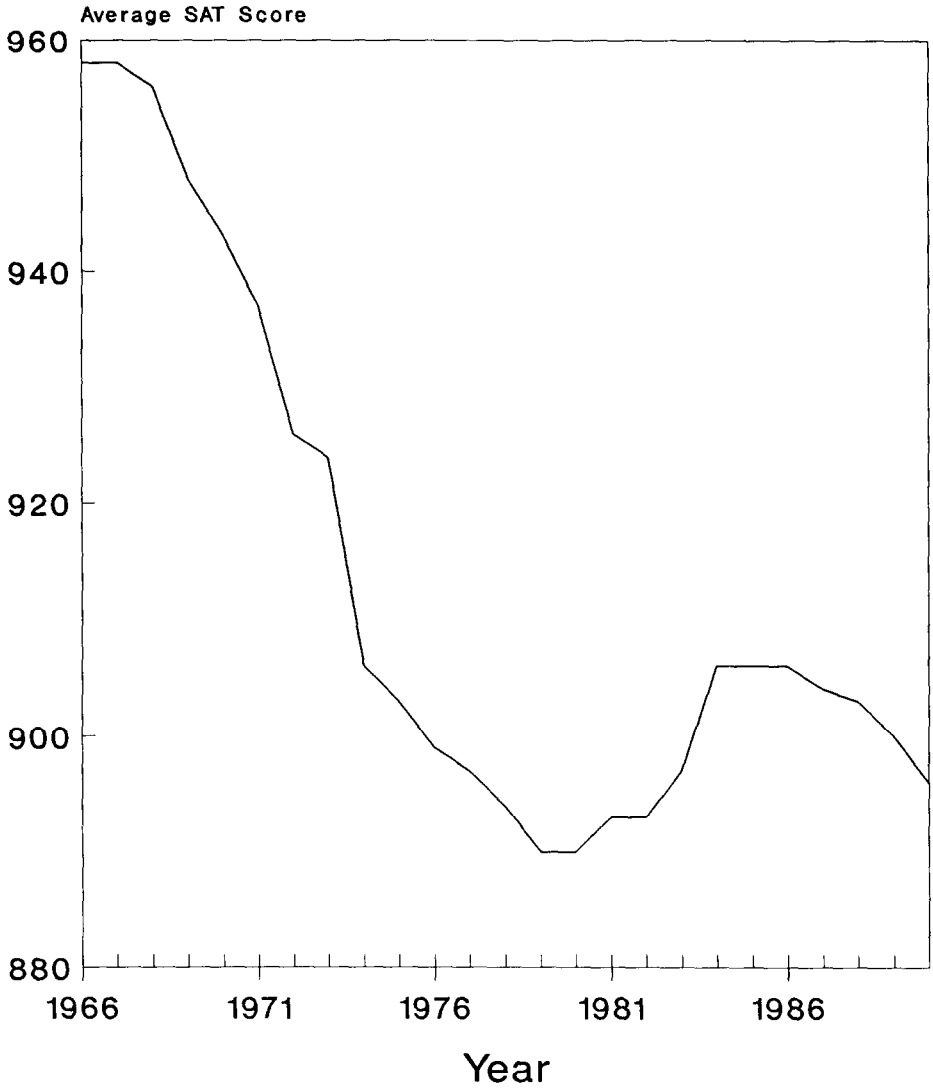
The general trend information on test scores is easily described. (See Figure 1). Average test scores on the SATs fell from 1963 through 1979, amounting to a drop in the mean of almost .5 standard deviations in verbal and .3 s.d. for the math portion. There was a recovery in the early 1980s, but, perhaps surprisingly, this stopped and regressed. The SAT test is potentially highly affected by changing composition of test-takers. The proportion of high school seniors taking the test rose rather steadily until just the past few years. Further, since the SAT is only taken by older students, it is difficult to detect the timing and character of changes at earlier grades until considerable time has past.

Other tests which are not as subject to selection problems also provide more precise information about the changes in performance at different grades. Specifically, Iowa achievement tests have been given over a long period of time and provide a time series for tests that do not involve individual self-selection. The story of the score changes on these tests appears to be one of systematic cohort difference in performance. As shown in the analysis of the Congressional Budget Office (1986), each of the different grade-level tests dips in the 1970s. The low point, however, occurs systematically when students at or near the birth cohort of 1964 reach a specific grade.

If the SAT cycle is compared with the Iowa cycle, the troughs line up almost exactly after adjusting for age of the test-takers (see Congressional Budget Office, 1986). The onset of declines across tests is not, however, as precise as the dates for the troughs. The Iowa test scores had recovered to their prior peak in the 1980s. This leads to a forecast that the SATs should soon resume their improvement, perhaps getting back to their earlier peaks, as cohorts who have performed better age and move through the educational

[8]Tests, for example, often measure speed of completion of tasks, something that might not be highly related with labor-market skills.

Figure 1
Scholastic Aptitude Test Scores
1966-1990



system.[9] Note, however, that the recent flattening and fall-off in the SATs after 1985 is not consistent with the earlier Iowa scores.

The question is: Will the improving trend continue? This would imply a resumption of the secular improvement in test scores that was seen before the mid-1960s. Or, having recovered, will test scores simply stagnate or enter a new period of decline? The answers to these questions have clear implications for the contribution of education to economic growth in the future. Unfortunately, the evidence for the youngest cohorts is not yet available to answer these questions.

International performance. One of the motivating factors behind concern about education, at least in the public discussions, is the role of education in determining economic competitiveness of the U.S. economy (vis-à-vis the economies of Japan or other economic trading partners). While the full implications of direct test-score comparisons are difficult to fit easily into international trade models, the ranking of the United States on common tests given to twelve different populations of 13-year-olds is interesting. The results from mathematics and science tests given in 1988 give a disheartening picture of U.S. student competencies. For example, "In Korea, 78 percent of Korean 13-year-olds can use intermediate math skills to solve two step problems...compared to only 40 percent...in the United States" (Lapointe, Mead, and Phillips 1989, p. 10). The picture is virtually identical for science achievement.[10]

Testing performance across countries and school systems is difficult, in part because the curricula of countries differ substantially. The important aspect of the previous comparisons is that they relied upon the international version of a standard U.S. test (the National Assessments of Educational Progress, or NAEP, test). In other words, we were at or near the bottom on our own tests.

Previous mathematics tests show similar results, although they rely upon different test instruments. McKnight *et al.* (1987) place U.S. 13- and 17-year-olds at the bottom of a larger sample of national tests given in the early 1980s. To complicate the interpretation, however, it should be noted that the U.S. students also placed very low on similar mathematics tests given in the early 1960s, a time when our schools were presumably performing better.

Racial differences. The comparisons of scores by race give two pictures. First, blacks have scored much lower than whites on standardized tests. This finding was first highlighted in the Coleman Report (Coleman *et al.* 1966)

[9]This forecast must, of course, be qualified by the general concerns about the special nature of the SAT and the test-taking population.

[10]It is also interesting to juxtapose performance with attitudes. In mathematics, two-thirds of the U.S. students felt "they are good at mathematics," while only 23 percent of the Koreans-the best performers-felt the same (Lapointe, Mead, and Phillips 1989, p. 10).

where blacks in the urban north were some three years behind their white counterparts in quality terms in 1965. The gaps were even larger when rural populations were compared. Other tests show similar pictures.

Second, the gap between blacks and whites has narrowed over time. This is seen on the SAT tests and on other tests that can be compared over time (see Congressional Budget Office 1986). Blacks did not take the same downturn that whites did, and that helped to narrow the gaps.

The interpretation of racial differences in performance is, of course, complicated by the geographic concentration of blacks and by other systematically different family conditions.

B. History of cost growth: 1890–1988

In many respects, the schooling process has changed little during the last 100 years. While there might be increasing use of audio-visual equipment, computers, and other learning aids, the extended use of educational technology does not show up as simple capital-labor substitution. Indeed, it has not increased the number of students for each teacher. Instead, the education sector actually reduced the student-teacher ratio in the face of rising labor costs, substantially increasing educational expenditure. This implies that the productivity of teachers, at least as measured by students taught per teacher, has not risen through time.

The economy as a whole experiences substantial growth in the productivity of labor during this period. Real wages rise roughly in proportion with the rise in labor productivity. This implies that the education sector, a heavy user of educated labor, will face rising labor costs. This rising cost of inputs will, however, be offset by the rising demand for its outputs (i.e., educated labor). An early assessment of factors influencing growth in education expenditure (Federal Council on Science and Technology 1972) addressed the extent to which these changing demands for educated labor were operating in the U.S. education sector. We expand on this analysis in what follows.

We examine the growth in educational expenditure between 1890 and 1988, both in absolute terms and relative to Gross National Product. Subsequently, we analyze the extent to which rising labor costs, declining pupil-teacher ratios, and other factors explain the expenditure increase.

Real public expenditure on primary and secondary education in the United States rises from $2 billion dollars in 1890 to over $170 billion in 1988.[11] This almost 100-fold increase is more than triple the growth rate of Gross National Product during this period: educational expenditure increased from less than 1 percent of GNP in 1890 to over 3 percent of GNP in 1988.

[11] All monetary measures are in GNP deflated constant 1990 dollars; educational expenditure refers to current educational expenditure and excludes capital costs.

Increasing enrollment accounts for a substantial portion of the rise in spending. But rising per student expenditure explains the bulk of the change in educational outlays. Figure 2 plots the increase in per student expenditure that occurs between 1890 and 1988. Per student expenditure is $164 in 1890, $772 in 1940, and $4,253 in 1988, roughly quintupling in each fifty-year period.

Why has education become so much more expensive over time? We begin to answer this question by decomposing changes in educational expenditure into a number of factors, which are loosely grouped into three categories: quantity, intensity, and cost. "Quantity" captures expenditure changes related to student enrollment, which in turn reflect increases or decreases in either the school-age population, school enrollment rate, or the division of students among public and private schools. "Intensity" refers to factors that affect the level of student inputs to schooling over the year. These include the length of the school year and the pupil-teacher ratio. Finally, "cost" refers to changes in the price of instructional personnel.

Based upon the results of the decomposition, we examine in detail the changes in the single most important factor in the increase in per student instructional expenditure: the price of instructional personnel.[12] Increases in the price of instructional personnel are separated into a cost component, reflected by rising wages of comparably skilled professionals, and a quality component, indicated by changes in the earnings of teachers in comparison to other skilled professionals. These discussions focus on recent years, due to both greater data availability and relevance to present-day policy discussions.

Table 1 traces the growth in real educational expenditure between 1890 and 1988.[13] The top panel divides expenditure into instructional staff expenditure and noninstructional expenditure. Over the entire period, noninstructional expenditure increases most rapidly. Note, however, that, while we employ these categories, they are not well-defined for policy purposes. Noninstructional expenditure includes such things as teacher benefits and health insurance, fixed charges by the existing accounting conventions but part of instructional expenditure for most purposes.

The bottom portion of this table shows changes in major factors that affect expenditure. Indeed, we will use these six characteristics to decompose the rise in educational instructional expenditure. We define the following

[12]While there has been much recent debate about the role of noninstructional expenditure, available data do not allow for accurate analysis of this component. Therefore, the following analysis concentrates on instructional costs.

[13]Decennial censuses are used to provide data on the school-age population. The remaining information is taken from various tables in issues of the Digest of Education Statistics and from Historical Statistics of the United States: Colonial Times to 1970.

Figure 2
Real Expenditure per Student: 1890-1988

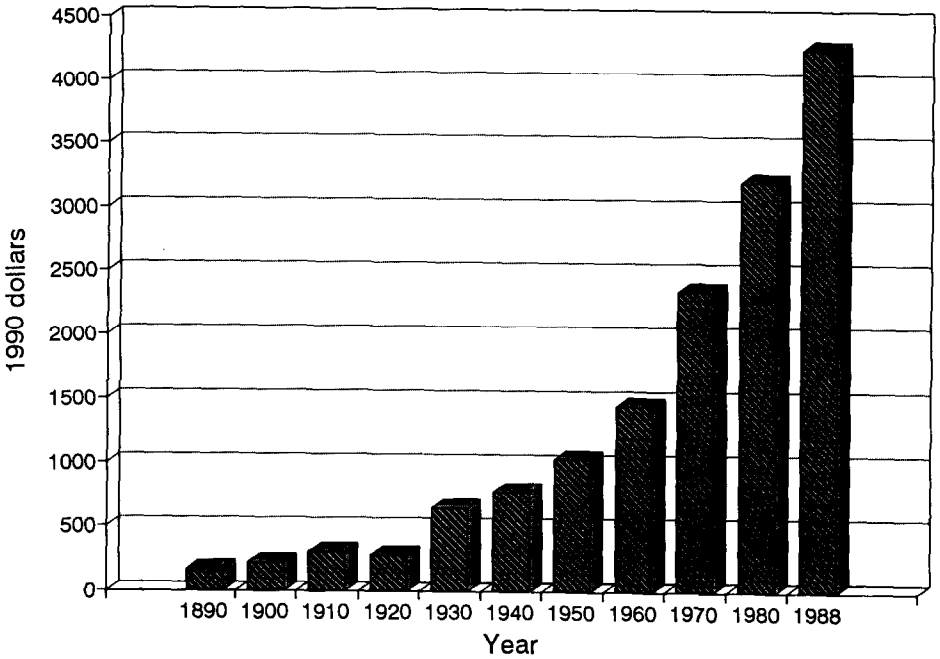


Table 1:
Values of Total Current Educational Expenditure
and the Explanatory Variables:
1890–1988 (1990 dollars)

	1890	1940	1950	1960	1970	1980	1988
Total Current Expenditures (billions)	2.09	19.65	25.79	52.47	107.1	133.5	170.4
Instructional Staff Expenditure (billions)	1.68	13.22	15.78	31.95	61.48	61.41	80.43
Noninstructional Staff Expend (billions)	0.41	6.43	10.01	20.52	45.62	72.09	89.97
Quantity							
School Age Population (millions)	21.2	34.8	34.9	48.7	59.8	56.1	52.6
Enrollment Rate (%)	68.4	80.7	81.6	85.8	85.8	83.3	86.3
Public School Enrollment (%)	87.8	90.7	88.1	86.4	88.9	89.1	88.2
Intensity							
Pupil/Teacher Ratio	35	28.1	26.3	24.9	20.5	17.4	15.8
Days Per Year	135	175	178	178	179	179	179
Input Cost							
Daily Wage Of Teachers	34.2	83.3	93.1	123.7	154.7	143.4	177.2

identity as a basis for decomposing changes in educational expenditure:

$$\text{TOTAL INSTRUCTIONAL EXPENDITURE} = \text{POP}\cdot\text{ENRATE}\cdot\text{PUBLIC}/\text{PTRATIO}\cdot\text{DAYS}\cdot\text{TPRICE} \quad (1)$$

where POP = school age population; ENRATE = enrollment rate of school-age population; PUBLIC = proportion public school enrollment; PTRATIO = pupil/teacher ratio; DAYS = school days per year; and TPRICE = average daily wage of teachers and staff.

Taking the log of equation 1 gives:

$$\text{Ln(TOTAL INSTRUCTIONAL EXPENDITURE)} = \text{Ln(POP)} + \text{Ln(ENRATE)} + \text{LN(PUBLIC)} - \text{Ln(PTRATIO)} + \text{Ln(DAYS)} + \text{Ln(TPRICE)} \quad (2)$$

Changes over time are calculated by subtracting Ln(TOTAL INSTRUCTIONAL EXPENDITURE) for year $i - n$ from Ln(TOTAL INSTRUCTIONAL EXPENDITURE) for year i, where n is the length of time between the two years being compared. The proportion of the change in TOTAL INSTRUCTIONAL EXPENDITURE accounted for by a particular variable equals the natural logarithm of that variable for year i minus the natural logarithm of that variable for year $i - n$, divided by the natural logarithm of TOTAL INSTRUCTIONAL EXPENDITURE for year i minus the natural logarithm of TOTAL INSTRUCTIONAL EXPENDITURE for year $i - n$. Using this accounting framework implies that factors changing proportionately more will account for a greater percentage of expenditure growth.

The proportions of change explained by each of the six variables for different time periods are presented in Table 2. We divide the years 1890–1988 into three time periods: 1890–1940, 1940–1970, and 1970–1988. The post-war period is divided between years of rapid growth in student enrollment and years in which enrollment declines. Within 1970 and 1988 changes are further given for the decade of the 80s. (Note from Table 1 that total instructional expenditure is virtually constant during the 70s, and, therefore, the decomposition is not useful.)

1890–1940. The rapidly-rising school age population and increasing public school enrollment rate account for roughly one-third of the $11.5 billion increase in educational expenditure between 1890 and 1940. The school-age population grows by 13.6 million, the enrollment rate rises from 68.4 percent to 80.7 percent, and the percentage of students attending public schools increases by 3 percentage points, yielding an overall increase of 12.7 million public elementary and secondary school students during this period.

The cost of teachers accounts for the 43 percent of expenditure growth between 1890 and 1940. The price of teachers increases by a factor of 2.5.

Table 2:
Percentage of Change in Instructional Staff
Expenditure as Explained by Changes in the Following Variables:
1890–1988 (by periods)

	PERIOD				
VARIABLES	1890–1988	1890–1940	1940–1970	1970–1988	1980–1988
QUANTITY					
School Age Population	23.5	24	35.3	-48.2	-24.2
Enrollment Rate	6	8	4	2.1	12.7
Public School Enrollment	0.1	1.5	-1.3	-2.9	-3.5
INTENSITY					
Pupil Teacher Ratio	20.6	10.7	20.3	98.5	36.7
Days Per Year	7.3	12.7	1.4	-0.2	0
INPUT COST					
Price of Teachers	42.5	43.1	40.3	50.7	78.3
TOTAL	100	100	100	100	100

Changes in the length of the school year and the pupil-teacher ratio (the two elements of intensity) account for over one-fifth of the expenditure increase. The average school year lengthens by 40 days. The pupil-teacher ratio declines from 35 students per teacher in 1890 to 28 students per year in 1940, a decrease of roughly 20 percent in 50 years.

1940–1970. The immediate postwar period, while differing in details, has a very similar pattern of cost increases to that in the prewar period. Increases in simple numbers of students attending public schools between 1940 and 1970 account for somewhat more of the expenditure increase compared to the earlier period. Public school enrollment increases by 25 million students, accounting by itself for 35 percent of the expenditure increase. The overall school enrollment rate rises by 5 percentage points during this period, while the proportion of students attending public schools falls by 2 percentage points. But, the role of these latter changes, much more important before 1940, pales in comparison to the rapid growth of the school-age population during the baby boom.

The cost of teachers and staff accounts for forty percent of the expenditure increase. The price of instructional staff nearly doubles in real terms, increasing from $83 per day in 1940 to $155 per day in 1970. This price increase has the largest impact on expenditure.

The two intensity components account for the remaining 20 percent of the expenditure increase, almost all of which is explained by the declining pupil-teacher ratio. Between 1940 and 1970 the school year lengthens only slightly on average and explains little of the expenditure rise. In contrast, the pupil-teacher ratio decreases from 28.1 to 20.5 during this 30-year period, which exceeds the decrease in the previous 50 years.

1970–1988. The character of cost changes over the past two decades changes dramatically from the earlier periods of the century. Public school enrollment in 1988 is 5.6 million students less than enrollment in 1970. The decrease is due almost entirely to a decline in the school-age population. The overall enrollment rate remains at 86 percent, and roughly 11 percent of students attend public schools. Nevertheless, expenditure continues to grow, reflecting the rapid pace of per-student expenditure growth.

Teacher cost factors account for a substantial portion of the increase in per-student expenditure, though as Table 2 shows, its impact differs in the two decades. Between 1970 and 1980 the price of teachers declines by an average of over 10 dollars per day. But between 1980 and 1988 the price of teachers jumps by over 30 dollars per day, imposing tremendous cost pressures on schools. Over the 1970–1988 period, teacher cost increases roughly equal the total cost savings from reduced quantity of school children.

The largest factor in expenditure growth is the decline in the pupil-teacher ratio. This factor by itself is sufficient to account for the aggregate expenditure change of the period. As Table 1 shows, the rate of decline slows in the 1980s, but its impact on expenditure remains strong. As in the prior 30-year period, the length of the school day changes little between 1970 and 1988 and consequently has little impact on expenditure growth.

The rising price of instructional staff. The average daily wage of instructional staff increases from $34 in 1890 to $83 in 1940 and to over $177 in 1988. Much of the increase results from the rise in real wages occurring throughout this period. The earnings of teachers may also change relative to other similarly skilled workers. Any increase (decrease) in teacher wages beyond that occurring in other sectors reflects a change in where teachers are drawn from the distribution of workers in the labor force. This is not to say that nonpecuniary factors are unimportant in determining whether individuals choose to teach. Rather, we assume that nonpecuniary benefits or costs of teaching have not changed in comparison to those in other occupations, in which case changes in relative earnings function as a good index of where teachers fall in the labor force.

We use yearly earnings data for full-time wage and salary workers taken from the five decennial Censuses of Population between 1940 and 1980 and the 1988 Current Population Survey. This earnings measure includes money

teachers receive in other occupations in addition to teaching.[14]

Teacher earnings are compared to the earnings of college graduates who do not teach. Private school and public school teachers are grouped together; therefore some movement in relative teacher earnings is due to changing earnings of private school teachers. But since only roughly 10 percent of students attend private schools throughout the period, it is unlikely that movement in the earnings of private school teachers will have a large impact on the overall relative wages of teachers.

Between 1940 and 1970 male and female teacher earnings decline monotonically at every age in comparison to other workers. The relative earnings of older teachers increase during the 1970s, a period in which the real wages of college-educated workers decline. The earnings of new entrants in the teaching profession, however, erode even during this period, and the downward trend in teacher earnings resumes during the 1980s.

We use these cross-sectional relative earnings profiles to decompose the change in the price of teachers into quality and cost components. The average earnings of teachers is compared to the average earnings of nonteachers, weighted by the sex/age composition of teachers. By weighting the earnings of nonteachers in this way, we are comparing teachers with nonteachers of the same age and sex composition. Therefore, relative teacher earnings will not change because teachers become younger or older on average than nonteachers; they will only change when the earnings of teachers change relative to nonteachers of the same age/sex category, which we believe is the best

[14]We are not comparing teacher salaries to salaries in other occupations. But that would be the wrong comparison. Teachers enjoy much longer vacations than most other workers. Overall earnings better reflect the monetary benefits of being a teacher as opposed to having a different primary occupation.

Using yearly earnings ignores changes in hours and weeks worked which affect hourly compensation rates. We use yearly earnings for two reasons. First, the census does not ask for an hourly or weekly wage. Instead, it asks for information on yearly earnings, number of weeks worked, and average hours worked per week. Hourly or weekly wages must be constructed from that information. In the case of teachers, their work weeks are likely to be very different during school vacations when they are working in other occupations, introducing a great deal of noise into the construction of weekly and hourly wage rates for teachers. Second and more importantly, the question on weeks worked in the Current Population Survey is more probing than that in the Censuses, encouraging people to count more weeks in which they worked only a few hours. Since teachers are more likely than nonteachers to work part-time weeks, this likely has a greater impact on teachers. A comparison of the 1980 Census and 1980 Current Population Survey reveals that while teacher/nonteacher yearly earnings ratios are similar in the two surveys, the teacher/nonteacher weeks worked ratio is much lower in the Census. Yearly earnings is chosen as the earnings measure in order to improve the compatibility between the 1988 Current Population Survey and the Censuses. It is important to note that the teacher/nonteacher ratio of average hours worked per week does not change substantially over time, suggesting that the yearly earnings measure is a good measure of the relative monetary benefits of teaching.

measure of relative teacher quality.

Table 3 shows the ratio of earnings of teachers to earnings of college graduates not in teaching. The ratio declines from 1.05 in 1940 to 0.85 in 1970 and remains relatively constant until 1988. This suggests that between 1940 and 1970, the full cost of teachers actually rises substantially more than the price of teachers that we employ. If relative earnings of teachers had remained constant between 1940 and 1970, the price of teachers would have risen by an additional 24 percent. Between 1970 and 1988 relative teacher earnings remain roughly constant, indicating that the entire movement in the price of teachers reflects changes in the cost of teachers.

Table 3:
Average Yearly Earnings of Teachers as a
Proportion of Earnings of Nonteaching College Graduates:
1940–1988[a]

Year	Men	Women	Total
1940	0.92	1.16	1.05
1950	0.86	1.03	0.94
1960	0.80	0.99	0.88
1970	0.78	0.92	0.85
1980	0.77	0.91	0.84
1988	0.78	0.88	0.84

Note: a. Average earnings of nonteachers is a weighted average of earnings based on the sex and age composition of teachers.

Comparisons of cross-sectional age earnings profiles provide a baseline measure of the potential change in teacher quality. While it may be natural to expect that the decline in teacher quality chronologically follows the decline in relative teacher earnings, the true effect depends on a number of unmeasured factors including the substitutability between teaching skill and other activities and the ability of schools to hire and retain effective teachers.

Summary. Changes in public school enrollment substantially impact educational expenditure. But even if the student population had remained constant throughout the last 100 years, expenditure would have risen by a factor of 25. Specific factors have had relatively greater impacts at different points in time, though two stand out as being of primary importance throughout the entire period: the rising price of instructional staff and the declining pupil-teacher ratio (see Table 2).

Moreover, the cost movements that are observed in Table 1 suggest that

the changes do not reflect adjustments in relative prices: the pupil-teacher ratio declines steadily, regardless of whether the price of instructional personnel increases or decreases. Part of the decline in the nationwide average pupil-teacher ratio might reflect changes in schools serving more difficult-to-educate students such as children from low-income families. But there has also been a general nationwide decline in the pupil-teacher ratio affecting schools in all types of communities. The only concession to cost pressures has been a decrease in the quality of teachers, as we discussed. But the magnitude of this decrease is overwhelmed by the decline in the pupil-teacher ratio, causing per student expenditure to march steadily upward.

3. Evidence on the efficacy of existing resource allocations

The previous discussions of cost movements pinpoint the standard answer to "school reform" that has been pursued over an extended period of time. Simple resource policies, closely linked to reducing class sizes and, in recent times, attracting different sets of teachers, have been the backbone of attempts to improve schools. The aggregate data, presented above, that link student performance and expenditure indicate that the traditional policies have not worked.

There is another, more persuasive set of studies that gives the same message. Over the past quarter century, there have been a large number of attempts to estimate an underlying educational production function. These confirm that there is no simple or systematic relationship between expenditure and student performance. The evidence is reviewed in the next section.

A. Production function estimates

Systematic investigation of the relationship between school inputs and student performance began in earnest with the publication of the *Equality of Educational Opportunity*, or, more commonly, the Coleman Report after its principal author, James Coleman. That massive study of the U.S. Office of Education was published in 1966 and had an immediate impact on both the research and policy communities.

The reason for the immediate impact of this research was its conclusion that school inputs had a minor influence on student achievement. Instead, variations in family background and in the backgrounds of other students in the school were the primary determinants of students' performance. These conclusions appear, as discussed below, to result largely from a misinterpretation of their results.

The more important impact of the Coleman Report, from both the policy and the scholarly perspective, is to be found not in the specific conclusions but in the approach. Through statistical methods, variations in student

achievement were related directly to the various factors thought to influence performance. While viewing education as a straightforward production process now seems obvious, such an approach was quite novel twenty-five years ago.

After the Coleman Report, in part spurred by its conclusions, a large number of investigations of educational production relationships were undertaken. These studies, when taken together, offer considerable support to the proposition made previously on the basis of macro data. Expenditure and performance are not systematically related.

The underlying model guiding most analysis is very straightforward. The output of the educational process—that is, the achievement of individual students—is directly related to a series of inputs. Some of these inputs—the characteristics of schools, teachers, curricula, and so forth—are directly controlled by policymakers. Other inputs—those of families and friends plus the innate endowments or learning capacities of the students—are generally not controlled. Further, while achievement may be measured at discrete points in time, the educational process is cumulative; inputs applied sometime in the past affect students' current levels of achievement.

A majority of studies into educational production relationships measure output by standardized achievement test scores, although significant numbers have employed other quantitative measures including post-school earnings, student attitudes, school attendance rates, and college continuation or dropout rates. The general interpretation, particularly with the test scores, is that these are indicators of future success, either in further schooling or in the labor market. This interpretation is supported by previous labor-market studies, as reviewed above.

Empirical specifications have varied widely in details, but they have also had much in common. Family inputs tend to be measured by socio-demographic characteristics of the families, such as parental education, income, and family size. Peer inputs, when included, are typically aggregate summaries of the socio-demographic characteristics of other students in the school. School inputs include measures of the teachers (education level, experience, sex, race, and so forth), of the school organization (class sizes, facilities, administrative expenditure, and so forth), and of district or community factors (for example, average expenditure levels). Except for the original Coleman Report, most empirical work has relied on data constructed for other purposes, such as the normal administrative records of schools that might be supplemented in some manner.[15]

[15] As discussed elsewhere (Hanushek 1979, 1986), a variety of empirical problems enter into estimation and the subsequent interpretation of results. The most significant general problems are the lack of measurement of innate abilities of individuals and the imprecise measurement of the history of educational inputs. Both the quality of the data and the

There is a "core" set of factors—those that determine basic expenditure—that has been broadly investigated in the production function context. Instructional expenditure makes up about two-thirds of total school expenditure. Given the number of students in a school district, instructional expenditure is in turn determined mostly by teacher salaries and class sizes. Finally, most teacher salaries are directly related to years of teaching experience and educational levels completed by the teacher. Thus, the basic determinants of instructional expenditure in a district are teacher experience, teacher education, and class size, and most studies, regardless of what other descriptors of schools might be included, will analyze the effect of these factors on outcomes. (These are also the factors most likely to be found in any given data set, especially if the data come from standard administrative records.)

This commonality in specification permits easy tabulation of the effects of these expenditure parameters. Table 4 presents the results of 187 separate "qualified studies" found in 38 separate published articles or books.[16] Since not all studies include each of the expenditure parameters, the first column in Table 4 presents the total number of studies for which an input can be tabulated–for example, 152 (of the 187) studies provide information about the relationship between teacher-student ratio and student performance. The available studies provide regression estimates of the partial effect of given inputs, holding constant family background and other inputs. These estimated coefficients have been tabulated according to two pieces of information: the sign and the statistical significance (5-percent level) of the estimated relationship.

The table makes it clear that the conventional views about school policy are simply wrong. Smaller classes, more educated teachers, and more experienced teachers are not systematically related to higher student performance.[17] This first set of findings is particularly important because these three factors tend to identify variations in instructional expenditure per student. Teacher

estimation techniques are very important in interpreting any specific findings, but, as discussed below, these problems have less impact on the overall findings illuminated here.

[16]The studies included in the table are defined as production function estimates: (1) published in a book or refereed journal; (2) relating some objective measure of student output to characteristics of the family and the schools attended; and (3) providing information about the statistical significance of estimated relationships. A given publication can contain more than one estimated production function by considering different measures of output, different grade levels, or different samples of students (but different specifications of the same basic sample and outcome measure are not duplicated). The table includes all studies that could be located through the end of 1988 that satisfied the above criteria. This tabulation is described in more detail in Hanushek (1986, 1989).

[17]Teacher experience is somewhat stronger, although hardly overwhelming. This stronger relationship may partly reflect the ability of more experienced teachers to select their school, which in turn may lead experienced teachers to select schools where student performance is otherwise high. See Greenberg and McCall (1974).

Table 4:
Summary of Estimated Expenditure Coefficients from
187 Educational Production Function Estimates

Input	Number of Studies	Statistically Significant +	Statistically Significant -	Statistically Insignificant Total	Statistically Insignificant +	Statistically Insignificant -	Statistically Insignificant Unknown Sign
Teacher/pupil ratio	152	14	13	125	34	46	45
Teacher Education	113	8	5	100	31	32	37
Teacher Experience	140	40	10	90	44	31	15
Teacher Salary	69	11	4	54	16	14	24
Expenditure/pupil	65	13	3	49	25	13	11
Administration	61	7	1	53	14	15	24
Facilities	74	7	5	62	17	14	31

Source: Hanushek (1989).

education and experience determine teacher salary, and pupil-teacher ratios merely indicate across how many students the salary of the teachers must be spread.

The remaining rows summarize information on other expenditure factors, including administration, facilities, teacher salaries, and expenditure per student.[18] Administration and facilities also show no systematic relationships with performance. This could partly be explained by variations in how they are measured. The quality of administration is measured in a wide variety of ways, ranging from characteristics of the principal to expenditure per pupil on noninstructional items. The character of facilities similarly is identified through both spending and many specific physical characteristics. Nevertheless, the available evidence again fails to support the conventional wisdom.

Finally, and not surprisingly, explicit measures of teacher salaries and expenditure per student provide no separate indication of their importance in determining achievement.[19] After all, the underlying determinants of these

[18]Information on each of these is less frequently available. This is partially explained by common reliance on administrative records which do not record each. The form of the analysis offers an additional explanation; for example, since expenditures per student are generally measured for districts, any of the 60 analyses for individual districts would find no variation in this input and thus could not include it.

[19]The expenditure and salary estimates are generally more difficult to interpret. Their interpretation is sometime clouded by including them in addition to teacher experience, education, and/or class size. Additionally, because prices can vary across the samples in the separate studies, it is more difficult to interpret the dollar measures than the real input measures. Finally, in terms of the results in Table 4, eight of 13 significant positive

expenditure were themselves unrelated to achievement.

B. Counter evidence on expenditure

The most widely discussed argument against the previous position refers to the work by David Card and Alan Krueger (1992). Their work, instead of starting with the school and looking at individual outcomes, begins with people in the labor market. It then attempts to fill in the school experiences relevant to workers observed in census data. Specifically, they use the public-use sample from the 1980 census to estimate rates of return to schooling for specific states of birth and birth cohorts (holding constant current region of residence). A second stage analysis then regresses these estimated returns on pupil-teacher ratio, length of school year, and relative teacher salaries for each state. This produces significant results pointing to the productive nature of educational expenditure–something quite at odds with the research previously reported.

A variety of factors could explain the apparent differences. The Card and Krueger analysis trades off better measures of educational outcomes (earnings variations over the life cycle) with much poorer measures of school characteristics (average state characteristics in state of birth). It is implausible, as Burtless (forthcoming) points out, that school officials are truly effective at producing something that they are not trying to produce (earnings) while being ineffective at producing something they are actively trying to produce (cognitive skills as measured by test performance). The plausible explanations of the difference in results are either that complicated measurement errors lead to the Card and Krueger findings[20] or that the schooling circumstances are quite different. In terms of school experiences, the individuals in the sample attended school between approximately 1925 and 1967, and separate consideration is given to segregated black schools in the south. Therefore, much larger variations in school factors are found in their analysis than in any more recent analysis. It could be that the wider variation permits detection of effects that cannot be separated from the background noise in other samples. Or, it could be that there are real thresholds of expendi-

expenditure results also come from the different estimates of Sebold and Dato (1981). In this study, imprecise measurement of family inputs suggests that school expenditures may in fact mainly be a proxy for family background.

[20]The models estimated are quite complicated not ony by virtue of being two-stage estimates but also be virtue of the complicated transformations and fixed-effect models estimated in the first stage. Thus, the simple errors in measurement story of biasing the estimates toward zero is not clearly appropriate in this case. Moreover, having no measure of family backgrounds, the story becomes even more complicated, because family backgrounds tend to be positively correlated with within-state spending patterns (the cause of most current school finance litigation). Finally, the methodology does not allow for overall (intercept) differences in earnings across states.

ture and resources, below which variations are important and above which variations are not.

This study provides some insights into schooling that need to be pursued further. It is, for example, possible that some of the low-resource areas in the country today would be below some quality threshold. Identification of such a threshold, if it exists, would be important in considering overall school policies. On the other hand, this argument should not be taken too far, because the previously discussed production function studies clearly have observations across a very wide range of schooling circumstances. They include rural areas in the south, central cities throughout the nation, and representative samples. There is no indication that the results given previously differ by any of these characteristics, which also index absolute levels of resources.

C. Variations in school and teacher quality

The previous sections concentrated on whether expenditure and the components of expenditure are related to student performance. The lack of relationship does not, however, imply that there are not differences among teachers and schools. There are very important differences among teachers.

A number of studies provide direct analyses of this overall question of differential effectiveness of teachers. They do this by estimating differences in the average performance of each teacher's students (after allowing for differences in family backgrounds and initial achievement scores).[21] The findings of these studies (Hanushek 1972, 1992; Murnane 1975; Armor et al. 1976; and Murnane and Phillips 1981) are unequivocal: Teachers and schools differ dramatically in their effectiveness.[22] The formal statistical tests employed in these studies confirm that there are dramatic differences in average gain in student achievement across teachers.

The faulty impressions left by the Coleman Report and by a number of subsequent studies about the importance of teachers have primarily resulted from a confusion between the difficulty of explicitly measuring components of effectiveness and true effectiveness. In other words, existing measures of characteristics of teachers and schools are seriously flawed and thus are poor indicators of the true effects of schools; when these measurement errors are avoided, schools are seen to have important effects on student performance.

[21] These studies are analyses of covariance or, equivalently, the use of individual teacher dummy variables in addition to measures of prior student achievement, family background factors, and other explicitly identified inputs.

[22] The findings for the United States are also supported by findings in developing countries; see, e.g., Heyneman and Jamison 1980 and Harbison and Hanushek (1992).

4. Policy implications

We take the conclusions of the previous evidence to be the following. First, there are real problems with the current operations of the schooling system. Performance, or student quality, appears to be down while resources devoted to schools are up dramatically. Second, we find that both the aggregate evidence and the production function evidence suggest that inefficiency in schools is an extremely important issue. Spending on schools is currently not closely linked to school performance. Moreover, simply devoting more resources to what we have been doing is unlikely to lead to much improvement. From these results, we consider a variety of alternatives.

Our policy considerations at this point are limited in ways that may be inappropriate. Specifically, we do not expand our thinking much beyond the school as it exists today. This is unfortunate because influences outside of the school are extremely important. Families have a dramatic influence on a student's performance. Moreover, the readiness of a student for school may have lasting implications for a student's subsequent performance. Nonetheless, we restrict our attention to schools and ignore pre-school and extra-school policies. This limitation is one of convenience, because it is extraordinarily difficult to compare the efficacy of all kinds of programs that might relate to children.

A. What not to do

Two policy conclusions spring immediately from these overall results. First, since within the current institutional structure expenditure is not systematically related to performance, policies should not be dictated simply on the basis of expenditure. Second, common surrogates for teacher and school quality—class size, teachers' education, and teachers' experience among the most important—are not systematically related to performance within the current institutional structure, yet expenditure on surrogates has driven up student costs by a large multiple over the past several decades. Stopping or reversing these trends is a natural direction for policy.

These conclusions appear obvious and indeed seem to be subscribed to in principle by many policymakers. But violations of these principles also appear frequently and go unchallenged.

Take for example the area of clearest policy by both state legislators and the courts: the financing of local schools. Virtually all of the discussions (and court cases) related to school finance are phrased entirely in terms of the pattern of expenditure variations across districts. The argument for this is frequently one of expediency—since there is ambiguity about what factors affect performance and since legislators cannot realistically run local schools,

expenditure offers the only reasonable policy instrument.[23]

Or, just as obviously, local school boards are content to focus on class sizes and to negotiate contracts setting teacher salaries exclusively on the basis of teacher education and experience. State legislators themselves enter into regulating salaries and class sizes in different programs and mandating that teachers obtain a master's degree. And, as the cost evidence suggests, the clearest change in school policy over recent decades has been the reduction in class sizes.

B. Incentives and organizational changes

The other side of our consideration of policy issues is the obvious importance of getting incentives right. The current system has virtually no incentives for schools (or students) to preform well. The way we currently set policy is to dictate through contract, rules, or legislation what we believe the pattern of inputs should be. The point of the previous discussion was that we do not know how to specify the best set of inputs in order to increase quality.[24] This is not the only way to operate schools, however.

The alternative is to develop output incentives, i.e., mechanisms to induce teachers and schools to do the right thing as opposed to trying to specify exactly how education should be conducted. A variety of mechanisms fall into this category. Merit pay schemes for teachers are based on the idea that teachers who do well in the classroom should receive more rewards than those who do poorly. Merit school plans extend this to the school level. Choice plans, which themselves come in a variety of forms, attempt to develop performance incentives by having students and parents identify schools that are performing well.

Performance incentives generally share three aspects. First, they are conceptually appealing. Second, their operation depends crucially on the details of their specification and implementation. And, third, with the exception of merit pay, we know virtually nothing about how well they work in reality.

The conceptual appeal is that they are aimed directly at the object of policy–student performance. As such, they do not require detailed specification of the how-to's, something that has eluded us. They also admit to the possibility that there is not just one approach; simply put, different approaches may be appropriate for different people in different settings.

The generic categories nevertheless contain many possible structures, and each is likely to have different results. The case of choice is instructive. "Choice" is a term that is applied to the use of magnet schools, to intradistrict

[23]See Hanushek (1991) for a discussion of the details of school finance cases and policy.

[24]We obviously know, however, that our previous attempts have had large cost implications.

school selection plans, to interdistrict school selection plans, to tuition tax credits, and to vouchers.[25] The key to each is how resources move in response to the choices of students and parents. For example, if there is a system of intradistrict choice, all effect could be eliminated if high demand in a specific school is met by reassigning surplus teachers from low-demand schools. In other words, if teacher contracts or system rules guarantee that people in bad schools always have a job and are simply moved where students have chosen to go, one would expect very little net impact even though there is ostensibly choice. The situation, even with intradistrict choice, would be very different if reassignment of teachers were not guaranteed.

Similarly, interdistrict choice plans depend crucially on the rules of financing. Do state funds follow the student? Do local funds follow the student? Can a district charge tuition to out-of-jurisdiction students? The specific rules define the incentives for sending districts and for receiving districts.

Unfortunately, we do not know very much at all about the operation of different incentive plans, particularly of the choice variety.[26] Where plans have been introduced, little research or evaluation has been conducted.[27] Moreover, there is no mechanism by which we can learn from prior (or future) attempts at choice plans. The U.S. Department of Education, the logical place for such research and evaluation, does not have any system of developing or disseminating broad-based information about incentives in schools.

[25]These terms refer to standard categories of organizations. Magnet schools, used in a large number of urban districts, allow students to choose between a geographically assigned school and some other school drawing from a wider geographic area. The magnet schools, which may have special themes, are frequently developed to attract a racially diverse audience and thus to meet desegregation objectives. Intradistrict choice, such as practiced in Harlem or Cambridge, allows students to select from any school within the district, although choice is generally constrained by capacity and rules for choosing from queues. Interdistrict choice, as developed in Minnesota and transported to several other states, permits schools to cross jurisdictional boundaries to select a desired schooling situation. Tax credits and vouchers open up private school competition either through rebates on taxes or from actual checks that students can apply to whichever school is chosen. Within each category, many alternatives also exist.

[26]The one exception where some evidence is available is the merit pay plan. Merit pay has been attempted in various forms a great number of times, but essentially no school systems keep such schemes over a long period of time. If they are kept, they usually evolve into "extra pay for extra work." See Cohen and Murnane, 1986. At the same time, it is unclear just what forms of merit pay have been employed and whether, in this era of renewed interest in performance, an acceptable scheme could be developed.

[27]A serious research and evaluation program would also include an international component which compared systems with different organizational structures.

C. The challenge of technological change

We return to the cost decompositions. Simply to "keep up" with wages in the economy, expenditure on schools has risen 25 percent over this century. The increase has been even faster in recent decades. The current interpretation of expanding wage differentials by schooling as a reflection of increased skill demand suggests that this cost disadvantage is likely to expand. This implies that doing nothing (i.e., attempting to maintain current quality with current organization) will require continually increasing expenditure.[28] This suggests that a renewed look at productivity enhancement through alternative technologies may be very important. This point, discussed in an early assessment of educational reform (Federal Council on Science and Technology 1972), has been neglected in recent analyses.

There is by now extensive experience in the United States and abroad with efforts to utilize radio, television, and computer technologies in the classroom. Similarly, print media, with and without audiovisuals, have been used in distance instruction to dispense with traditional classrooms while still providing structured instruction. Student performance acceptance, usually examination-based, has provided for equivalency in accreditation with traditional programs. Reviews and evaluations can be found in Jamison, Suppes, and Wells (1974) and Office of Technology Assessment (1989).

Four conclusions are noteworthy. First, there is massive evidence to suggest that television or radio can provide a very close substitute for teacher time in the educational production process.[29] Second, there is mounting evidence that radio, television, and computers can be deployed in ways that increase the rate of learning, i.e., in ways that increase the productivity of student time. This evidence, while strongly suggestive of potential, is, however, presently based on very limited experience. Third, when schools based on classroom instruction have been replaced by distance education institutions (such as the British Open University), there is a well-established record of reducing direct costs but no evidence of improving the productivity of student time (except insofar as travel time is reduced). Finally, outside of some parts of higher education, there appear to be no incentives within the current U.S. structure for systems to adopt either cost-saving or productivity-enhancing innovation.

We conclude by noting, then, the interdependence between technologi-

[28]Since the value of output, educated labor, is also rising, this should not be interpreted as inefficiency in production. It does, however, imply that more goods will be traded off for schools and that increased attention should be given to the possibility for capital substitution.

[29]These results have the same flavor that the production function studies do in finding little differential influence (plus or minus) on student test scores. As such, they suggest that cost considerations should govern.

cal improvement and the incentive structure. Absent a better menu, incentive improvements face natural limitations. Equally important, even major productivity-enhancing innovations can be expected to languish absent incentives to develop and apply them.

References

Armor, D., Conry-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., Pauly, E., and Zellman, G., (1976). *Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools.* R-2007-LAUSD. Santa Monica, CA: Rand Corporation.

Baumol, W.J., (1957). Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis, *American Economic Review*, **57**: 415–526.

Baumol, W.J., and Bowen, W.G., (1965). On the Performing Arts: The Anatomy of Their Economic Problems, *American Economic Review*, **57**: 495–502.

Bishop, J.H., Did the Test Score Decline Cause the Productivity Growth Decline? *American Economic Review*, March, 178–197.

Bishop, J., (1991). Achievement, Test Scores, and Relative Wages, Marvin H. Kosters (ed.), *Workers and Their Wages*, Washington, D.C.: The AEI Press, 146–186.

Burtless, G., (forthcoming). Current Proposals for School Reform: An Economists's Assessment, Anthony O'Brien and Robert J. Thornton (ed.), *Economic Consequences of American Education*, Greenwich, CT: JAI Press.

Card, D. and Krueger, A., (1992). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States, *Journal of Political Economy*, **100**: 1–40.

Cohen, D. and Murnane, R.J., (1986). Merit Pay and the Evaluation Problem: Understanding Why Most Merit Pay Plans Fail and a Few Survive, *Harvard Education Review*, **56**: 1–17.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D., and York, R.L., (1966). *Equality of Educational Opportunity.* Washington: U.S. Government Printing Office.

Congressional Budget Office, (1986). *Trends in Educational Achievement.* Washington, D.C.: Congressional Budget Office.

Congressional Budget Office, (1987a). *The Economic and Budget Outlook: Fiscal Years 1988-92.* Washington, D.C.: Congressional Budget Office.

Congressional Budget Office, (1987b). *Educational Achievement: Explanations and Implications of Recent Trends.* Washington, D.C.: Congressional Budget Office.

Council of Economic Advisers, (1991). *Economic Report of the President.* Washington, D.C.: GPO.

Dugan, D.J., (1976). Scholastic Achievement: Its Determinants and Effects in the Education Industry, Joseph T. Froomkin, Dean T. Jamison, and Roy Radner, (eds.), *Education as an Industry*, Cambridge, MA: Ballinger.

Educational Testing Service, (1991). *Performance at the Top: From Elementary Through Graduate School*, Policy Information Report, Princeton, NJ: Educational Testing Service.

Ehrlich, I., (1975). On the Relation between Education and Crime, F. Thomas Juster (ed.), *Education, Income, and Human Behavior*, New York: McGraw-Hill.

Federal Council for Science and Technology, (1972). Committee on Automation Opportunities in the Service Areas (FCST), *Productivity and Efficiency in Education: A Report by the Educational Panel*, (mimeo).

Freeman, R.B., (1976). *The Over-Educated American*, New York: Academic Press.

Greenberg, D. and McCall, J., (1974). Teacher Mobility and Allocation, *Journal of Human Resources*, **9(4)**: Fall, 480–502.

Griliches, Z., (1974). Errors in Variables and Other Unobservables, *Econometrica*, **42(6)**: November, 971–998.

Grossman, M.J., (1975). The Correlation Between Health and Schooling, N.E. Terleckyj (ed.), *Household Production and Consumption*, New York: NBER, 147–211.

Hanushek, E.A., (1972). *Education and Race: An Analysis of the Educational Production Process*, Cambridge: MA: Heath-Lexington.

Hanushek, E.A., (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions, *Journal of Human Resources*, **14(3)**:

Summer, 351–388.

Hanushek, E.A., (1986). The Economics of Schooling: Production and Efficiency in Public Schools, *Journal of Economics Literature*, **24(3)**: September, 1141–1177.

Hanushek, R.A., (1989). The Impact of Differential Expenditures on School Performance, *Educational Researcher*, **18(4)**: May, 45–51.

Hanushek, E.A., (1992). The Trade-off Between Child Quantity and Quality, *Journal of Political Economy*, **100**: February, 84–117.

Harbison, R.W. and Hanushek, E.A., (1992). *Educational Performance of the Poor: Lessons from Rural Northeast Brazil*, New York: Oxford University Press, forthcoming.

Hartigan, J.A. and Wigdor, A.K.,(ed.), (1989). *Fairness in Employment Testing: Validity Generalization, Minority Issues, and the General Aptitude Test Battery*, Washington, D.C.: National Academy Press.

Haveman, R.H. and Wolfe, B.L., (1984). Schooling and Economic Well-Being: The Role of Nonmarket Effects, *Journal of Human Resources*, **19**: Summer, 378–407.

Heyneman, S. and Jamison, D.T., (1980). Student Learning in Uganda: Textbook Availability and Other Factors, *Comparative Education Review*, **24**: 206–220.

Jamison, D.T., Suppes, P. and Wells, S., (1974). The Effectiveness of Alternative Instructional Media: A Survey, *Review of Educational Review*, **44**: 1–67.

Juhn, C., Murphy, K.M. and Pierce, B., (1991). Accounting for the Slowdown in Black-White Wage Convergence, Marvin H. Kosters (ed.), *Workers and Their Wages*, Washington, D.C.: The AEI Press, 107–143.

Kosters, M.H., (1991). Wages and Demographics, Marvin H. Kosters (ed.), *Workers and Their Wages*, Washington, D.C.: The AEI Press, 1–32.

Lapointe, A.E., Mead, N.A., and Phillips, G.W., (1989). *A World of Differences: An International Assessment of Mathematics and Science*, Princeton, New Jersey: Educational Testing Service.

Leibowitz, A., (1974). Home Investments in Children, *Journal of Political Economy*, Mar/Apr. **82(2)**: Pt. II, S111–S131.

Lucas, R.E., (1988). On the Mechanics of Economic Growth, *Journal of Monetary Economics.*, July **22**:1, 3–42.

Manski, C.F. and Wise, D.A., (1983). *College Choice in America*, Cambridge: Harvard University Press.

McKnight, C.C. et al., (1987). *The Underachieving Curriculum: Assessing U.S. School Mathematics from an International Perspective*, Champaign, IL: Stipes Publishing Co.

Michael, R.T., (1982). Measuring Non-Monetary Benefits of Education: A Survey, W. McMahom and R. Geske (ed.), *Financing Education: Overcoming Inefficiency and Inequity*, Urbana: University of Illinois Press.

Michael, R.T., (1972). *The Effect of Education on Efficiency in Consumption*, New York: Columbia University Press.

Murnane, R.J., Willett, J.B. and Levy, F., (1991). Skills, Skill Payments, and the Mismatch Hypothesis, (mimeo).

Murnane, R.J., (1975). *Impact of School Resources on the Learning of Inner City Children*, Cambridge, MA: Ballinger.

Murnane, R.J. and Phillips, B., (1981). What Do Effective Teachers of Inner-City Children Have in Common? *Social Science Research*, **10(1)**: March, 83–100.

Murphy, K.M. and Welch, F., (1989). Wage Premiums for College Graduates: Recent Growth and Possible Explanations, *Educational Researcher*, **18(4)**: May, 17–26.

Murphy, K.M. and Welch, F., (1991). The Role of International Trade in Wage Differentials, Marvin H. Kosters (ed.), *Workers and Their Wages*, Washington, D.C.: The AEI Press, 39–69.

National Center for Education Statistics, (1991). *Digest of Education Statistics: 1990*, Washington, D.C.: GPO.

Nelson, F.H., (1991). *International Comparison of Public Spending on Education*, Research Department of the American Federation of Teachers, Washington, D.C.

Niemi, R. and Sobieszek, B.I., (1977). Political Socialization, *Annual Review of Sociology*, **3**: 209–233.

O'Neill, J., (1990). The Role of Human Capital in Earnings Differences Between Black and White Men, *Journal of Economic Perspectives*, **4(4)**: Fall, 25–46.

Office of Technology Assessment, (1989). *Power On! New Tools for Teaching and Learning*, Washington, D.C.: OTA.

Penner, R.G., (1988). Economic Growth, Isabel V. Sawhill (ed.), *Challenge to Leadership*, Washington, D.C.: Urban Institute.

Rivkin, S.G., (1991). *Schooling and Employment in the 1980s: Who Succeeds?* unpublished Ph.D. dissertation, University of California, Los Angeles.

Romer, P., (1989). Human Capital and Growth: Theory and Evidence, (mimeo), April.

Scitovsky, T. and Scitovsky, A., (1959). What Price Economic Progress? *Yale Review*, 95–110.

Sebold, F.D., and Dato, W., (1981). School Funding and Student Achievement: An Empirical Analysis, *Public Finance Quarterly*, **9(1)**: 91–105.

U.S. Bureau of the Census, (1975). *Historical Statistics of the United States, Colonial Times to 1970, Bicentennial Edition*, Parts 1 and 2, Washington, D.C.: GPO.

U.S. Bureau of the Census, (1973 through 1991). *Statistical Abstract of the United States*, various years 1972–1990, Washington, D.C.: GPO.

Young, K.H. and Jamison, D.T., (1974). The Economic Benefits of Schooling and Compensatory Education, paper presented at the annual meetings of the Econometric Society, December.