Can Equity Be Separated from Efficiency in School Finance Debates?

Eric A. Hanushek
*University of Rochester*

Chapter 3 (pp. 35-74) in: *Essays on the Economics of Education*  
Emily P. Hoffman, ed.  
Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 1993  
DOI: 10.17848/9780880995542.ch3

Copyright ©1993. W.E. Upjohn Institute for Employment Research. All rights reserved.
Can Equity Be Separated from Efficiency in School Finance Debates?

Eric A. Hanushek*  
University of Rochester

School finance discussions have concentrated on equity and rest on a few elementary premises. Poor children, often residing in decaying cities, do worse in terms of achievement, jobs, and overall success than children from better environments. Schools are society's designated institution to remedy this situation—but schools serving the disadvantaged are hampered in this task by a lack of sufficient resources. With more funding, these schools could put in place the successful programs that are available, and the cycle of poverty could be broken. It is only equitable then to support poor schools at the level at which schools for otherwise more advantaged students are financed. The more recent variant of the discussion, focusing on an adequacy version of equity, begins by noting the need for high quality education in order for an individual to compete successfully in the labor market, and then turns to a statement of how overall funding for schools must be increased to provide everybody with acceptable opportunities.

These common arguments are simple, straightforward, and compelling. Unfortunately, they are also seriously flawed. The quest for equity has generally pointed to policies that neither promote greater equity nor help deal with the serious schooling problems facing the United States.

For over two decades, courts and legislatures have been embroiled in debate and controversy over the way in which local public schools are financed. Interestingly, this has been an area where the states have completely dominated policy deliberations, and the federal government has never played an important role. Indeed, as a direct result of the U.S. Supreme Court ruling in San Antonio Independent School Dis-

*This is a revised and expanded version of Hanushek 1991.
strict v. Rodriguez, the court discussion has been conducted exclusively at the state level. Each state has followed a different course based on the requirements of its state constitution, the preferences of its citizens and legislators, and the wisdom of its courts. Nevertheless, while sometimes obscured by the details of specific state actions, there are common elements to the school finance policy developments in the states.

One important lesson learned over time is that school finance court cases, legislative decisions, and school policies in general are more complicated than was previously thought. The framework for deliberations on school finance reform was developed in the 1960s and given national attention through the landmark case in California, Serrano v. Priest. This case, which has been transported elsewhere, set out what is now the standard argument:¹ (1) Traditional funding of schools, which relies heavily on local funds raised substantially by property taxes, leads to large disparities in the education available to rich (suburban) students and to poor (urban and rural) students; and (2) The inequities in the quality of schooling resulting from the fiscal system must be corrected, and the courts are an obvious route to forcing the legislature to provide the economically and educationally disadvantaged with better schools.

An updated version of these arguments is found in Kozol (1991), where the disparities in schools between some of the nation's best and worst schools are described in vivid detail. Armed with this descriptive information, Kozol proceeds directly to the policy conclusion that all schools should be moved to duplicate the very best, a conclusion that merges both the equity and the adequacy arguments.

We have now discovered that many of these simplistic views are misleading, if not just plain wrong. The required actions involve more fundamental adjustments than merely redirecting funds, and these fundamental changes are difficult to implement directly from the court or from the capitol. These complications are addressed in the subsequent discussion.

This paper considers the overarching public policy issues involved in searching for improved equity through altering school financing arrangements, concentrating on the central policy issues that transcend state boundaries. Moreover, it avoids all consideration of legal theories
and interpretations that have surrounded the court cases except as they intersect with larger matters of broader educational policy.

Most school finance discussion, as opposed to school policy discussion, has focused almost exclusively on variations in expenditure per student. A variety of reasons can be cited to explain this emphasis. First, expenditure levels are easily measured and easily modified by the court or legislature. Second, it seems reasonable to presume that what a school can offer in terms of quality or breadth is directly correlated with the resources devoted to the task. Finally, even if there is some doubt about how well money is currently being spent, money well spent would surely make a difference. Each of these premises is reasonable if schools are operating efficiently. Given efficient school operation, expenditure is a good index of performance. On the other hand, if schools are not operating efficiently, the interpretation of expenditure differences becomes totally ambiguous, because expenditure variations need no longer be directly correlated with variations in school quality. Moreover, added funding of schools may lead to no gains in student performance.

One fundamental observation underlies the discussion in this paper: There is no systematic relationship between school expenditure and student performance. This observation implies a significant level of inefficiency in schools. Given that, legal arguments and policy decisions based on expenditure variations are simply suspect, at least from an educational perspective. Indeed, many popular changes, both proposed and adopted, no longer look like "reform" but instead tend to move us away from good policy.

School finance discussions have not totally ignored the potential pitfalls of concentrating on expenditure alone. After passing references to issues of efficiency along with assertions that the research is ambiguous, pragmatism is frequently claimed as the underlying justification: expenditure differences appear to be such a reasonable measure of differences in schools, and they are measurable. I argue later that this logic is likely to cause serious distortion in policies.

The plan of this paper is straightforward. It begins with a discussion of the evidence about expenditure and school performance. It then considers how this evidence relates to court cases and overall judgments about a state's schools. It concludes with an examination of how court
cases, and the related legislative actions, relate to effective policies toward schools.

**What We Know About School Expenditure**

Because the interpretation of expenditure differences is so central to all discussions of school finance, this section provides evidence concerning the relationship between expenditure and student performance. It is simply not possible to ignore these data in setting school policy when the objective is either to improve overall student performance or advance the cause of true educational equity.

**Aggregate Data**

Much of the current concern about the performance of our schools is motivated by the fact that student performance has remained constant or actually fallen during a period in which school spending has continually increased. Figure 1 illustrates this by superimposing the trend in student performance on the trend in educational expenditure. Real expenditure per pupil has risen steadily and dramatically over the past two decades. Specifically, after allowing for inflation, expenditure per pupil more than doubled between 1967 and 1991; this corresponds to about a 3.5 percent compound annual growth rate. At the same time, performance as measured by Scholastic Aptitude Test (SAT) scores fell to a level significantly below those attained during the mid-1960s. Moreover, while there was some recovery from the 1979-80 trough, the improvements of the early 1980s have now ceased.

There are reasons for quibbling about these specific statistics for both achievement and spending. The measurement of performance by SAT scores has been questioned because the test does not rely on a representative sample, because the test-taking population has changed over time, and because the content of the test itself may have changed. Analysis of these objections, however, indicates clearly that the observed achievement decline is not simply an artifact of that specific test. Further, declines have been registered on a variety of other tests given over the same time (see Congressional Budget Office 1986,
Continued international evidence also places U.S. students behind a surprisingly wide range of foreign students on math and science performance. For example, in tests of advanced algebra for twelfth graders in 1982, U.S. students trail students from Hong Kong to Hungary, bettering only the students from Thailand in fifteen countries sampled (McKnight et al. 1987). Thus, there is no doubt that students are performing worse now than they did in the past, when spending on schools was noticeably less.

**Figure 1**

Real School Expenditure and Achievement 1967–1991

![Graph showing the relationship between average SAT scores and real expenditure per pupil from 1967 to 1991.](image)

*Current expenditure in 1992 dollars per student in average daily attendance (ADA).*

Similarly, some have argued that the tasks facing schools have changed over time so that the comparisons of expenditures are not strictly appropriate. For example, increased expenditure may partly
reflect attempts to educate more expensive students—handicapped students, immigrants, and other educationally disadvantaged. Again, however, while these changes in student populations undoubtedly have some influence on costs, they are insufficient to explain the substantial aggregate increases that have transpired.

Moreover, it is important to note that the expenditure patterns reflect a number of underlying adjustments, which mirror common policy recommendations. Pupil-teacher ratios have fallen steadily for the past three decades. While there were twenty-five students per teacher in public elementary and secondary schools in 1965, there were fewer than eighteen in 1985.8 Over the same period, the proportion of teachers holding a master's degree or above went from under a quarter to over half. Median teacher experience also almost doubled, going from eight years in 1966 to fifteen in 1986.9 The only aggregate input not to follow this steady pattern is teacher salaries. Real teacher salaries, as best we can tell, have cycled: average salaries rose through the 1960s, fell back in the mid to late 1970s, and rose again during the 1980s.10

The aggregate picture is clear. School spending has increased dramatically since the mid-1960s, largely through the instituting of policies that educational decision makers have proposed as a way of improving student performance—reducing class sizes and upgrading the education and experience of the teaching force. Yet student performance has actually fallen over the same period.

**Individual- and School-Level Analyses**

Although research into the determinants of students' achievement takes various approaches, one of the most appealing and useful is what economists call the production function approach, or in other disciplines the input-output or cost-quality approach. In this approach, attention is focused primarily on the relationship between school outcomes and measurable inputs into the educational process.

The origin of estimating input-output relations in schools is usually traced to the monumental U.S. study, *Equality of Educational Opportunity*, or what is more commonly known as the Coleman Report. Explicitly designed to study equity, this report was the U.S. Office of Education's response to a requirement of the Civil Rights Act of 1964 to investigate the extent of inequality (by race, religion, or national ori-
gin) in the nation's schools. The study's fundamental contribution was to direct attention to the distribution of student performance—the outputs with which I am concerned here. Instead of addressing questions of inequality simply by producing an inventory of differences among schools and teachers according to race and region of the country, the Coleman Report sought to provide an understanding of which, if any, of the observed differences in school resources were important for student learning. This very different perspective—the right one when student educational performance is the concern—set a standard in the school policy debate. Unfortunately, in the subsequent financial equity debate, this important innovation has been largely ignored.

The Coleman Report was widely interpreted as finding that schools are not very important in determining student achievement. Families and, to a lesser extent, peers were seen to be the primary determinants of variations in performance. The findings were clearly controversial and immediately led to a substantial research effort to compile additional evidence about the relationship between school resources and school performance.11

The underlying model guiding the Coleman Report and most subsequent studies is very straightforward. It postulates that the output of the educational process—that is, the achievement of students—is related directly to a series of inputs. Policy makers directly control some of these inputs—for instance, the characteristics of schools, teachers, and curricula. Other factors, such as families and friends plus the innate endowments or learning capacities of the students, generally cannot be affected by public policy. Further, although achievement is usually measured at discrete points in time, the educational process is cumulative; past inputs affect students' current levels of achievement.

Based upon this model, statistical techniques, typically some form of regression analysis, are employed to identify the specific determinants of achievement and to make inferences about the relative importance of the various inputs into student performance. This summary highlights the overall findings from the research.

These studies of educational production relationships measure output not only by student scores on standardized achievement tests but also by other quantitative measures, such as student attitudes, school attendance rates, and college continuation or dropout rates. The general interpretation is that they are all plausible indicators of future suc-
cess in the labor market. This interpretation has been confirmed by other research into labor market outcomes (see review in Hanushek, Rivkin, and Jamison 1992).

Empirical specifications of production function models have varied widely in details, but they have also had much in common. Family inputs tend to be measured by sociodemographic characteristics of the families, such as parental education, income, and family size. Peer inputs, when included, are typically aggregate summaries of the sociodemographic characteristics of other students in the school. School inputs include measures of the teachers' characteristics (education level, experience, sex, race, and so forth), of the school's 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, such as the normal administrative records of schools, that were constructed for other purposes.

Empirical Results for Expenditure Effects

The production function approach has been broadly employed to investigate the impact on school performance of the core factors determining expenditure on education. Instructional expenditure makes up about two-thirds of total school expenditures. Instructional expenditure is in turn determined mostly by teacher salaries and class sizes. Finally, in most U.S. school districts, teacher salaries are directly related to the years of teaching experience and educational level of the teacher. Thus, the basic determinants of instructional expenditure in a district are teacher experience, teacher education, and class size. Most studies, regardless of what other school characteristics might be included, 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.)

Because the analyses have such common specifications, the effects of the expenditure parameters can easily be tabulated. Here I present data from a reasonably exhaustive search that uncovered 187 separate "qualified studies" found in thirty-eight separate articles or books through the middle of 1988. These studies, while restricted to public schools, cover all regions of the United States, different grade levels,
different measures of performance, and different analytical and statistical approaches. About one-third draw their data from a single school district, while the remaining two-thirds compare school performance across multiple districts. A majority of the studies (104) use individual students as the unit of analysis, whereas the remainder rely upon aggregate school-, district-, or state-level data. The studies are split about evenly between primary schooling (grades one through six) and secondary schooling (grades seven through twelve). Over 70 percent of the studies measure school performance by some kind of standardized test. However, those using nontest measures (such as dropout rates, college continuation, attitudes, or performance after school) are for obvious reasons concentrated in studies of secondary schooling. There is no indication that differences in sample and study design lead to differences in conclusions.13

According both to conventional wisdom and to generally observed school policies, each factor should have a positive effect on student achievement. More education and more experience on the part of the teacher cost more and are presumed to improve individual student learning; smaller classes (more teachers per student) are also expected to be beneficial.14 More spending in general, higher teacher salaries, better facilities, and better administration should also lead to better student performance. The quantitative magnitudes of estimated relationships are ignored at this point, and attention is focused on the direction of any estimated effect.

The data in table 1 provide a picture of how well conventional wisdom and common school policies hold up to analysis. The columns in the table divide the available estimates by direction of effect and statistical significance. Since not all studies contain estimates of each expenditure component, the first column simply indicates the total number of estimates available. Thus, for example, 152 of the 187 studies include an estimate of the effect of teacher-pupil ratios, or class sizes. Of the 152 estimates of the effects of class size, only 27 are statistically significant. Of these, only 14 show a statistically significant positive relationship, whereas 13 display a negative relationship.15 An additional 125 estimates show that class size is not significant at the 5 percent level.16 Nor does ignoring statistical significance help to confirm the benefits of small classes, since the insignificant coefficients have the "wrong" sign by a 46 to 34 margin.17
The entries for teacher education tell a similar story. The statistically significant results are split between positive and negative relationships, and in a vast majority of cases (100 out of 113) the estimated coefficients are statistically insignificant. Forgetting about statistical significance and looking just at estimated signs again does not make a case for the importance of added schooling for teachers.18

Table 1. Summary of the Estimated Relationship Between Student Performance and Various Components of School Expenditure (187 studies)

<table>
<thead>
<tr>
<th>Input</th>
<th>Number of studies</th>
<th>Statistically significant</th>
<th>Statistically insignificant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Teacher/pupil</td>
<td>152</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>Teacher education</td>
<td>113</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Teacher experience</td>
<td>140</td>
<td>40</td>
<td>10</td>
</tr>
<tr>
<td>Teacher salary</td>
<td>69</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Expenditure/pupil</td>
<td>65</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Administrative input</td>
<td>61</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Facilities</td>
<td>74</td>
<td>7</td>
<td>5</td>
</tr>
</tbody>
</table>


Teacher experience is slightly different. A clear majority of estimated coefficients point in the expected direction, and about 29 percent of the estimated coefficients are both statistically significant and of the conventionally expected sign. But these results only appear strong relative to the other school inputs; they are hardly overwhelming in an absolute sense. Moreover, they are subject to interpretive questions. Specifically, these positive correlations may result from senior teachers having the ability to locate themselves in schools and classrooms with good students. In other words, causation may run from achievement to experience and not the other way around.
Overall, the results are startlingly consistent. No compelling evidence emerges that teacher-pupil ratios, teacher education, or teacher experience have the expected positive effects on student achievement. There is no reason for confidence that hiring teachers with more education or having smaller classes will improve student performance. Evidence of the effect of teacher experience appears marginally more convincing, at least when no consideration is given to the magnitude of any relationship.

The remaining rows of table 1, summarizing information on other expenditure components, including administration, facilities, teacher salaries, and total expenditure per student, provide poorer evidence on the relationship of resources and performance, but what evidence does exist is consistent with the previous results. The quality of administration is measured in a wide variety of ways, ranging from characteristics of the principal to noninstructional expenditure per pupil. Similarly, the quality of facilities is identified through spending and many specific physical characteristics. If only because of the preponderance of positive signs among the significant coefficients, administration appears marginally stronger in its relationship to student achievement than facilities. Nevertheless, the available evidence on both again fails to support convincingly the conventional wisdom.

Finally, and not surprisingly, explicit measures of teacher salaries and expenditure per student do not indicate that they play an important role in determining achievement. After all, the underlying components of these expenditures were themselves unrelated to achievement. While negative expenditure effects—in which funds are not only unproductive but also harmful—are difficult to interpret, it is much easier to believe that differences in spending have little or no impact on student performance.

Without systematic tabulation of the results of the various studies, it would be easy to conclude that the findings are inconsistent. But there is a consistency, though it does not match the conventional wisdom. The research reveals no strong or systematic relationship between school expenditure and student performance. This is the case both when expenditures are decomposed into their underlying determinants and when they are considered in the aggregate.

Given the general biases toward the publication of statistically significant estimates, the paucity of results confirming the conventional
wisdom is notable. The common calculation of statistical significance is inappropriate when a series of sequential tests of alternative formulations of the achievement relationship is conducted. A sequential approach built on the calculated statistical tests will yield biased estimates of significance. In reality, too many estimated parameters will be judged to be significant.22

These results reflect the structure and operating procedures of schools observed in existing settings. 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 or not schools can systematically choose and retain the best teachers from the pool. The results on salary differentials presented previously might be very different if schools faced a greater incentive to produce student achievement and if mechanisms for teacher selection were altered. In other words, there seems little question that money could count. It just does not systematically do so with the current organization of schools.

Moreover, the consistency criterion used to judge the results and the potential for policy improvements does not suggest that money never counts. The results are entirely compatible with the notion that some schools use funds effectively and others do not. But unless some way is found to change the districts that would squander additional funds into districts that would use them effectively, added resources are not likely to lead to any improvement in average performance. Good uses of funds are balanced by bad uses within the current structure.

Other Inputs into Education

Since the publication of the Coleman Report, intense debate has surrounded the fundamental question of whether schools and teachers are at all important to the educational performance of students. The Coleman Report has been commonly interpreted as finding that variations in school resources explain only a negligible portion of the variation in student achievement. If this were true, it would not matter which particular teacher a student had or which school a student attended—a conclusion that most people would have difficulty accepting.
The findings of direct analyses of differences among teachers are unequivocal and indicate a very different conclusion: teachers and schools differ dramatically in their effectiveness. A number of studies provide analyses of the differential effectiveness of teachers and schools based on estimation of the average gain in performance of each teacher's (or school's) students. These studies confirm that there are striking differences in teacher performance as measured by average gain in student achievement.

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

These input-output analyses have also investigated a wide variety of other school and nonschool factors. Although it is difficult to be specific in any summary of other factors because the specifications of the various inputs employed in the statistical analyses vary widely, three generalizations are possible. First, family background is clearly very important in explaining differences in achievement. Second, while considerable attention has been given to the characteristics of peers or other students within schools, the findings about their effects are ambiguous. Finally, studies have examined many additional measures of the effects of schools, teachers, curricula, and especially instructional methods on achievement, but no simple characterization of good teachers emerges.

While not systematically addressed by existing research, one plausible interpretation of the combined results of these studies is that an important element of “skill” is involved in being a successful teacher. Skill refers simply to the ability of some teachers to promote higher achievement among their students. The evidence previously presented then indicates that it is currently impossible to identify, much less to measure, components or elements of this skill with any precision. Moreover, the direct evidence casts doubt on whether any form of teacher training course could be organized to foster high skill levels in
teachers. In simplest terms, if we cannot define or measure it, how can we teach it?

Implications for Equity and School Finance Reform

I now turn to the application of this evidence to consideration of school finance reform. Here I sketch some obvious and some less obvious implications of the preceding findings. Again, while school finance policy frequently contains many state-specific nuances, this discussion concentrates on two common elements of “reform.” Most school finance reform programs, based on simplistic equity notions, assume that a basic objective is to limit local variations in school expenditure or, if variations are to exist, to insure that such variations are not related to the property wealth of the district.

The Central Implication

The evidence on school performance indicates that variations in school expenditure are exceedingly poor measures of the variations in education provided to students. Most directly, when students’ learning is the concern, the conventional evidence about inter-district disparities in spending does not identify where educational deficiencies are to be found, and such evidence is generally irrelevant for either an equal protection or an educational disparity court case. Such evidence about expenditure simply does not indicate differential provision of education. Therefore, showing how expenditures vary, either absolutely or in accordance with characteristics of districts and students, does not have much use.

We must be quite precise about the interpretation of expenditure. As previously noted, most economists, including myself, would readily accept that differences in spending would be directly related to the education provided if schools were operating efficiently. The previously presented evidence indicates clearly, however, that assuming efficiency in spending is entirely inappropriate.

While there are many alternative ways to define and measure educational equity, only the most narrow of these would call for paying
attention to expenditure variations in the face of the evidence that such expenditure variations are unrelated to the education provided. The standard employed would have to be a rigid one linked to dollars, with total disregard for the quality of schooling received by students.

In other words, equity and efficiency are inextricably linked. It is not possible to ignore efficiency issues under the guise of being concerned solely with equity.

Other Implications

There is another side to this discussion: What is likely to happen if we disregard the evidence on the interpretation of expenditure differences and simply make policy on the basis of expenditure differences? This consideration is prompted by a few arguments that are sometimes heard, such as: “The educational problem of the poor is serious, and equalizing expenditure cannot hurt;” or “We should at least give everyone the same chance to make mistakes.” The policies flowing from such notions do, unfortunately, have a down side to them.

First, a likely reaction to any move to lessen variation in expenditure is to increase the total level of expenditure on schools. The reason is simple: a state legislature, faced with a need to alter expenditure patterns, finds it much easier to redistribute a larger pie than a fixed pie. In the school finance debate, this is frequently referred to as “leveling up,” or bringing the low-spending districts up to the spending levels of the top districts. The arguments behind the policy are generally based either on the need to do better or on pure political necessity. On the other hand, because of the potential for disruption and the obvious divisiveness of “leveling down,” there is seldom much interest in this idea. The previous evidence indicates, however, that added funds will on average be dissipated on things that do not improve student achievement (at least unless other, larger changes are also made). Teachers, administrators, and perhaps taxpayers in some districts gaining funds will probably be happier, but the average state taxpayer and parent will not find that the resulting changes do much more than increase tax bills.

Second, there is no assurance that the new funds will go to the schools of poor children. As indicated previously, one of the pervasive views of finance “reform” is that poor children will be helped (or at
least will have a better chance by virtue of greater funding). However, reform schemes designed to follow district wealth patterns can lead to unexpected outcomes because frequently there is not a strong relationship between district wealth and the concentration of student poverty. Some states find that wealthier districts in terms of property wealth per student also have concentrations of poorer families and children. New York State provides a good illustration. Consider the six largest cities in New York State: New York City, Buffalo, Rochester, Yonkers, Syracuse, and Albany. Albany and Yonkers have tax bases in which real property per student is greater than the state average; New York City, Rochester, and Syracuse have tax bases per student only slightly below the state average; and Buffalo is left with a tax base 30 percent below the state average. Yet all of these districts except Yonkers have poverty rates for children above the state average. For example, while the average poverty rate in New York State for children 18 or younger in 1980 was 19 percent, it was over 36 percent in New York City and over 30 percent in Buffalo.27 The largest districts in the state intervened (unsuccessfully) on the side of the plaintiffs in the Levittown case and introduced a new argument, municipal overburden,28 in order to protect their funding. In other states, property wealth and poverty may be negatively correlated—that is, high property wealth tends to be found in districts with a small poverty population, but even in these states the overall pattern clearly does not hold jurisdiction by jurisdiction.29 Therefore, while not inevitable, it is likely that many districts serving poor children are hurt in spending terms by plans to neutralize expenditure on the basis of district wealth. Moreover, because of a combination of federal and state grants, districts with concentrations of poor students frequently have above average spending, regardless of their property wealth or overall economic health.30 Programs to limit variations in expenditure could operate to cut back existing compensatory spending for disadvantaged students.

Third, spending differences may not even accurately reflect the real resources each district is able to deliver (i.e., the actual educational inputs). This is the simple result of possible cost differentials facing individual districts. That is, if districts face different prices for things they might buy, from teachers to buildings and equipment, dollar variations themselves do not indicate variations in available real resources. As a simple example, if the schools in one city were less pleasant and
desirable than those in other cities, it would be necessary to pay a higher salary to hire a teacher of equal quality. An extension of this notion involves districts faced with concentrations of students who are more difficult to educate because of a variety of pre-existing educational deficiencies. These, like cost differences for inputs, lead to expenditure variations in districts behaving in an otherwise identical manner. (Indeed, many state funding formulae recognize such issues and attempt to adjust for input cost differences or for differences in student preparation, handicap status, and the like, even though the magnitude of any real cost differences is poorly understood).

Fourth, districts themselves are not entities to which educational policies should be geared. Individuals choose among districts when they enter an area and move among districts after they live in an area. In fact, there is extensive evidence that individuals make choices among districts in part to satisfy their demands for various public services. Some people who place considerable weight on schooling search for districts that seem to emphasize quality schooling. Others who emphasize other goods or even low public expenditure seek districts that provide an agreeable level and pattern of the services they are looking for. Certainly this system has some drawbacks. Moving can be expensive, and some might find it difficult to move to the districts they would like, for example, because of housing prices, commuting costs, or discrimination. Nevertheless, the fundamental fact for this discussion is that individuals generally have considerable latitude in choosing schools. They are not inextricably tied to a particular district and are not doomed to whatever expenditure levels currently exist in a specific district. Finally, individual districts change their expenditure in line with the desires of the population and with population shifts, so that districts may increase or decrease their expenditure over time. For example, it is possible to trace the movement of district expenditure in the State of Indiana between 1977 and 1987. Only forty-three of the seventy-six top spending districts in 1977 remained in the top quartile in 1987; only forty-two of the seventy-six bottom quartile districts remained there from 1977 through 1987. Thus, policy discussions that speak generally of the population as captives of districts with undesirable spending patterns tend to miss an extremely important feature of the political economy of local jurisdictions. (The special problems of "mobility-constrained" groups, such as the poor, are discussed below).
Fifth, the preferences and movements of citizens across district boundaries have direct ramifications for the observed distribution of property wealth. Specifically, districts that appear to offer a particularly favorable tax and school quality package will appear relatively attractive to many people. This will lead to a bidding up of housing prices in such desirable jurisdictions, because they are in demand, other things being equal. In fact, it is well documented that "otherwise identical" houses will sell for different amounts because of citizens' evaluations of the taxes and the schooling being offered. (See Tiebout 1956, Oates 1969, Rosen and Fullerton 1977, and Wendling 1981). Another way of saying this is that some people pay for their schooling up front through the capitalization of school advantages into the price of homes. Some places that initially look attractive from the vantage point of the tax rate alone are really less attractive because the low rate is multiplied times a high valuation (relative to the other attributes of the home). This has, among other things, a direct effect on the property tax base of the community—something that is often entered into the discussion of the "inequities" of the school finance system. Moreover, reform changes in the funding formula of the state imply distributing somewhat arbitrary capital gains and losses across the jurisdictions in the state. Some places will be made more fiscally attractive and some less by major changes in the financing laws, leading to changes in the capitalization of fiscal differences.

Sixth, in most states spending levels reflect a wide variety of things, including the preferences of the citizens. While it is common to argue that local property wealth is the primary determinant of expenditure differences, that simply is not the case. For example, even though New Jersey and Indiana have relied on local property taxes to fund schools, rough estimates indicate that less than a fifth of the variation in expenditure would be eliminated by totally equating local property wealth per student. A combination of local preferences, differences in student needs, curricular choices, cost differentials, and a variety of other factors completely dominate property wealth in the determination of the pattern of expenditure.

Seventh, differences in tax rates across communities bear no direct relationship to the degree of educational equity. Most importantly, school finance reform has been based on perceived differences in the quality of education available, and the quality of education is not
related in any simple way to tax rates. The tax rate provides an indication of the price that residents face to raise funds for schools, and high tax rates might indicate that some districts find it more difficult than others to raise funds through the property tax. But tax rates differ according to a variety of factors, including community preferences, community income and wealth, the amount of nonresidential wealth in the tax base, and so forth. The pattern of tax rates may be an issue from the standpoint of various notions of “taxpayer equity,” but tax rates seldom have much to do with considerations of equity in education. Further, while the education clauses of state constitutions may require states to provide certain levels of education, they never indicate that school tax rates must be equalized across a state.

This list of likely ramifications underscores the point that simple alterations in expenditure patterns can have consequential and undesirable effects. What is already known about the educational process and about behavior of local jurisdictions leads to the inescapable conclusion suggested in my introduction: the general assumptions behind early school finance reform are misleading at best.

**Magnitude of Expenditure Effects**

The evidence presented in Table 1 did indicate that a majority of studies found a positive relationship between aggregate expenditure and student performance, albeit few statistically significant relationships. While this finding might suggest a potential for equity improvement by means of adding resources to low-spending districts, moving to such policy deliberations requires consideration of the magnitude of any expenditure effects. Specifically, how much could achievement in low-spending districts be altered by an infusion of new resources?

Two alternative estimates, representing very different circumstances, illustrate why the magnitude of performance change associated with expenditure increases must enter into policy considerations. First, Wendling and Cohen (1981) conducted a study of expenditure effects in 1977–78 in New York State, the state with the second highest rate of spending (behind Alaska) during that year. Second, new evidence for 1991 from the State of Alabama, with the 46th highest spending rate in 1990, provides information about expenditure relationships at low levels of expenditure.
New York has consistently been at or near the top of spending on schools across all states. Its average expenditure in 1978 was 45 percent above the average for the country. Wendling and Cohen (1981) examine whether or not expenditure differences among districts in New York matter for student achievement, and they conclude that indeed expenditure is important. They analyze average third-grade reading and mathematics achievement for 1,021 districts. While they examine various model formulations and different groupings of districts, the results for the entire state are representative and provide a clear indication of how expenditure relates to student performance.

The estimated effect of approved operating expenditure on student performance is uniformly statistically significant, but the magnitude of the estimates shows the difference between statistical significance and policy significance. The estimated expenditure parameters are .001 and .002 for third-grade reading and mathematics performance, respectively. This implies that a $1,000 increase in expenditure per student yields an additional point on the reading test and an additional two points on the mathematics test. While absolute scores are difficult to interpret, some idea of magnitude can be gained by looking at movement in the distribution of spending and performance. An increase of $1,000 is a 50 percent increase in state school spending and is 2.2 standard deviations in expenditure within the sample of schools, but it yields an increase in performance of only .2 to .3 standard deviations. This is equivalent to moving the average student up to around the sixtieth percentile or to moving a student starting at the 10th percentile to the 15th percentile. In simple terms, attempting to increase performance through simple increases in expenditure is very expensive.

Alabama is a relatively low-spending state, falling at the other end of the distribution from New York State. The state's highest-spending district in 1991 (Mountain Brook) had total current expenditure per student of $5,113. This is slightly below the mean level for the entire nation in 1991, for which the preliminary estimates are $5,237. The minimum spending in Alabama was about $2,900 per student. Thus, variations in spending in Alabama should provide some insight into whether or not there exists some threshold expenditure below which changes in spending have strong and noticeable effects on achievement—the issue that comes up in discussions of "adequacy" (Celis 1992).
The statistical analysis attempts to explain variations across districts in Alabama's Basic Competency Tests (BCTs). These are criterion-referenced tests adopted by the state board of education to measure whether or not students are accomplishing what was expected of them according to the curriculum for different grade and subject areas (reading, mathematics, and language arts). The performance measure is the percentage of a district's students meeting the minimum standards for the specific tests ("passing"). Weighted least squares regression analysis is employed to estimate the effect of current expenditure per student in average daily attendance (ADA) on performance, while holding constant the influence of family background and school district type (i.e., city or county district).

The results of estimates for the State of Alabama can be easily summarized. Table 2 presents the estimated expenditure effects. None of the nine estimated relationships is statistically significant at the 5 percent level, and one is even negative.

Table 2. Estimated Change in BCT Pass Rate for Expenditure Increase of $1,000 per Student in Average Daily Attendance (ADA): Alabama, 1990-91

<table>
<thead>
<tr>
<th>Test and Grade</th>
<th>Read gr 3</th>
<th>Read gr 6</th>
<th>Read gr 9</th>
<th>Math gr 3</th>
<th>Math gr 6</th>
<th>Math gr 9</th>
<th>Lang gr 3</th>
<th>Lang gr 6</th>
<th>Lang gr 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in percent passing</td>
<td>0.568</td>
<td>0.119</td>
<td>2.733</td>
<td>2.562</td>
<td>0.423</td>
<td>1.574</td>
<td>1.656</td>
<td>-0.133</td>
<td>2.565</td>
</tr>
<tr>
<td>t-ratio</td>
<td>0.51</td>
<td>0.07</td>
<td>1.95</td>
<td>1.70</td>
<td>0.19</td>
<td>0.78</td>
<td>1.00</td>
<td>-0.07</td>
<td>1.54</td>
</tr>
</tbody>
</table>

NOTE: Estimates are the weighted least squares regression estimates of the determinants of percent passing the Alabama Basic Competency Test (BCT) for different subject areas and grades for 127 school districts in 1990-91 (Hoover and Mt. Brook excluded). Estimates equations include percent of students receiving free or reduced lunch; percent nonwhite; city school district indicator; and current expenditure per ADA. Weights are the number of students in average daily attendance.

Abbreviations:
AVA: per student in Average Daily Attendance
BCT: Basic Competency Test

Table 3 presents the results of two alternative policy scenarios. The first brings all of the spending of districts below the median in the state up to the median. The second, which is sometimes called "full leveling
Table 3. Predicted Change in State Pass Rates from Increased Spending: Alabama Districts, 1990–91

A. Partial leveling-up by bringing all low-spending districts to median (cost=$74 million)

<table>
<thead>
<tr>
<th></th>
<th>Reading</th>
<th>Mathematics</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>grade 3</td>
<td>grade 6</td>
<td>grade 9</td>
</tr>
<tr>
<td>1990–91 pass rate</td>
<td>81.3</td>
<td>64.3</td>
<td>73.1</td>
</tr>
<tr>
<td>Change in pass rate</td>
<td>0.1</td>
<td>0.0</td>
<td>0.3</td>
</tr>
<tr>
<td>Projected pass rate</td>
<td>81.4</td>
<td>64.3</td>
<td>73.4</td>
</tr>
</tbody>
</table>

B. Full leveling-up by bringing all districts to top (cost=$1.05 billion)

<table>
<thead>
<tr>
<th></th>
<th>Reading</th>
<th>Mathematics</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>grade 3</td>
<td>grade 6</td>
<td>grade 9</td>
</tr>
<tr>
<td>1990–91 pass rate</td>
<td>81.3</td>
<td>64.3</td>
<td>73.1</td>
</tr>
<tr>
<td>Change in pass rate</td>
<td>0.9</td>
<td>0.2</td>
<td>4.2</td>
</tr>
<tr>
<td>Projected pass rate</td>
<td>82.2</td>
<td>64.5</td>
<td>77.3</td>
</tr>
</tbody>
</table>
"up," brings all districts up to the level of the highest spending district in the state. Bringing all states to the median expenditure level would cost an additional $74 million (compared to total spending of $2.4 billion). While this increase of 3 percent is not a large relative change in expenditure, the top panel of table 3 indicates that it would yield imperceptible changes in performance on most of the tests. The bottom panel provides estimates of the achievement effects of full leveling up—i.e., bringing all district spending up to that of Mountain Brook. Such a policy would cost $1.05 billion and would yield at most a 4 percentage point increase in students passing the BCT in the state. The net impact of leveling up is shown in figure 2.

Figure 2
Alabama Performance with Leveling-Up

The importance of these results is clear. A policy of bringing all districts to the top in spending would place Alabama schools at approximately the mean for the nation, up dramatically from its current
position in the national distribution of spending. But this increase of the state's school spending by over one-third would have very small effects on aggregate school performance, at least if the schools behave in a way consistent with current practice. The resulting performance would remain very far below the state's goals of a 95 percent passing rate on the separate BCT tests.

There is also no evidence from analyzing the schools of Alabama that there exists a minimum threshold for school spending. There is no apparent range of stronger influence of spending on achievement than is found for the entire set of schools.

Policy Alternatives

Concerns about the implications of school finance reform do not, of course, vitiate the undeniable need to improve our public schools. The deplorable conditions described in Kozol (1991) require addressing. The intentions of finance reformers have been, in my opinion, good. Only their approach is questionable. Three general factors lead to the judgment that structural change is essential. First, in absolute terms students are not performing up to expectations. Performance, as measured by standardized tests over time, international comparisons of tests, various measures of workplace performance, and common perceptions, is currently unacceptable. Second, as indicated by the previous evidence, there is overwhelming evidence that the resources devoted to schools—which have been both large and increasing--have not been effectively used. Third, the significantly skewed distribution of educational success, which leaves poor and minority students behind the rest of the population, is incompatible with most people's views on the goals of our society.

The previous sections of this paper indicate why "reform" as commonly included in school finance considerations is unlikely to address any of these causes of concern. The primary focus on the distribution of state financial aid or the limits on local fiscal options distracts attention from the issues of school organization, incentives for performance, and the goals of the system. Because of the contentiousness of issues surrounding the distribution of funds, school finance debates
have the potential for absorbing all energy related to school policy. Thus, in addition to generally offering few solutions to the problems previously identified, there is a significant opportunity cost in stalling some of the fundamental restructuring that must proceed if we are to deal with the current problems of schools. This problem of distraction, of course, is not inevitable, but there are strong forces pushing in that direction.

The concentration of courts and legislatures on finance reform while skirting away from more fundamental policy considerations does follow a certain logic. Expenditures are readily measurable; there is a plausible argument behind their importance; there is no obvious alternative focus of policy; and operating on expenditure at least represents doing something. In other words, there are serious problems, and it is perceived that at least some attempt to remedy them should be undertaken.

This logic is supported by the lack of convincing evidence that any specific approaches or policies will bring about significant improvements in student performance. As reviewed previously, no set of simple changes involving either resources or programs shows a consistently strong relationship with performance.

But that is just the problem. Concentrating solely on dollars or resources does not confront the basic structural problems in the operations of schools. For example, while Jonathan Kozol (1991) points to the lack of achievement of impoverished students and calls for increased funding, he ignores the fact that the increased funding will be administered by the same school boards and administrators that he railed against in *Death at an Early Age* (Kozol 1967). An alternative approach, which suffers from many of the same problems, is to argue that additional funds would not be utilized in old unproductive ways but would be used in highly directed ways that insured achievement. This approach is often accompanied by the description of a specific program that has been shown to work in the few places where it has been tried. The problem, of course, is that we really do not know in general terms what will work, and the successful programs that have been identified have not been broadly introduced by districts with the funds to do so (for example, through general federal compensatory education funding or more generous local support). There is simply no reason to believe that a centrally directed system of increased funding
relying on the simple identification of productive programs or spending patterns would be more productive than what has been currently observed.

An alternative formulation of educational policies avoids the pitfalls of previous approaches and offers considerably more promise of improvement. The alternative is moving to organizations and incentive systems that directly reward performance. The current set of policies, almost exclusively pursued, involves providing or requiring certain inputs—expenditures, class sizes, teacher attributes, and the like. These input policies are essentially pursued and continued without regard to their effectiveness, either in the aggregate or in specific instances. The proposed alternative is to concentrate on student performance instead of factors thought (or hoped) to be important in determining student performance.

Various systems have been used or suggested to promote performance-based policies, including merit pay for teachers, merit awards for schools that perform well, and a variety of plans emphasizing choices among educational institutions. Essentially, the common ingredient of such plans is that resources are directly related to performance: if performance is high, resources are high; if performance is low, resources are commensurately low. For example, merit pay for teachers operates by increasing salaries of those who perform well and not of those who perform poorly. Similarly, choice plans, which operate by allowing students and parents to choose among alternative schools, work by reinforcing parental judgments about quality schools, with suitable flows of resources to the schools that attract students.

The orientation of these policies is based on finding the correct incentives. If tangible incentives for improved performance are offered, most decision making can be expected to improve. Actual operations of hiring, promotion, curriculum, student placement, and the like—while not specified or regulated by a central authority—can be expected to respond to incentives. This has been demonstrated by wide-ranging research, both in education and elsewhere. Determining effective incentives, however, will take experimentation, bargaining, and evaluation.

There are many different versions of these performance-based plans, particularly of the choice plans. Commonly discussed choice plans range from magnet or special schools (which are fairly wide-
spread in some urban districts) to full vouchers, which provide parents with funds to pay tuition at either public or private schools of their choice. Between these extremes are free choice within public school districts, open enrollment in any public school in the state, and tuition tax credits to rebate a portion of any tuition payments to the parents.

Performance-based options have been discussed widely and will not be reviewed here (see Chubb and Hanushek 1990). Instead I will merely highlight two features. Each option has conceptually appealing elements. And there is little historical evidence for each option that would provide details of either how it should be implemented or the magnitude of gains that might be expected. In other words, there is also considerable uncertainty, particularly about details of implementation, because these approaches are largely untried. The uncertainty should not, however, be taken as a reason for avoiding them but should dictate a more interactive approach to policy making instead. Moreover, as indicated, each performance-based option has considerable appeal, especially as an option to the almost universally employed input policies that have had such a dismal record.

The performance-based view of educational policy is very different from the current view of how to make policy. It also is not very amenable to the simple remedies and simple tracking of responses so appealing within a court context. Nevertheless, for all its messiness and uncertainty, it offers some realistic hopes for improvement—something that is absent from narrow decisions on expenditure and other inputs.

The use of performance-based plans is supported by the research into educational performance reviewed above. This research indicated extremely large and significant variations in the performance of individual teachers and schools. It is also very important to reiterate here that research has concentrated on the value-added of teachers and not on absolute performance levels of students. The research demonstrates that there can be low value-added in a "good" suburban school where the absolute level of achievement is quite high. Similarly, there can be high value-added within "bad" central city schools where students come to school quite unprepared but leave with marked increases in their achievement (see, for example, Hanushek 1992). While research cannot identify the components of successful and unsuccessful teaching, it supports the simple but powerful notion that good performance
can be identified by school administrators (see Murnane 1975 and Armor et al. 1976). Further, if this capacity can be extended to individual parents—who frequently at least act as if they can tell the difference between good and bad teachers—the groundwork for performance-based policies is established.

The overall point is straightforward: a range of effective policies appears to be available. They are, however, almost certainly very different from the traditional policy focus and the orientation of traditional school finance "reform" efforts. Moreover, instituting some of these fundamental reforms might take added funds, particularly in the implementation and learning phases. There is a huge difference, as should be clear, between expenditures directly linked to improved incentives and student performance and expenditures made in the hopes that something good will happen.

Finally, the restructuring of incentives in schools appears to be the only feasible answer to dealing with the gloomy record schools have in improving the performance of educationally and economically disadvantaged youth. Various input-oriented programs have been mounted to deal with the disadvantaged, including a large portion of all federal spending on schools, but there is little evidence that this has had much impact. The alternative to restructuring incentives as proposed here is to continue to expand the programs that have thus far been unsuccessful.

The evidence from past analyses demonstrates that good teachers exist in what are commonly thought to be bad urban districts. Their existence, however, is masked by generally low achievement levels; that is, even though an individual teacher may elicit more than one year of achievement growth within a one-year period, low absolute levels of performance could hide it. The policy problem is that we have not been able to attract, to identify, and retain sufficiently large numbers of such good teachers so as to have the kind of influence that is needed. This is just the appeal of performance-based incentive schemes. They are designed to reinforce good performance. We should, at the same time, not have overly optimistic expectations. As has been thoroughly documented, family influences are very powerful in determining achievement levels; so while specific teachers might have a substantial influence on achievement, they might not immediately overcome the deficits arising from factors outside the schools.
Indeed, it may take the continued efforts of many good teachers over the course of the student's school career. This reality, however, should not deter our efforts to provide the best possible education.

Conclusions

School finance reform, as commonly espoused in courts and legislatures across the country, is likely to work against the very improvements most needed in public schools. By its nature, emphasizing primarily the distribution of expenditure per student, financing reform is almost certain to exacerbate existing problems of inefficiency in school operation.

Discussions of school finance reform typically attempt to separate considerations of efficiency from issues of "equity." Such a distinction is impossible, however, if the definition of equity involves the learning of children, which depends directly on the ability of school districts to translate resources into student achievement. If schools are ineffective at this task, little can be done to improve equity in student performance by simply heaping more resources on poorly performing districts.

Research into the relationship between resources and student performance, conducted over the past quarter century, has demonstrated conclusively that, within the current organization and operation of schools, there is no consistent relationship between resources and student performance. Common policy proposals—ones that are used to justify pleas for added resources to school districts—simply are not supported by evidence about their performance within schools.

Ignoring the evidence on performance is likely to worsen the problems of performance and inefficiency. The current incentive structure in schools does not promote efficient use of resources. Therefore, while additional funds might be used effectively by some districts, other districts will probably use them ineffectively—which, if past history is a guide, leads to no aggregate improvement in performance from increased funds.

Large differences in performance exist among teachers and schools. Past evidence further indicates that parents and administrators can identify the best and worst among these. What is missing is an effec-
tive structure for channeling knowledge about performance into overall improvement in the schools. A variety of mechanisms for this have been proposed; although there is little operational experience with them, the key to each is that resources are more directly linked to performance of the schools. This is an entirely different perspective from what has been found in most discussions of school finance reform, which gives no weight to student performance in making resource allocations.

NOTES

1. See, for example, Coons, Clune, and Sugarman (1970) or Wise (1968).
2. The term "efficient" here is used in the economist's sense of obtaining the maximum possible performance from any given expenditure of resources. This definition is very different from that which appears to have been employed in a number of legal arguments emanating from state constitutional requirements to provide an efficient system of public schools.
3. See, for example, Coons, Clune, and Sugarman (1970, p. 29) where they discuss Ribich (1968). They state, "Ribich's painstaking analyses suggest, if anything, a variety of sometimes conflicting relationships between cost and purely economic benefits from added dollar increments." They go on to indicate, "There are similar studies suggesting stronger positive consequences from dollar increments, and there are others suggesting only trivial consequences, but the basic lesson to be drawn from the experts at this point is the current inadequacy of social science to delineate with any clarity the relation between cost and quality. We are unwilling to postpone reform while we await the hoped-for refinements in methodology which will settle the issue" (p. 30).
4. For example, after discussing the difficulty of employing alternative measures of real resource differences (such as education levels of teachers), Coons, Clune, and Sugarman (1970, p. 26) state: "We have no stomach for such an imbroglio. Ultimately we will need a standard appropriate to the rigors of judicial proof, and the only convincingly quantifiable item in the spectrum is money available for the general task of education in each district."
5. This section draws extensively on the presentation in Hanushek (1989) which in turn updates previous analyses in Hanushek (1981, 1986).
6. Current expenditures per student are deflated by the consumer price index. See Digest of Education Statistics, 1989, tables 88 and 114, and updates.
7. On the other hand, evidence from international tests in 1964 suggest that U.S. students have historically done relatively poorly (Husen 1967).
9. Moreover, only 3 percent of teachers in 1986 were in their first year of teaching (Digest of Education Statistics, 1988, table 54). The aging and stagnation of the teaching force have, however, been the subject of separate concerns.
10. The teacher salary data over time that are normally cited are provided by the National Education Association, and the sample and reliability of these are unknown (Digest of Education Statistics, 1988, table 57). An alternative source, the decennial population censuses, indicates smooth increases in salaries of teachers by decade, but these fall relative to annual earnings of all college graduates. See Hanushek, Rivkin, and Jamison (1992).
11. There were also extensive analyses of the report's methodology and of the validity of its inferences. See, for example, Bowles and Levin (1968), Cain and Watts (1970), and Hanushek and Kain (1972).

12. A qualified study was defined as a production function estimate that: (1) is published in a book or refereed journal; (2) relates some objective measure of student output to characteristics of the family and the schools attended; and (3) provides information about the statistical significance of estimated relationships. Note that 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 count as only one study). Search procedures for articles involved using the national educational data base (ERIC), searching the bibliographies of included and related articles, and scanning the tables of content of likely journals. No articles uncovered in this search and meeting the above criteria were excluded, but inevitably some were missed. Articles from this time period that were overlooked in the search process but that have subsequently been discovered include Brown (1972), Walberg and Rasher (1974), Wendling and Cohen (1981), and Walberg and Fowler (1987). And, of course, there have been publications subsequent to construction of this table. The inclusion of these articles does not change the weight of the evidence or the substantive conclusions reached.

13. The tabulations, when stratified by grade level, by whether individual or aggregate data were used, by output measure, and by value-added or level forms of estimation, yield the same qualitative conclusions.

14. Tabulated results are adjusted for variables being measured in the opposite direction; for example, the sign for estimated relationships including student-teacher ratios, instead of teacher-student ratios, is reversed.

15. Teacher-pupil ratios are treated here as synonymous with class sizes. This is not strictly the case and, in fact, could be misleading today. Several changes in schools, most prominently the introduction of extensive requirements for dealing with handicapped children in the mid-1970s, have led to new instructional personnel without large changes in typical classes. Since much of the evidence here refers to the situation before such legislation and restrictions, it is reasonable to interpret the evidence as relating to class sizes.

16. In any statistical analysis, which necessarily relies on a sample of all possible students and classroom environments, an estimated relationship may not be real but only perceived to be so because of the specific sample. Standard regression techniques provide ways of estimating the likelihood of being fooled by the sampling into thinking there is a relationship when in fact there is not. The shorthand term, "statistical significance," implies that less than 5 percent of the time when there is really no relationship would we get an estimate as large as the one obtained. In other words, when the estimate is "statistically significant," we are quite confident that some relationship does indeed exist. In all cases, however, the estimates of statistical significance assume that the "correct" relationship is being estimated; that is, that the model of achievement is properly specified to include the relevant factors determining performance.

17. Note that not all studies report the sign of insignificant coefficients. For example, 45 studies report insignificant estimated coefficients for teacher-student ratios but do not report any further information.

18. Note that only 113 studies report evidence about teacher education. Since data on teacher education are so readily available, it seems likely that a number of additional studies investigated teacher education effects but discarded the results without reporting them after finding negative or insignificant effects.

19. Information on each of these is less frequently available. This is partially explained by common reliance on administrative records which do not record them (except perhaps teacher sal-
Can Equity Be Separated from Efficiency in School Finance Debates?

aries). The level of the analysis and sampling frame for some studies offer another explanation; for example, since expenditures per student are generally measured for districts, the analyses that rely on individual student data for a single district would find no variation in this input and thus could not include it. More recent studies have generally concentrated on the analysis of individual student and classroom data and thus have not considered aggregate expenditure effects.

20. The interpretation of expenditure and salary estimates is sometimes clouded by including them in addition to teacher experience, education, and class size. Because multiple regression coefficients indicate the effect of a specific variable when all other variables are held constant, direct measures of expenditure, for example, in models also including the prime determinants of instructional spending would be interpreted as the effect of noninstructional expenditures on achievement. Also, because prices can vary across the samples in the separate studies, it is more difficult to interpret dollar measures than real input measures. Finally, eight of the thirteen significant positive expenditure results in table 1 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.

21. There are several obvious reasons for caution in interpreting this evidence. For any individual study, incomplete information, poor quality data, or faulty research could distort the statistical results. Even without such problems, the actions of school administrators could mask any relationship. For example, if the most difficult students to teach were consistently put in smaller classes, any independent effect of class size could be difficult to disentangle from the mismeasurement of a students' characteristics. Finally, the statistical insignificance of estimates can reflect no relationship, but it also can reflect a variety of data problems, including high correlations among the different measured inputs. In other words, as in most research, virtually any of the studies is open to some sort of challenge.

22. This issue is discussed in Hedges (1990). “The published literature is particularly susceptible to the claim that it is unrepresentative of all studies that may have been conducted (the so-called publication bias problem). There is considerable empirical evidence that the published literature contains fewer statistically insignificant results than would be expected from the complete collection of all studies actually conducted. . . . There is also direct evidence that journal editors and reviewers intentionally include statistical significance among their criteria for selecting manuscripts for publications. . . . The tendency of the published literature to overrepresent statistically significant findings leads to biased overestimates of effect magnitudes from published literature. . . .” (p. 19, listed references omitted).

23. These studies are analyses of covariance or, equivalently, regression analysis using individual teacher (or school) dummy variables in addition to measures of prior student achievement, family background factors, and other explicitly identified inputs in a regression format. See Hanushek (1971, 1992); Murnane (1975); Armor et al.(1976); and Murnane and Phillips (1981).

24. Perhaps the closest thing to a consistent conclusion across the studies is that “smarter” teachers, ones who perform well on verbal ability or achievement tests, do better in the classroom. Nonetheless, while plausible, there remains mixed evidence on how good teacher tests are at indicating teaching ability. Tabulations similar to those in table 1 indicate thirty-one studies that have analyzed teacher verbal scores. Of these, eight find positive and significant relationships, and another ten find positive but insignificant relationships. These overall findings have been extended by a recent study by Ronald Ferguson (1991), which finds teacher ability as measured by scores on the Texas teacher test to be related to student performance, although that study is insufficient to change the weight of the evidence.

25. The idea of skill differences among teachers is not the only possible interpretation of the data. Differences in achievement across classrooms could reflect differences in teachers, in other classroom-specific factors, or a combination of both. The teacher skill interpretation is suggested
by the fact that principals' ratings of teachers are correlated with the covariance estimates of classroom differences; see Murnane (1975) and Armor et al. (1976). Evidence on the stability of teacher effects across grades, test area, and years for individual teachers further supports the interpretation based on teacher skill; see Hanushek (1992). A discussion of skill differences in the production function context can be found in Hanushek (1986).

26. School finance court cases have typically contained two elements. First, an equal protection argument is employed, which asserts that school expenditure differences related to variations in the local property tax base are discriminatory. Second, the "education clause" usually found in the state constitution is used to back an assertion that large variations in expenditures are impermissible. In both instances, the direct evidence provided for the alleged wrong involves variations in expenditures (sometimes linked to other things such as property tax wealth).


28. The argument of municipal overburden is that excessive demands for nonschool expenditures faced by urban districts subtract from what otherwise would be available for schools. Therefore, the state funding formula should recognize these other expenditures in allocating school support. See the arguments in Levittown. For an economic analysis, see Brazer and McCarty (1987).

29. As described, there is considerable variation in tax bases and poverty rates within a state. Thus, for example, Albany had a property tax base per student that was 34 percent above the state average, and yet it also had a poverty rate above the state average. Cutting back on funds for this "wealthy" district would potentially harm sizable numbers of poor children.

30. For example, in the situation in New York State, each of the six large districts except New York City had expenditures per student above the state average. See Financial Data for School Districts, 1982.

31. This situation, known to economists as "compensating differentials," can exist whenever jobs or job locations include different attributes such as riskiness, opportunities for learning, or, in the case of cities, favorable living conditions. For a general description, see Ehrenberg and Smith (1991) or Hamermesh and Rees (1988). In the context of teachers, see Toder (1972), Antos and Rosen (1975), and Kenny and Denslow (1980). Differences in the attractiveness of areas can also lead to differences in housing and land prices, thus affecting other inputs to education. See, for example, Roback (1982).

32. These calculations rely on estimates of the relationship between expenditures per student and wealth per student in districts in these states. The $R^2$ of simple regression in each state was less than .20, indicating that one-fifth would be an estimate of the upper bound on the potential for equating spending by eliminating property tax base differences.

33. This study is entitled "Education Resources and Student Achievement: Good News for Schools," presumably because it was one of the few studies that ever found statistically significant relationships between expenditure and policy.

34. The basic regression models estimated include median years of schooling and percent below poverty in the district, percent minority students, district size, and pupil/teacher ratio in addition to expenditure. Expenditure is measured in several alternative ways, and some formulations include treatment of geographic location of districts.

35. Approved operating expenditure excludes certain categories included in total current expenditure, such as some transportation, employee benefits, etc. The average approved expenditure in 1978 for the sampled schools was $2,064, compared to an average total current expenditure for the state of $2,527.
36. The results are unchanged in a qualitative sense if performance is measured by average scores in the district as opposed to the percentage of students passing the BCT examinations. The magnitude of changes in average scores is somewhat less than of the changes in pass rates predicted for spending changes, a finding that is consistent with the notion that the average performance is relatively close to the established passing score on most of the separate BCTs.

37. Family background is measured according to the percentage of students receiving free or reduced lunch and to what percentage is nonwhite. In 1990–91, there are 129 separate school districts of which 67 serve countywide populations (outside of city districts), while the remaining 62 serve individual cities. The estimates are weighted by the number of students in average daily attendance, in order to deal with the heteroscedasticity introduced by averaging performance across populations of different sizes.

38. The estimates presented exclude two districts (Hoover and Mt. Brook) that are significant "outliers in terms of expenditure levels. Because they are noticeably distant from most other districts they have an undue influence on the estimated expenditure effects. Including these districts yields somewhat larger expenditure effects (ranging from 1.0 to 3.7 percent passing per $1,000 as opposed to -.1 to 2.7 percent passing per $1,000 in table 2). Three of the nine estimated coefficients are significant at the 5 percent level when the two outliers are included. These estimates do not, however, provide reliable information about the effects of increasing expenditure because there is little or no pattern to expenditures for the remaining 127 districts included in the analysis.
References


72


