



Measuring Investment in Education

Eric A. Hanushek

The Journal of Economic Perspectives, Volume 10, Issue 4 (Autumn, 1996), 9-30.

Stable URL:

<http://links.jstor.org/sici?sici=0895-3309%28199623%2910%3A4%3C9%3AMIIE%3E2.0.CO%3B2-1>

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The Journal of Economic Perspectives is published by American Economic Association. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/aea.html>.

The Journal of Economic Perspectives
©1996 American Economic Association

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2002 JSTOR

<http://www.jstor.org/>
Thu Apr 11 20:31:09 2002

Measuring Investment in Education

Eric A. Hanushek

Political campaigns, from the U.S. presidency to local city elections, are routinely fought on the grounds of which candidate most favors investing in the education of our children. Lacking any definitive way of demonstrating such sentiments, however, the issue frequently comes down to which candidate is willing to spend the most tax dollars supporting local public schools. Reinforcing this position is a series of judicial proceedings that have for 30 years revolved around state constitutional requirements mandating the equitable provision of public education, a requirement generally cast in the courts purely in terms of equality of financial support for local school districts. On an international scale, the foundation of the aid programs of national governments and international agencies interested in fostering development of less developed countries has been investment of added resources aimed at boosting the human capital of the youth and thus ensuring the future growth of these economies.

Given these policy positions, it would at the very least be an embarrassment, and at the worst a potential policy disaster, to find that variations in resources devoted to schooling are not the primary factor determining student performance. But that appears to be the case. Three decades of intensive research leave a clear picture that school resource variations are not closely related to variations in student outcomes and, by implication, that aggressive spending programs are unlikely to be good investment programs unless coupled with other fundamental reforms. In simplest terms, how money is spent appears to be much more important than how much is spent—at least within the normal variations in spending levels in today's

■ *Eric A. Hanushek is Professor of Economics and Public Policy and Director of the Wallis Institute of Political Economy, University of Rochester, Rochester, New York. His e-mail address is hanu@troi.cc.rochester.edu.*

schools.¹ While such a perspective is now commonplace in, say, consideration of health care, it has been slow coming to education.

While this article concentrates on direct policy ramifications of issues related to school spending, it also has a subtext related to economic research and modeling. While measures of inputs are frequently thought of as convenient summaries of investment in human capital, they rest on a series of suspect assumptions. The discussion of measurement of human capital parallels some of the earlier debates about the choice of input or output measures of physical capital, but it also has unique components. Specifically, one must believe that inputs are converted efficiently to outputs and that measured school inputs comprise the bulk of all inputs into human capital—two assumptions that appear far from true.

This article begins by reviewing the main body of evidence about the effectiveness of resources devoted to schools. It then discusses some of the central implications for policy and for economic research. The implications for educational policy draw heavily on the discussion of the Panel on the Economics of Education Reform (PEER), a group of economists who met over a four-year period to discuss economic aspects of educational investment.² Those recommendations are only sketched here, but are detailed in its report, *Making Schools Work* (Hanushek and others, 1994).

The Aggregate Picture

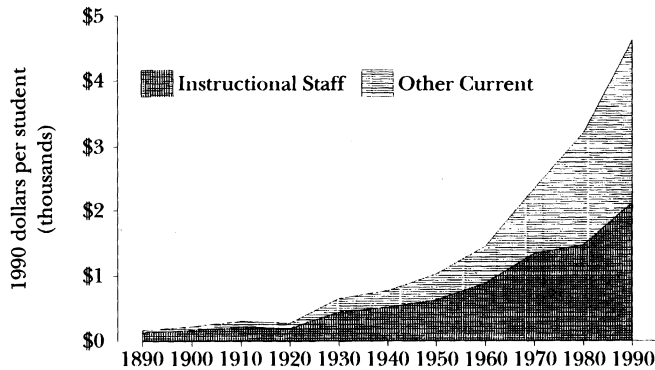
A discussion of investment in schooling quite naturally begins with an overview of the past record. In many respects, U.S. policy toward schools has been a real success story, as historic investment has produced a labor force of unrivaled skill and has contributed to the extraordinary economic growth of this century. The long-term picture, however, neither accurately reflects the current picture nor provides clear guidance about the policy decisions of today. Part of the confusion arises from translating considerations of the quantity of schooling—the educational attainment of the population—into implications about investments in improved quality of schooling. Part of the confusion also seems to reflect a lack of awareness of basic facts about the level and character of past investment in schooling.

There has been a dramatic rise in real expenditure per pupil over the entire twentieth century. Figure 1 shows that, after allowing for inflation, expenditures

¹ Sensing slippage between spending and performance, some political actors have gone to the opposite pole of the “investors” and have argued that all spending is wasteful and should be reduced. Interpreting calls for less school spending is, however, difficult, because part of it appears to reflect pure sentiments to limit government spending independent of any views about the productivity of school spending.

² The Panel on the Economics of Education Reform (PEER), which met over the period 1989–1993, included Eric A. Hanushek (chair), Charles S. Benson, Richard B. Freeman, Dean T. Jamison, Henry M. Levin, Rebecca A. Maynard, Richard J. Murnane, Steven G. Rivkin, Richard H. Sabot, Lewis C. Solmon, Anita A. Summers, Finis Welch and Barbara L. Wolfe. Its final report, *Making Schools Work*, was published by the Brookings Institution in October 1994.

Figure 1

Instructional Staff and Other Current Expenditure Per Student, 1890–1990

per pupil have increased at almost 3.5 percent per year for 100 years (Hanushek and Rivkin, 1997). The figure also divides expenditures into total instructional staff salaries and other expenditures per student. The spending increases that have occurred have come from three basic sources. The first two affect direct instructional staff expenditure: declines in pupil-teacher ratios and increases in the real salaries of teachers. Pupil-teacher ratios have fallen partly because of direct programs to reduce class sizes, and partly because of the introduction of new supplementary programs, which often result in more individualized or small-group attention.³ Real teacher salaries have also grown, although in a somewhat complicated way. The increases in teacher salaries have not been uniform, as periods where salaries do not keep up with inflation (the decade of the 1970s) are offset by periods of more rapid increase (the decade of the 1980s). Moreover, salary growth for teachers has not kept up with growth in salaries in other occupations. Thus, while rising teacher wages have put cost pressure on schools, school salaries have also been competitive with a smaller proportion of outside jobs for college graduates over time.⁴

The top portion of Figure 1 identifies in a general way the third source of cost increases. Expenditures outside those for instructional staff have increased even more rapidly than those for aggregate instructional staff salaries. For example, between 1960 and 1990, salary expenditures fell from 61 percent to 46 percent of total current

³ Considerable recent attention has focused on expenditures and services for students with various mental and physical handicapping conditions. These services, commonly referred to as special education, have increased steadily since the mid-1970s, but cannot explain a majority of the overall growth in spending over the last two decades (Hanushek and Rivkin, 1997).

⁴ An additional complication is that the competitiveness of teaching salaries differs for men and women. Teaching has historically offered better relative salary opportunities to women than to men, but this is ending as outside opportunities for women expand. For men, teaching has kept up with outside wages over the last two decades. For women, the relative pay of teaching to alternative jobs has declined sharply over the same period, as their other opportunities have expanded (Hanushek and Rivkin, 1997).

expenditures. What underlies this change is unclear, because data on these expenditures are very poor. While it is often convenient to label this rise as “increased bureaucracy,” the available data neither confirm nor deny this interpretation, because these expenditures include a variety of items that are legitimately classroom expenditures—such as teacher health and retirement funds or purchases of books and supplies—in addition to administrative and other spending. The aggregate effects are clear, however: if these expenditures had grown between 1960 and 1990 at just the rate of instructional staff spending (which itself includes significant increases in resource intensity), total spending per pupil would have been 25 percent lower in 1990.

Three aspects of school expenditure deserve highlighting. First, real per pupil spending has increased steadily and dramatically. Commonly held views about how schools have struggled with a climate of fiscal stringency and financial pressure must be kept in a larger perspective of just what has been happening. As displayed in Figure 1 and provided for more recent years in Table 1, the nation has been running an aggressive school investment program for an extended period of time. Spending growth has outstripped income growth over the century, and this difference is particularly large in recent decades. Second, as Table 1 also shows, substantial increases in resources have been devoted to just the areas commonly advocated: pupil-teacher ratios have fallen dramatically (from 26:1 in 1960 to 17:1 in 1990); teacher education has risen to new heights (from 23 percent with a master’s degree in 1960 to 53 percent in 1990); and teacher experience has lengthened (from a median of 11 years in 1960 to 15 years in 1990). While the previously noted shift away from classroom salaries may be questioned, most advocates of school investment support smaller classes and better teaching training—just what has been happening. Third, the spending pattern and growth represent the outcomes of the existing political and decision-making system for schools. Absent changes in this environment, these historical patterns would seem to provide reasonable information on what we might expect in the future—in terms not only of natural growth but, more importantly, in terms of its effectiveness.

The central issue for this discussion is what we have gotten from this expenditure. The historical record on performance, however, is difficult to piece together. The clearest data come from the record for school attainment (U.S. Bureau of the Census, 1975; Goldin, 1994). At the turn of the twentieth century, barely 5 percent of the population graduated from high school. School attainment rose steadily, so that graduation rates exceeded 50 percent by 1940. What is less appreciated is that the growth in school completion essentially ceased during the 1960s, resulting in very stable levels of school attainment, at least since the mid-1970s. The median years of school completed for people aged 25–29 was 12.8 in 1975 and 13.0 in 1993 (U.S. Department of Education, 1995).⁵ The expansion of schooling for the labor

⁵ The change toward providing relatively more secondary education clearly contributed to the growth in per student spending over the century, although this would not be an important component of expenditure increases for the past quarter century. The precise magnitude of its effect on spending growth is nonetheless difficult to estimate.

Table 1
Public School Resources, 1961–1991

<i>Resource</i>	1960– 61	1965– 66	1970– 71	1975– 76	1980– 81	1985– 86	1990– 91
Pupil-Teacher Ratio	25.6	24.1	22.3	20.2	18.8	17.7	17.3
Percentage Teachers with Master's Degree	23.1	23.2	27.1	37.1	49.3	50.7	52.6
Median Years Teacher Experience	11	8	8	8	12	15	15
Current Expenditure/ADA (1992–93 dollars)	\$1,903	\$2,402	\$3,269	\$3,864	\$4,116	\$4,919	\$5,582

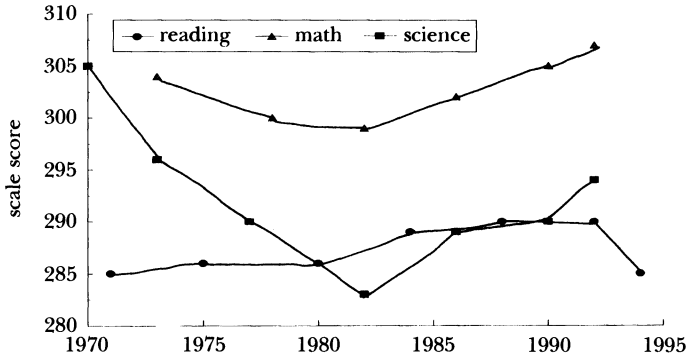
Source: U.S. Department of Education (1995).

force was undoubtedly an important part of the economic growth story in the United States (Denison, 1974; Romer, 1990; Jorgenson and Fraumeni, 1992; Goldin, 1994). The United States, leading the rest of the world in schooling investment throughout the century, became a model for other countries who wished to promote economic growth.

Information about quality changes in schooling over the long period is hard to come by. Some have suggested declines in average quality of graduates over time because as a larger group was drawn into schools, the school population would tend to include a higher proportion of less able students. Suspicions about some aspects of school quality were driven home by the Soviet launching of the Sputnik space vehicle in 1957, which raised doubts about whether U.S. education matched that of the Soviets or other countries. But no consistent evidence on quality is available until more recently.

One source of information about performance of a large segment of students comes from the college admissions examinations of the Scholastic Aptitude Test (SAT). Average SAT scores began falling sharply from the mid-1960s. There are legitimate concerns about the comparability of these scores over time, particularly for years before 1970. The main issue is that the proportion of students interested in admission to selective colleges expanded, which meant that on average, a lower caliber of student was probably taking the tests. But even after taking these changes into account, it appears that these average score changes do indicate real and significant performance declines (Congressional Budget Office, 1986).

Since 1970, consistent data for a random selection of U.S. students indicate that student performance has at best stayed constant, and it may have fallen. While aggregate performance measures are somewhat imprecise, all point to no significant gains in student performance over time. For example, consider the performance over time of a representative sample of 17-year-olds on the various

*Figure 2***NAEP Performance: 17-Year-Olds: Reading, Math, Science, 1970–1994**

components of the National Assessment of Educational Progress (NAEP) as summarized in Figure 2. Over the past quarter century, math performance is slightly up, reading is flat, and science is down (U.S. Department of Education, 1994). Optimists might point to rises in math and science performance over the 1980s, but at the very least it seems premature to conclude that there has been a general turnaround in student performance.

Recent comparisons of U.S. students with those from other countries have proved unflattering to America. Comparisons of United States and Japanese students in the early 1980s showed, for example, that only 5 percent of American students surpassed the average Japanese student in mathematics proficiency (McKnight et al., 1987; National Research Council, 1989). In 1991 comparisons, Korean 9-year-olds appeared closer to United States 13-year-olds than to U.S. 9-year-olds, hardly the kind of performance that would suggest that U.S. students will soon be top in the world in mathematics performance, as called for by legislated goals for U.S. schools (U.S. Department of Education, 1994). Careful study of international examinations reveals some bright spots, such as high relative performance of U.S. students on recent international literacy examinations (U.S. Department of Education, 1996). Nevertheless, a more complete comparison of performance on international mathematics and science examinations not only places U.S. performance below the average of the 39 countries that have participated in such examinations over the past quarter century, but also shows that performance of a country's population on these more analytical tests has direct and strong implications for economic growth (Hanushek and Kim, 1995).

The problems of performance appear particularly acute when considered by race or socioeconomic status. Although there has been noticeable narrowing of the differences in performance, the remaining disparities are huge, and

incompatible with most notions of equity. During the 1980s, there was a broad-based convergence of black-white score differences. While the gap in science achievement of 17-year-olds has closed little and remained about one standard deviation, the 1.3 standard deviation gaps in mathematics and reading each closed by about 60 percent.⁶ The most recent data suggest that convergence may have ceased, with the NAEP reading scores, for example, showing significant recent widening.

The aggregate story of spending and performance seems quite clear over the past quarter century. Increases in real expenditure per pupil have been steady and large, amounting to over a 70 percent increase between 1970 and 1990. School attainment over this period has been constant. Quality, as best we can measure it, has been flat. This provides a strong *prima facie* case for the ineffectiveness of pure resource policies. Of course, ambiguities abound in the aggregate time series data, because many things have changed over time, and it is difficult to separate the various possible influences on costs and performance. For example, one explanation is that the student population has become more difficult to educate over time, say, because of increases in divorce rates, child poverty and, perhaps, female labor force participation rates. Of course, there are countertrends that would tend to make students easier to educate, such as the fact that the average parent has a higher level of education and the average family size is smaller. It is difficult to say *a priori* how these factors balance out in terms of aggregate student performance.⁷ Thus, support from microlevel analyses is helpful in solidifying the interpretation of the aggregate data.

Basic Micro Evidence

The core of the relevant micro evidence comes from attempts to estimate production functions for education. This work began seriously with the Coleman report (Coleman et al., 1966). The Coleman report, a response to the Civil Rights Act of

⁶ Robert Hauser kindly provided the black-white trend data.

⁷ Grissmer et al. (1994) suggest that on net, families have actually gotten *better* in educational terms, but these conclusions have considerable uncertainty attached to them. They result from applying aggregate changes in family characteristics to the estimated importance of individual family factors derived from cross-sectional models of student achievement. These cross-sectional models concentrate on family characteristics but ignore any direct schooling effects. This analysis suggests that actual student performance for the population and for whites is not as high as would be expected, based on family improvements. For blacks, the opposite holds, raising the possibility that schools for blacks have been responsible for their better-than-expected gains, but the opposite for the majority of the student population. The authors attribute the difference between actual and predicted performance to schools, although there is no explicit analysis of this. Direct analysis of the closing performance gap of blacks and whites on the NAEP nonetheless provides little support that the relative gains are the result of improved school-level resources (Cook, 1995). The analysis in Congressional Budget Office (1987) provides details of information about test score convergence by race through the mid-1980s. None of these incorporate the recent divergence in NAEP reading scores in the 1990s.

1964, was a mandated study of inequality in the provision of education. That work attracted the attention of policymakers and researchers because of its controversial conclusions that schools seemed relatively unimportant in determining student achievement, while families were the key element of student success. Its most important scholarly contribution was to focus attention on how various inputs to education—schools, parents, peers and others—combined to affect student performance. The ensuing research went in a variety of directions, but an important and continuing strand focused on how different resources supplied to schools translated into student outcomes. This research provides a strong case supplementing the aggregate data that differences in resources are not very closely related to student performance.

A compilation of results of this estimation provides little confidence that adding more resources to schools as currently operating is likely to boost student achievement. The overall results are summarized in Table 2. This summary table reports results from production function estimates of the effects of resources that were published through 1994. A total of 377 separate estimates, published in 90 separate articles and books, contribute to the knowledge base about the effects of resources.⁸ These studies combine estimates of the effect of resources on a range of different outcome measures in different grades and regions of the country. About three-quarters employ standard achievement tests as the measure of student performance, while the remainder combine a variety of objective measures including dropout status, continuation into college and wages when in the labor force. All studies include some measure of family background, and all had to provide information about the statistical significance of any estimates to be included.

The most reliable estimates come for the real resources that are the prime determinants of variations in spending per pupil: teacher-pupil ratio, teacher education and teacher experience. Teacher education and teacher experience are the primary determinants of variations in teacher salaries and, when combined with the number of pupils for each teacher, indicate how instructional spending per pupil varies. These data are frequently available in databases that supply background and performance information for individual students, providing a solid basis for estimation of achievement relationships.

The results in Table 2 are now reasonably well known: the primary resources for schools are not consistently related to student performance. Consider teacher-pupil ratios. Of the 277 total estimates available, only 15 percent find a positive and statistically significant relationship with student achievement, and 13 percent find a negative and statistically significant relationship. The insignificant results are also almost evenly split between positive and negative, although 20 percent do not provide information about the sign of insignificant results. Even fewer estimates of the effects of teacher education are positive and significant (9 percent of the 171 total

⁸ See Hanushek (1996a) for a description of these studies along with more detailed compilations of the results than those presented here.

Table 2
Percentage Distribution of Estimated Effect of Key Resources on Student Performance, Based on 377 Studies

<i>Resources</i>	<i>Number of Estimates</i>	<i>Statistically Significant</i>		<i>Statistically Insignificant</i>		<i>Unknown Sign</i>
		<i>Positive</i>	<i>Negative</i>	<i>Positive</i>	<i>Negative</i>	
Teacher-Pupil Ratio	277	15%	13%	27%	25%	20%
Teacher Education	171	9	5	33	27	26
Teacher Experience	207	29	5	30	24	12
Teacher Salary	119	20	7	25	20	28
Expenditure Per Pupil	163	27	7	34	19	13
Administrative Inputs	75	12	5	23	28	32
Facilities	91	9	5	23	19	44

Source: Author's tabulations.

estimates). Teacher experience tends to have more positive and more statistically significant results, but 70 percent of the estimates still are either negative or statistically insignificant at the 5 percent level. (There is also ambiguity about the direction of causation, since more senior teachers frequently can choose their schools and even their classes.)

It is important to put these results in the context of the rising inputs to education documented in Table 1. The steady and large expansion of resources devoted to schools has come precisely in areas that are quite unrelated to student performance. Thus, it is not surprising to find that aggregate spending trends have nothing to do with aggregate student performance. Moreover, it calls into question the commonly espoused resource solutions to schooling problems.

The remainder of Table 2 presents the estimates for other resource measures. The pattern of these tends to mirror that for the primary real resources in that there are no consistent impacts on student performance. Each of these inputs, however, is measured much more imprecisely. To begin with, these latter resource measures are seldom available for classrooms or even individual schools but instead are aggregate measures for the district or even the state.⁹ (The estimates of the effects of expenditure per pupil are especially likely to be drawn from highly aggregated studies, and, as shown below, this aggregation has a dramatic impact on the overall results.) These resource measures are more difficult to interpret, because they tend to mix together a variety of specific measures. For example, the category of facilities includes financial measures (such as the value of physical plant)

⁹ As discussed below, some of the estimates of the effects of the primary resource measures also come from aggregated data. This aggregation, particularly when coupled with specification errors, is likely to lead to substantial biases in the estimates.

and a wide variety of physical measures, such as characteristics of libraries, science labs, or building and grounds. Nonetheless, these further results add little confidence that resource differences will affect achievement.

These results taken as a whole are consistent with resources sometimes being used effectively, but most of the time not. Knowing that added resources are sometimes effective is not very useful, because none of the research is able to characterize when resources are used effectively and when they are not. The results simply demonstrate that the current operations of schools do not generate consistent improvements in student performance.

While no consensus exists about what specific factors affect student performance, there is overwhelming evidence that some teachers and schools are significantly better than others. The important distinction is between whether measured characteristics of the sort that are the subject of educational policy and that appear in econometric studies are important, as opposed to true underlying differences—even if these true differences cannot be specified and measured in terms of observable characteristics or traits. For example, within one inner-city school system serving an entirely black population, a good teacher was found to surpass a bad teacher by more than a full grade level of student achievement over a single academic year, even after holding constant the family characteristics of students and the level of achievement at which they started the class (Hanushek, 1992). Such dramatic differences in performance are simply not determined by the training of teachers, the number of students in the classroom, the salaries of teachers, or the overall level of spending. A primary task of school reform is increasing the likelihood that a student ends up in a high-learning environment.

The general policies of past, even if dressed up in new clothing, are unlikely to lead to student performance gains, even though cost pressures will continue to mount. While it may be appropriate to increase spending on schools in the future, the first priority is restructuring how existing resources are being used. The challenge is to identify the characteristics that make certain schools and teachers better, remembering that these characteristics don't seem to have much to do with standard inputs, and then figure out how worse schools can take on these positive characteristics.

Alternative Views of the Importance of Pure Resources

Before considering possible restructuring approaches, alternative views of this evidence need to be mentioned.

In some disciplines, like figuring out the health effects of a certain drug therapy, there is frequently an interest in compiling results from a variety of trials. Specialized techniques to combine the results of separate studies and thus assess the magnitude and significance of some relation have been developed. These approaches go under the general title of "meta-analysis." Clearly, the preferred approach to assessing disparate results would involve combining the underlying data

of the studies to develop statistical inferences and tests of hypotheses across the studies. Unfortunately the original data are seldom available for reanalysis—and even when they are, combining data from different sources can be difficult—which forces a variety of compromises in the aggregation of results. The previous data on studies in Table 2 is one form of aggregation of results, which relies on the minimal set of factors standardly reported.

However, using a subset of the studies of estimated resource effects in Table 2, Hedges, Laine and Greenwald (1994) have applied other statistical approaches to the prior results. Specifically, their analysis develops a series of statistical tests of the hypothesis that all of the underlying resource parameters relevant for the basic studies are simultaneously zero. In essence, they rely on the fact that, if the underlying resource parameters were all identically zero, an unbiased statistical test would find 2.5 percent of the estimates to be statistically significant in a negative and 2.5 percent in a positive direction (for a 5 percent test).¹⁰ The results in Table 2, say for teacher-pupil ratios, find noticeably more rejections than this. In such a case, their combined significance testing is designed to reject the null hypothesis that all underlying parameters are zero, which it does.

These results are sometimes interpreted as a refutation of the conclusion that educational inputs don't affect performance. But in my view, this work both confirms the previous substantive results and points to the same policy conclusions. The formal statistical testing of Hedges, Laine and Greenwald (1994) really emphasizes that instead of thinking of a single common resource parameter—like that for the effect of more spending per student—it is best to think of an underlying distribution of resource parameters. This focuses attention on the need for an appropriate structure of the educational environment to ensure that added resources deliver positive effects. It shows that productive results are possible, even if seldom achieved currently. Moreover, it highlights the inappropriateness of simple resource policies within the context of current schools.

Taken as a group, the production function studies give little indication that variations of resources have anything to do with present variations in student performance. However, the widely publicized findings of Card and Krueger (1992) indicate variations in school resources are related to earnings differences among workers.¹¹ Several issues could contribute to reconciling these conclusions: differences in levels of resources considered; differences in measurement of student

¹⁰ An alternative explanation of the data is that the underlying distributional assumptions for the statistical tests are incorrect. This could occur if the underlying estimates were biased because of specification errors or if there were a publication bias leading more statistically significant results to be published than were appropriate from the underlying data. There are a series of assumptions and manipulations that are required for their formal testing and that do make interpretation of the formal tests problematic, but they are not important for the considerations here. See Hanushek (1994, 1996b).

¹¹ The Card and Krueger (1992) analysis of school resources and earnings is the most discussed, but it follows a larger line of research. See, for example, Welch (1966), Johnson and Stafford (1973) and Wachtel (1976). An insightful review of past studies that considers underlying characteristics of the studies is Betts (1996).

performance; differences in specification; and aggregation bias in the statistical analysis.

The Card and Krueger (1992) analysis begins with a sample of adult workers from the 1980 Census of Population and fills in information about the schooling circumstances of individuals from information about their year and state of birth. The workers in their sample attended schools between the 1920s and the 1970s, implying variations in the level of resources going far beyond what is found today. This suggests one reconciliation: if added resources have diminishing effects on student achievement, current school operations may be largely “on the flat” of the production function, while Card and Krueger observe ranges from the past where resources had stronger effects.¹² A related possibility might be that the political economy of schools has changed over time. For example, with the rise of teachers’ unions and the resulting change in bargaining positions, resources might be used in different ways and have different implications for student achievement now than in the past. In other words, it is quite possible that the enormous changes in educational resources did have an effect on outcomes in the first half of this century, but that more recent studies are also correct in finding “no effect” for the sorts of resource changes discussed in current schools.

A second suggested reconciliation revolves around the measurement of outcomes. The previously compiled production function estimates are heavily weighted toward analyses of standardized test scores, while the Card-Krueger analysis concentrates on labor market earnings. It is possible that schools do not affect test performance of students, but do affect skills and earnings. The previous conclusions from production function estimates in Table 2, however, hold equally when results are divided between studies that use test scores as measure of outcomes and other measures of outcomes like college continuation or earnings (Hanushek, Rivkin and Taylor, 1996). Also, as Burtless (1996) points out, it seems implausible that schools do not affect what they explicitly are attempting to do (improve test performance) but do affect earnings, something they seldom measure or even consider a direct objective.

Moreover, there is considerable evidence that test scores are increasingly related to labor market performance (O’Neill, 1990; Bishop, 1991; Grogger and Eide, 1993; Murnane, Willett and Levy, 1995; Neal and Johnson, 1996). It seems unlikely that school resources affect just the component of earnings that is uncorrelated with cognitive skills. Moreover, school resources are not consistently related to earnings (Betts, 1996). This finding is particularly clear when direct measures of the school resources relevant to individuals are available (Betts, 1995; Grogger, 1996). As an overall summary, the lack of relationship with school resources is more generally true for recent studies of earnings than for earlier investigations, while more recent studies have tended to find stronger effects of cognitive skills on earnings.

¹² While not a direct test of this on-the-flat thesis, the lack of significantly stronger resource effects in developing countries introduces some question about this hypothesis; see Hanushek (1995) or, in a growth context, Hanushek and Kim (1995).

The final set of reasons that could help to explain the different conclusions involve specification issues. To begin with, many of the direct analyses of earnings include just the level of school resources, but none of the other factors that might influence student achievement and skill development. For example, it is plausible that students attending schools with high levels of resources also have parents who contribute more time, energy and money to their education. If parental inputs are left out of the calculation, any estimated effects of school resources would tend to overstate the true independent effect of resources.¹³

Aggregation of school inputs is also likely to exacerbate any biases due to specification issues (Hanushek, Rivkin and Taylor, 1996). Most of the earnings analyses observe school resources measured only at the aggregate state level. A key concern in the study of schools is the autonomy of individual states in setting regulations and policies, in providing financial support and incentives for schools, and in overall organization of schools and labor laws. Virtually no studies incorporate any measures of the state policy environment, leading to potentially severe specification problems in studies that draw data from different states and policy environments. If states that provide a higher level of funding also tend to have more productive policy environments, then a regression analysis that doesn't control for the policy environment will tend to exaggerate the effect of funding on performance.¹⁴ Table 3 shows the previous production function estimates for teacher-pupil ratio and expenditure per student, arrayed according to whether the input data are collected at the state level, or at a lower level such as district or classroom, and also whether the observed schools are drawn across different states. (If all data are drawn from a single state, each school in the state faces the same state policy environment, and this particular specification bias won't be an issue.) The table shows that positive and statistically significant resource estimates for both teacher-pupil ratio and school spending come disproportionately from estimates employing interstate data and are especially heavily weighted in estimates relying on state aggregates. The Card-Krueger estimates come from resource data aggregated to the state level, but no measures of state policy differences are included, so their estimates are subject to this bias.

The end result of this comparison is that the estimates of Card and Krueger (1992) at most suggest that very low levels of resources—say those found in the

¹³ An additional specification issue is that Card and Krueger (1992) attempt to distinguish between the effects of schooling inputs and the effects of being in different local labor markets by assuming that migration across regions is nonselective. This assumption, however, runs counter to standard economic models and, as Heckman, Layne-Farrar and Todd (1996) demonstrate, counter to the data. Thus, the data do not support a key identifying condition for the Card-Krueger estimation of school resource effects.

¹⁴ There is no a priori reason to believe that omitted state factors lead to either an upward or downward bias in the estimated effects of school resources. In addition to the summary of production function studies below, Hanushek, Rivkin and Taylor (1996) present new empirical evidence from the High School and Beyond data on school achievement and college continuation that supports the general upward bias. That paper also demonstrates that the aggregation effect is not one due to correcting any measurement errors.

Table 3

Percentage of Estimated Effect of Teacher-Pupil Ratio and Expenditure per Pupil by State Sampling Scheme and Aggregation

State Sampling Scheme and Aggregation of Resource Measures	Number of Estimates	Statistically Significant		Statistically Insignificant		Unknown Sign
		Positive	Negative	Positive	Negative	
<i>Teacher-Pupil Ratio</i>						
Total	277	15%	13%	27%	25%	20%
Single State Samples ^a	157	12	18	31	31	8
Multiple State Samples ^b	120	18	8	21	18	35
With Within-State Variation ^c	109	14	8	20	19	39
Without Within-State Variation ^d	11	64	0	27	9	0
<i>Expenditure Per Pupil</i>						
Total	163	27	7	34	19	13
Single State Samples ^a	89	20	11	30	26	12
Multiple State Samples ^b	74	35	1	39	11	14
With Within-State Variation ^c	46	17	0	43	18	22
Without Within-State Variation ^d	28	64	4	32	0	0

Notes: Rows may not add to 100 because of rounding.

^a Estimates from samples drawn within single states.

^b Estimates from samples drawn across multiple states.

^c Resource measures at level of classroom, school, district, or county, allowing for variation within each state.

^d Resource measures aggregated to state level with no variation within each state.

Source: Hanushek, Rivkin and Taylor (1996).

poorest states before and during the Great Depression—may have an effect on student outcomes. But there is little reason to believe that this conclusion offers helpful policy advice given the current levels of resources.

Policy Implications

Because much of the concern about the effects of school resources is policy motivated, it is useful to sketch some of the policy implications of these results. One immediate implication of these results is that spending per pupil is not a good index of school quality. Thus, policies that key on the spending level of schools should not be interpreted as having anything to do with quality. The clearest set of such policies is the funding of schools. A majority of states have seen court suits focused on the distribution of state aid and whether existing formulae are equitable. A more recent version of these questions is whether the spending level is sufficient,

or adequate. Similarly, much of the discussion about education by state legislatures revolves around pure funding issues. While these discussions have distributional implications for taxes and subsidies, they have little to do with student achievement and educational outcomes. The same issue arises in political rhetoric about the importance of improving the nation's human capital, which invariably turns to a consideration of direct expenditures for schooling. The level of spending proves to be a terrible measure of investment in schooling, and thus the policy focus on spending or its correlates is most often badly misplaced.

The lack of overall relationship between resources and performance surprises many people, but it perhaps should not. The most startling feature of schools—a feature distinguishing schools from more successful parts of our economy—is that rewards are only vaguely associated with performance, if at all. A teacher who produces exceptionally large gains in student performance generally sees relatively little difference in compensation, career advancement, job status, or general recognition when compared with a teacher who produces exceptionally small gains. A superintendent who provides similar student achievement to that in the past while spending less is unlikely to be rewarded above what would be the case for spending the same or more.¹⁵ If few incentives exist to reward improved performance, it should not be surprising to find that resources are not systematically used in a fashion that improves performance.

The evidence for lack of a clear connection between educational resources and performance arises within the existing structure and operating procedures of schools.¹⁶ A different organizational structure with different incentives could produce very different results. For example, almost every economist would support the position that increasing teacher salaries would expand and improve the pool of potential teachers. However, whether this improves the quality of teaching depends on whether schools can systematically choose and retain the best teachers from the pool (Ballou and Podgursky, 1995). If schools faced a greater incentive to produce student achievement and mechanisms for teacher selection were altered, then perhaps altering teacher salaries would have significant effects on student learning. In other words, there seems little question that money could count. But given the current organization of schools, it just does not systematically count.

These results are entirely compatible with some schools using additional funds effectively—and that others do not. But within the current structure, good uses of funds are balanced by bad uses. Unless some way is found to change the districts that would squander additional funds into districts that would use them effectively, added resources in the aggregate are not likely to lead to any improvement in average performance. No one of course advocates wasting additional funds, and there is always a long list of potential “innovations” that are touted as breaking

¹⁵ As discussed below, the lack of incentives is not restricted just to school personnel. A student who gets high grades is not necessarily going to be rewarded by employment over the student with low grades (Bishop, 1991).

¹⁶ The discussion in this section draws upon *Making Schools Work* (Hanushek and others, 1994).

with past policies. Past innovations advertised with the same promise and conviction have nonetheless produced the historic ineffectiveness.

The suggestions for educational reform in the face of this evidence are actually very rudimentary, the kind of things that an undergraduate major in economics might be expected to suggest. These reforms are notable mostly for their complete absence from current discussions over education policy.

Almost certainly, the primary policy instrument must be improved incentives for performance. Public schooling is one of the largest remaining bastions of centralized, regulatory approaches to providing services. Education is a complicated task that requires the cooperation and ingenuity of teachers, principals and other school personnel. Research suggests that many equally effective approaches exist for learning various subjects and skills, differentiated only by how individual teachers and students adapt to specific tactics and techniques. Because there is no single best approach to performing specific educational tasks, it is not possible to design policies that are based on full descriptions of what is to be done and how it is to be done in the classroom. Incentives based upon student outcomes hold the largest hope for improving schools. This idea is radically different from most past policy, which has been based on a combination of regulations and fixed definitions of inputs to schooling: the resources, organization, and structure of schools and classrooms. Improvement seems more likely if policies are built on what students actually do and if good performance by students gets rewarded.

Simply saying "performance incentives," however, is easier than implementing incentives that have the desired outcomes. While many options have been vigorously discussed over a long period of time, ranging from merit pay for teachers to private contracting to vouchers, little experience about the details or the effects of these proposals has accumulated. Information about different incentives could come from a program of actual experimentation or from evaluations of various programs that are instituted around the country. However, neither approach is traditionally followed in education. The number of large-scale experiments with random assignment is small, and even though the results have been subject to serious debate, no follow-up experiments typically ensue.¹⁷ Evaluation of continuing or new programs is also surprisingly infrequent.

Finally, virtually all past considerations of school reform have either ignored costs or argued that the benefits are large enough to support any proposed increased costs. Downplaying costs tends to lead to proposals with higher costs. This in turn undoubtedly lowers the likelihood that proposals will be taken seriously,

¹⁷ A 1970 experiment with a faulty contracting structure provides most of our information about performance contracts with private firms (Gramlich and Koshel, 1975), although additional information from current demonstrations may soon be available. The effects of altered class sizes were the subject of a 1980s experiment in Tennessee (Word et al., 1990), but the design did not permit observing the outcomes of students who change from small to large classes, and there has been no comparable experiment since. Consideration of the effects of vouchers has been part of a small-scale and restricted demonstration program in Milwaukee, leading to controversial conclusions (Witte, Bailey and Thorn, 1993).

because policymakers and the public will certainly consider the price tag attached to any major restructuring of schools. Fundamentally, attention to both costs and benefits should not be restricted just to new programs. Many existing programs are inefficient and should be replaced by more cost-efficient programs. A more focused consideration of both costs and benefits is necessary before the cycle of increasing expenditure without improvements in student performance can be broken.

These ideas almost come down to a systematic application of common sense from an economist's viewpoint: introduce performance incentives into schools; experiment and evaluate alternatives; and compare costs of alternative approaches. They do underscore, however, that making better use of existing resources should receive a very high priority, as opposed to simply expanding existing resources on the notion that more spending will automatically bring rewards.

A Note on Implications for Economic Research

Many lines of economic research depend on having accurate measures of individual human capital, which in turn require measuring investment in education and other inputs for human capital. Measurement problems are particularly acute in the case of human capital. There is a natural parallel to earlier discussions about the measurement of physical capital, where much of the controversy revolved around the use of input and output measures in the face of quality differences (for example, Denison, 1972; Jorgenson and Griliches, 1972). The measurement of physical capital stock and services has been largely an issue of macroeconomics, keyed to questions of productivity change and economic growth. This also motivates the extension of output-based measures of human capital developed by Jorgenson and Fraumeni (1992). However, many applications of the concept of human capital relate to micro phenomena and require detailed measures by individuals or particular groups. This altered purpose by itself requires very different measurement strategies, typically making use of output proxies such as achievement test scores rather than aggregated market observations.

The preceding discussion in this paper, however, suggests two key issues that extend the measurement discussions into even more difficult territory.¹⁸ First, inefficiency in production of human capital (like that described previously) introduces natural measurement problems, because direct spending is no longer linked to a particular level of output. Second, it is well known that families have considerable influence on a student's achievement, implying that school resources are

¹⁸ One additional measurement issue has been discussed previously but is ignored here. Because of the dependence of schools on specialized labor inputs, the consumer price index or GDP deflator might not accurately reflect changes in input costs and thus real resource changes might not be accurately described over time if these price indices are used. For a discussion of these issues, which revolve importantly around questions of the particular purpose of cost adjustment, see Hanushek and Rivkin (1997).

only part of the inputs into human capital and, regardless of the efficiency of schools, cannot adequately capture human capital. Both factors suggest that employing input measures of school resources, while frequently an expedient approach to measuring human capital, could significantly distort any analysis.

There is wide recognition that differences in skills among individuals—and more broadly, the concept of human capital—are important not only for explaining earnings differences across individuals but also for investigating informational differences, managerial decision making, health investments and a wide variety of other behavioral issues. In general, empirical consideration of human capital effects has relied simply on differences among individuals in the quantity of schooling attained, leaving open the question of whether qualitative differences are important. Studies of wage determination, as discussed above, have investigated both input and output measures of schooling differences. The conclusion previously emphasized is that input measures are not a good proxy for school quality because of the inefficiency of schools. But there is also a larger problem that arises from imputing any causal interpretation when school resources appear to affect subsequent outcomes. School quality and individual skills are simply not synonymous. Since student achievement and variations in human capital across individuals result from family and peer influences in addition to school influences, analyses that neglect the other inputs will generally be biased. In fact, the bias will often be severe because family influences are very important in determining student achievement and because family influences tend to be positively related to measured school resources (which by themselves do not appear to be important). The introduction of differences in human capital, beyond quantity of formal schooling, into areas of research outside of wage determination is likely to occur more frequently, both because of increased recognition of qualitative differences and because of expanded data availability. A running message of this paper is that caution is required in both model specification and interpretation.

Investigations involving individual behavior require measures of the individual's human capital, but at other times the focus of attention is directly on school quality. Obvious examples—which have relied on spending as the measure of school quality—include investigations of the demand for local public schools (for example, Epple and Romer, 1991; Fernandez and Rogerson, 1996), of the effects of schools on housing prices (Oates, 1969), and of the equity of state support for local schools (Berne and Stiefel, 1984). These studies and others reflect the reality that citizens are interested in school spending, both because citizens care about school quality and because they care about taxes and governmental spending. The inefficiency of schools suggests a reinterpretation of such studies.¹⁹ On the one hand, such studies appear to provide insights into the positive aspects of school

¹⁹ Inefficiency concentrates on the investment and human capital aspects of spending. Spending that simply reflects consumption aspects—like fancy swimming pools or private music lessons—is not considered here, even though that might partly explain the character of parental demands for schools and the variation across schools in spending.

spending, reflecting issues like individual behavior in demand, residential choice, and the like. On the other hand, they appear to give a distorted view of any welfare implications that rely on an assumption that spending adequately measures school quality. A more limited set of studies has investigated how student achievement enters into demand or capitalization (Rosen and Fullerton, 1977). But here, too, caution is required. If one were interested in the effects of local school quality on housing prices, using average test scores as the measure of quality would confuse the value added of schools and the impacts of families on performance, and thus would also give a misleading answer to the question being investigated.

Sometimes the central issue is the value added of schools; sometimes it is overall performance levels. In either event, school resource measures are unlikely to suffice. The measurement issues flow from the character of the determinants of student achievement and the nature of public provision of school services. Considerably more attention must be given to the direct measurement of student skills and school performance than has typically been the case. Because it frequently complicates any analysis, this advice may not be what researchers looking into a variety of issues want to hear, but it is difficult to avoid.

Conclusions

The convenience of measuring educational investment in terms of resources or spending is undeniable. Yet the evidence suggests that spending and commonly used resources of schools are not good measures of quality. Moreover, simply adding more resources to schools as currently structured is unlikely to yield significant improvement in student performance. Many people recognize that past spending on schools has brought little reward, but they advocate still another round of investment. This is often coupled by mention of an approach or program that receives the current certification of educators. Such optimism about new and largely unexamined proposals are part of what lies behind a century of 3.5 percent annual growth in real spending per pupil, spending that has left student performance flat or declining over the past quarter century.

Economists should not be particularly surprised by these results. The public provision of schooling with minimal competition and with virtually no performance incentives for people in the system is not geared to efficient use of resources. These aspects also dictate that things are unlikely to change unless a fundamental restructuring is undertaken so as to introduce performance incentives of one sort or another into the school system. Input policies are simply an ineffective substitute for concentration on desired student outcomes.

Schools may currently be facing a different policy environment than that of the past quarter century. The falling student populations of the 1970s and 1980s held the growth in overall school budgets below the growth in per pupil spending and perhaps made expansion of school resources more palatable to the public. But the student population is again rising, putting added pressure on school budgets.

The result in the early 1990s has been a dramatic shift in spending trends to slow or no growth in spending per pupil. If this also reflects a growing public awareness of and dissatisfaction with trends in student performance, serious restructuring of schools might be a political prerequisite for relief from current fiscal stringency.

■ *This paper was partially supported by a grant from the William H. Donner Foundation. It benefitted from helpful comments and suggestions by Stanley Engerman, Robert Hauser, Fran Blau, Brad De Long and Timothy Taylor, although none is unduly implicated in the final version.*

References

- Ballou, Dale, and Michael Podgursky**, "Recruiting Smarter Teachers," *Journal of Human Resources*, Spring 1995, 30:2, 326-38.
- Berne, Robert, and Leanna Stiefel**, *The Measurement of Equity in School Finance: Conceptual, Methodological, and Empirical Dimensions*. Baltimore, Md.: Johns Hopkins University Press, 1984.
- Betts, Julian R.**, "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth," *Review of Economics and Statistics*, May 1995, 77:2, 231-47.
- Betts, Julian R.**, "Is There a Link Between School Inputs and Earnings? Fresh Scrutiny of an Old Literature." In Burtless, Gary, ed., *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. Washington, D.C.: Brookings Institution, 1996, pp. 141-91.
- Bishop, John**, "Achievement, Test Scores, and Relative Wages." In Kosters, Marvin H., ed., *Workers and Their Wages*. Washington, D.C.: AEI Press, 1991, pp. 146-86.
- Bishop, John**, "Signaling, Incentives, and School Organization in France, the Netherlands, Britain, and United States." In Hanushek, Eric A., and Dale W. Jorgenson, eds., *Improving American Schools: The Role of Incentives*. Washington, D.C.: National Academy Press, forthcoming 1996.
- Burtless, Gary**, "Introduction and Summary." In Burtless, Gary, ed., *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. Washington, D.C.: Brookings Institution, 1996, pp. 1-42.
- Card, David, and Alan B. Krueger**, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, February 1992, 100:1, 1-40.
- Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York**, *Equality of Educational Opportunity*. Washington, D.C.: U.S. Government Printing Office, 1966.
- Congressional Budget Office**, *Trends in Educational Achievement*. Washington, D.C.: Congressional Budget Office, 1986.
- Congressional Budget Office**, *Educational Achievement: Explanations and Implications of Recent Trends*. Washington, D.C.: Congressional Budget Office, 1987.
- Cook, Mike**, "Families or Schools?: Accounting for the Convergence in White and Black Test Scores since 1970," paper presented at meeting of Association for Public Policy Analysis and Management, Washington, D.C., November 1995.
- Denison, Edward F.**, "Final Comments," *Survey of Current Business*, May 1972, 52:5, Pt. 2, 95-110.
- Denison, Edward F.**, *Accounting for United States Economic Growth, 1929-1969*. Washington, D.C.: Brookings Institution, 1974.
- Epple, Dennis, and Thomas Romer**, "Mobility and Redistribution," *Journal of Political Economy*, August 1991, 99, 828-58.
- Fernandez, Raquel, and Richard Rogerson**, "Income Distribution, Communities, and the

Quality of Public Education," *Quarterly Journal of Economics*, February 1996, 111:1, 135-64.

Goldin, Claudia, "How America Graduated from High School, 1910 to 1960." Working Paper No. 1994, National Bureau of Economic Research, June 1994.

Gramlich, Edward M., and Patricia P. Koshel, *Educational Performance Contracting*. Washington, D.C.: Brookings Institution, 1975.

Grissmer, David W., Sheila Nataraj Kirby, Mark Berends, and Stephanie Williamson, *Student Achievement and the Changing American Family*. Santa Monica, Calif.: Rand Corporation, 1994.

Grogger, Jeffrey T., "Does School Quality Explain the Recent Black/White Wage Trend?," *Journal of Labor Economics*, April 1996, 14:2, 231-53.

Grogger, Jeffrey T., and Eric Eide, "Changes in College Skills and the Rise in the College Wage Premium," *Journal of Human Resources*, Spring 1993, 30:2, 280-310.

Hanushek, Eric A., "The Trade-Off Between Child Quantity and Quality," *Journal of Political Economy*, February 1992, 100:1, 84-117.

Hanushek, Eric A., "Money Might Matter Somewhere: A Response to Hedges, Laine, and Greenwald," *Educational Researcher*, May 1994, 23:4, 5-8.

Hanushek, Eric A., "Interpreting Recent Research on Schooling in Developing Countries," *World Bank Research Observer*, August 1995, 10:2, 227-46.

Hanushek, Eric A., "Assessing the Effects of School Resources on Student Performance: An Update." Working Paper No. 424, Rochester Center for Economics Research, University of Rochester, 1996a.

Hanushek, Eric A., "A More Complete Picture of School Resource Policies," *Review of Educational Research*, Fall 1996b, forthcoming.

Hanushek, Eric A., and Dongwook Kim, "Schooling, Labor Force Quality, and Economic Growth." Working Paper No. 5399, National Bureau of Economic Research, December 1995.

Hanushek, Eric A., and Steven G. Rivkin, "Understanding the 20th Century Growth in U.S. School Spending," *Journal of Human Resources*, forthcoming, Winter 1997.

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor, "Aggregation and the Estimated Effects of School Resources," *Review of Economics and Statistics*, November 1996, 78:4, forthcoming.

Hanushek, Eric A., and others, *Making Schools Work: Improving Performance and Controlling Costs*. Washington, D.C.: Brookings Institution, 1994.

Heckman, James S., Anne Layne-Farrar, and Petra Todd, "Does Measured School Quality Re-

ally Matter? An Examination of the Earnings-Quality Relationship." In Burtless, Gary, ed., *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. Washington, D.C.: Brookings Institution, 1996, pp. 192-289.

Hedges, Larry V., Richard D. Laine, and Rob Greenwald, "Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes," *Educational Researcher*, April 1994, 23:3, 5-14.

Johnson, George E., and Frank P. Stafford, "Social Returns to Quantity and Quality of Schooling," *Journal of Human Resources*, Spring 1973, 8:2, 139-55.

Jorgenson, Dale W., and Barbara M. Fraumeni, "Investment in Education and U.S. Economic Growth," *Scandinavian Journal of Economics*, 1992, 94, Supplement, 51-70.

Jorgenson, Dale W., and Zvi Griliches, "Issues in Growth Accounting: A Reply to Edward F. Denison," *Survey of Current Business*, May 1972, 52:5, Pt. 2, 65-94.

McKnight, Curtis C., F. J. Crosswhite, J. A. Dossey, E. Kifer, J. O. Swafford, Ken J. Travers, and T. J. Cooney, *The Underachieving Curriculum: Assessing U.S. School Mathematics from an International Perspective*. Champaign, Ill.: Stipes Publishing Co., 1987.

Murnane, Richard J., John B. Willett, and Frank Levy, "The Growing Importance of Cognitive Skills in Wage Determination," *Review of Economics and Statistics*, May 1995, 77:2, 251-66.

National Research Council, *Everybody Counts: A Report to the Nation on the Future of Mathematics Education*. Washington, D.C.: National Academy Press, 1989.

Neal, Derek A., and William R. Johnson, "The Role of Pre-Market Factors in Black-White Differences," *Journal of Political Economy*, October 1996, 104:5, 869-95.

Oates, Wallace, "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis," *Journal of Political Economy*, November/December 1969, 77:6, 957-71.

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

Romer, Paul, "Human Capital and Growth: Theory and Evidence," *Carnegie-Rochester Conference Series on Public Policy*, 1990, 32, 251-86.

Rosen, Harvey S., and D. J. Fullerton, "A Note on Local Tax Rates, Public Benefit Levels and Property Values," *Journal of Political Economy*, April 1977, 85:2, 433-40.

U.S. Bureau of the Census, *Historical Statistics*

of the United States, *Colonial Times to 1970, Bicentennial Edition*. Vols. 1 and 2, Washington, D.C.: U.S. Government Printing Office, 1975.

U.S. Department of Education, *The Condition of Education, 1994*. Washington, D.C.: National Center for Education Statistics, 1994.

U.S. Department of Education, *Digest of Education Statistics*. 1995 ed., Washington, D.C.: National Center for Education Statistics, 1995.

U.S. Department of Education, *Reading Literacy in the United States: Findings from the IEA Reading Literacy Study*. Washington, D.C.: National Center for Education Statistics, 1996.

Wachtel, Paul, "The Effect on Earnings of School and College Investment Expenditures," *Review of Economics and Statistics*, August 1976, 58:3, 326-31.

Welch, Finis, "Measurement of the Quality of Schooling," *American Economic Review*, May 1966, 56:2, 379-92.

Witte, John F., Andrea B. Bailey, and Christopher A. Thorn, "Third-Year Report: Milwaukee Parental Choice Program," mimeo, LaFollette Institute of Public Affairs, University of Wisconsin, 1993.

Word, Elizabeth, John Johnston, Helen Pate Bain, B. DeWayne Fulton, Jayne Boyd Zaharies, Martha Nannette Lintz, Charles M. Achilles, John Folger, and Carolyn Breda, *Student/Teacher Achievement Ratio (STAR), Tennessee's K-3 Class Size Study: Final Summary Report, 1985-1990*. Nashville: Tennessee State Department of Education, 1990.