DOCUMENT RESUME

ED 448 494 EA 029 205

AUTHOR	Fowler, William J., Jr., Ed.
TITLE	Developments in School Finance, 1997. Fiscal Proceedings from the Annual State Data Conference (July 1997).
INSTITUTION	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	NCES-98-212
ISBN	ISBN-0-16-049659-4
PUB DATE	1998-07-00
NOTE	187p.; For 1996 edition, see ED 409 634.
PUB TYPE	Collected Works - Proceedings (021)
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	Conferences; Cost Effectiveness; *Educational Equity (Finance); *Educational Finance; Elementary Secondary Education; *Finance Reform; *Outcomes of Education; *Productivity; Qualitative Research; *Resource Allocation
IDENTIFIERS	*Funding Formulas; National Center for Education Statistics

ABSTRACT

This document contains papers from the Annual State Data Conference and includes the following: (1) "Does Money Matter for Minority and Disadvantaged Students? Assessing the New Empirical Evidence" (David Grissmer, Ann Flanagan, and Stephanie Williamson), which demonstrates that additional money matters for minority and disadvantaged students, but may not matter for highly advantaged students; (2) "Rethinking the Allocation of Teaching Resources: Some Lessons from High Performing Schools" (Karen Hawley Miles and Linda Darling-Hammond), which shows how five schools support extraordinarily high student achievement by reallocating instructional resources to maximize individual attention for students and learning time for teachers; (3) "Financing Education in the District of Columbia from the Perspective of the Financial Authority" (Joyce Ladner), which describes the District of Columbia Financial Authority efforts to revamp the area's failing schools; (4) "Does Money Matter? An Empirical Study Introducing Resource Costs and Student Needs to Educational Production Function Analysis" (Corrine Taylor), which shows that geographic cost variations and special-needs student costs do not appreciably affect per student expenditures; (5) "School District Expenditures, School Resources and Student Achievement: Modeling the Production Function" (Harold Wenglinsky), which uses hierarchical linear modeling to show that expenditures on instruction and central-office administration affect teacher-student ratios, which, in turn, affect student achievement; (6) "The Development of School Finance Formulas To Guarantee the Provision of Adequate Education of Low-Income Students" (Andrew Reschovsky and Jennifer Imazeki), which explores the quandary of developing an equitable school finance formula; two "cutting-edge" papers, (7) "Using Cost and Need Adjustments To Improve the Measurement of School Finance Equity" (Lauri Peternick, Becky A. Smerdon, William Fowler, Jr., and David H. Mou); and (8) "A Proposal for Collecting School-Level Resource Data on the Schools and Staffing Survey" (Julia B. Isaacs, Michael S. Garet, and Stephen P. Broughman), examine effects of applying geographic-cost adjustments and student-need adjustments to traditional equity measures. (Most papers include several references.) (MLH)

National Center for Education Statistics

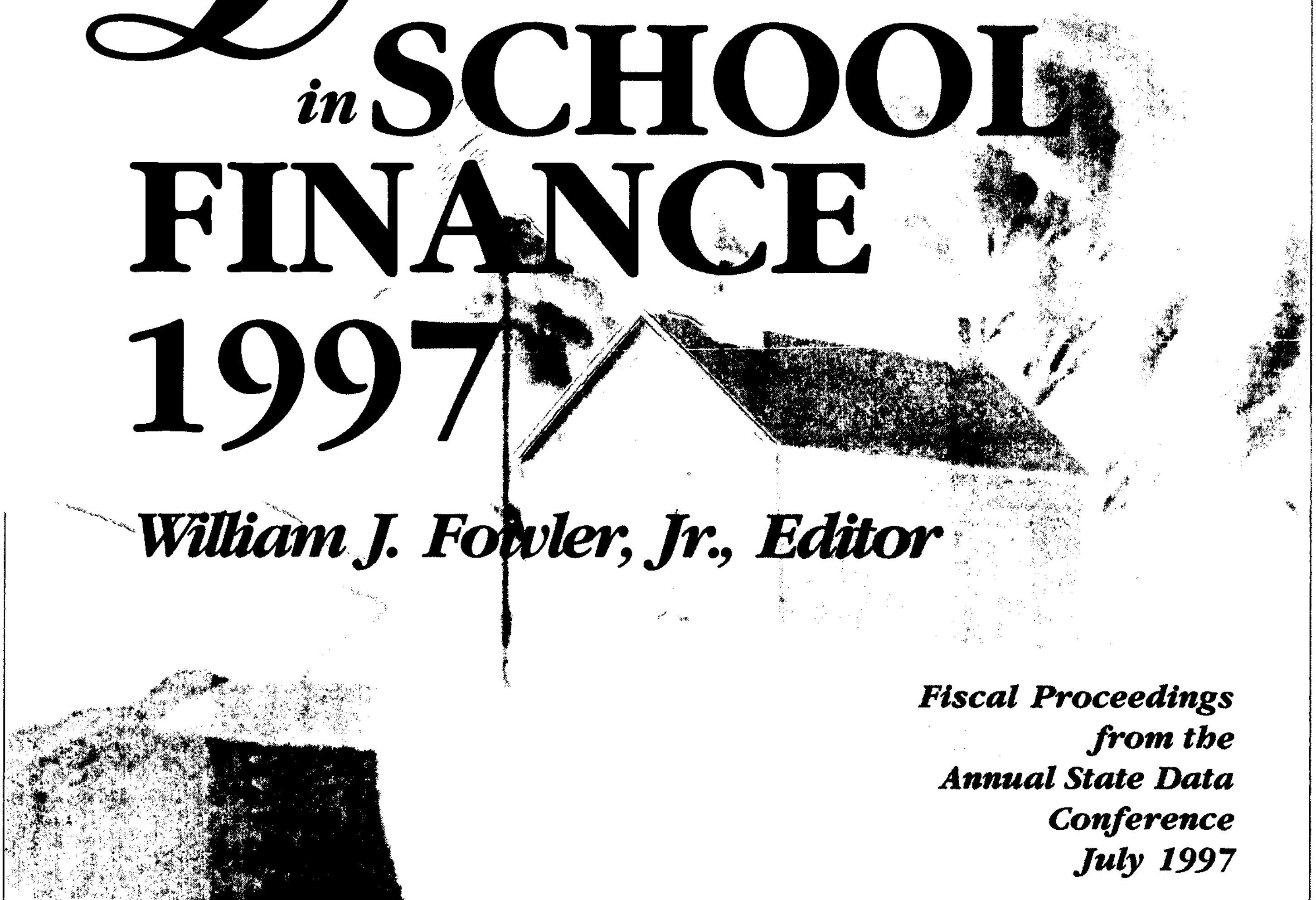
Developments in SCHOOL FINANCE 1997

U.S. Department of Education
Office of Educational Research and Improvement

NCES 98-212

National Center for Education Statistics

Developments in SCHOOL FINANCE 1997



William J. Fowler, Jr., Editor

Fiscal Proceedings from the Annual State Data Conference July 1997

U.S. Department of Education
Office of Educational Research and Improvement

NCES 98-212

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
C. Kent McGuire
Assistant Secretary

National Center for Education Statistics
Pascal D. Forgione, Jr.
Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue NW
Washington, DC 20208–5574

July 1998

The NCES World Wide Web Home Page address is
http://NCES.ed.gov

Suggested Citation

U.S. Department of Education. National Center for Education Statistics. *Developments in School Finance, 1997,* NCES 98-212, Edited by William J. Fowler, Jr. Washington, DC: 1998.

Contact:
William J. Fowler, Jr.
(202) 219–1921
Internet:
William_Fowler@ed.gov

Foreword

Paul D. Planchon, Associate Commissioner

Surveys and Cooperative Systems Group

One of the persistent dilemmas in education finance has been the inconsistent results obtained in educational research regarding the effectiveness of higher spending on student outcomes. Many of the scholars in the 1997 National Center for Education Statistics (NCES) Summer Conference addressed this issue in their presentations, and their insights and the diversity in their approaches and thinking regarding this seminal issue make this proceedings among the most captivating in the series.

Developments in School Finance contains papers by speakers at the annual NCES Summer Conference. The Conference attracts several state department of education policymakers, fiscal analysts, and fiscal data providers from each state, who are offered fiscal training sessions and updates on developments in the field of education finance. The presenters are experts in their respective fields, each of who has a unique perspective or interesting quantitative or qualitative research regarding emerging issues in education finance. The reaction of those who attended the Conference was overwhelmingly positive. We hope that will be your response as well.

This proceedings is the fourth education finance publication from the NCES Summer Conference. The papers included within present the views of the authors, and are intended to promote the exchange of ideas among researchers and policymakers. No official support by the U.S. Department of Education or NCES is intended or should be inferred. Nevertheless, NCES would be pleased if the papers provoke discussions, replications, replies, and refutations in future Summer Conferences.

Acknowledgments

The editor wishes to gratefully acknowledge the comments and suggestions of the reviewers: Michael P. Cohen, Marilyn M. McMillen, and Lee Hoffman of the National Center for Education Statistics (NCES). I also wish to acknowledge the contributions of Allison Pinckney and Carol Rohr of Pinkerton Computer Consultants, Inc. who incorporated the text, tables, and graphics into a published document, Rebecca Pratt, Ross Pfile, and Ginger Rittenhouse of Pinkerton Computer Consultants, Inc. who edited the manuscript, and Nikki Smith of Pinkerton Computer Consultants, Inc. who designed the cover.

Contents

Introduction and Overview

William J. Fowler, Jr.

National Center for Education Statistics

About the Editor

William J. Fowler, Jr. is the director of the Education Finance Statistical Center (EFSC) at the U.S. Department of Education, National Center for Education Statistics (NCES). He specializes in elementary and secondary education finance and education productivity research. His recent work has focused on the development of geographic and inflationary cost adjustments, and their effect upon measures of equity; school-level financial reporting; the construction of a student-level resource measure; and issues in analyzing NCES student achievement and finance data. His most recent publication is *Inequalities in Public School District Revenues*.

Dr. Fowler has worked for NCES since 1987, before which he served as a supervisor of school finance research for the New Jersey Department of Education. He taught at Bucknell University and the University of Illinois, and served as a senior research associate for the Central Educational Midwestern Regional Educational Laboratory (CEMREL) in Chicago and for the New York Department of Education.

Dr. Fowler received the Outstanding Service Award of the American Education Finance Association in 1997, and served on its Board of Directors from 1992 to 1995. He serves on the editorial board of the *Journal of Education Finance*. Dr. Fowler is a graduate of Columbia University with a doctorate in education (1977).

Introduction and Overview

William J. Fowler, Jr.

National Center for Education Statistics

The education finance scholars that assembled for the 1997 National Center for Education Statistics (NCES) Summer Conference brought with them diverse views of the effects of money on elementary secondary student outcomes. Some were sanguine because they had suggested new or alternative statistical or research designs from traditional "production function" studies that yielded empirical evidence of the positive effects of resources on student achievement. Others had examined in detail how a few schools with good student outcomes reallocated resources, in an effort to understand how changes in resources are linked to student effects. Still others demonstrated what they believed to be a consistent set of evidence that disadvantaged students who received more resources demonstrated higher gain scores. At least one scholar, however, found a large urban school district where, as a result of alleged mismanagement, high per-pupil spending was not reaching the classroom and student outcomes seemed below those of comparable school districts. Another academic sought to design a school finance system that would provide an "adequate" level of education, where adequacy is defined in terms of minimum standards of student performance.

Not all of the presentations involved finance and student outcomes. A few researchers were involved in what might be considered the "cutting edge" of education finance research, rather than reexamining the much-debated, albeit extremely relevant, question of the nexus of financing and student outcomes. One handful of researchers aspired to revolutionize the traditional manner of measuring fiscal equity. Their approach when measuring fiscal equity makes adjustments to equity measures for differences in geographic costs and the educational needs of students in a school district. Another set of enterprising scholars endeavored to devise a mechanism for collecting school-level financial data with the next administration of the NCES Schools and Staffing Survey (SASS), which is scheduled to be administered for the fourth time in 1999–2000. Should they succeed, the SASS finance survey would represent the first collection of traditional finance data from a nationally representative sample of private schools in 20 years, and the first ever for public schools (at the school level).

The first three papers presented in this collection of the proceedings reflect "real world" examinations of finances and student outcomes, rather than traditional education finance research designs that employ sophisticated statistical analyses of large-scale data bases that merge thousands of school district finances with thousands of students' achievements. While many in the statistical and research community would desire statistical support and replication for the "real world" findings, other researchers point out the value of qualitative and case studies in gaining insights that the "black box" statistical analyses cannot address.

David Grissmer, Ann Flanagan, and **Stephanie Williamson,** from RAND, pose the intriguing question of whether money matters for minority and disadvantaged students. They argue that evidence is accumulating that may replace the "Money Doesn't Matter" hypothesis. This new hypothesis asserts that additional money matters for students from less advantaged backgrounds and minority students, but may not matter for students from more highly advantaged backgrounds. They first explain the evidence they see in the NCES National Assessment of Educational Progress (NAEP) concerning resource growth and targeting. They then discuss large score gains of black students and class size changes. Finally, they look to new, experimental studies, rather than the quasi-experimental research that has been the cornerstone of the traditional education finance findings.

Grissmer, Flanagan, and Williamson argue that the widely-accepted evidence that real per-pupil resources doubled in education from the late 1960s to the early 1990s, while NAEP scores stagnated, is incorrect for a variety of reasons. First, disaggregating the national average scores suggests that scores for all racial/ethnic groups rose in reading and mathematics for all age groups. Second, using a different deflator from the Consumer Price Index (CPI) suggests the real increase (that is, after inflation) in educational expenditures was much lower than a doubling of resources. Third, most of this increase went for disabled students, who are not included in NAEP. Fourth, of the resources not directed toward disabled students, a disproportionate amount of resources was directed at minority students and students in poverty.

Since Grissmer, Flanagan, and Williamson assert that family changes can only explain one-third of the NAEP gains of black students, they examine what happened in the educational system. During these years, preschools and kindergartens flourished, desegregation occurred in the South, class sizes decreased, and teachers' age, experience, and education increased. Grissmer, Flanagan, and Williamson reject changes to preschools and kindergartens as having the kind of sustained gains in the test scores of black students they observe. Desegregation explains some of the gains, but not all. Instead, they believe that the school changes are the better candidates for explaining substantial parts of the NAEP gains, and of those, class size decreases seems the strongest.

Grissmer, Flanagan, and Williamson contend that class size effects have the virtue of experimental evidence supporting their relationship just as quasi-experimental research techniques are being called into serious question. The large, multi-district study in Tennessee where students were randomly assigned to smaller classes found significant positive effects on achievement, and larger effects for black students. Unfortunately, in the Tennessee experiment, students were returned to large classes after third grade, so we do not know what would have happened if students had remained in small classes until the end of school. In addition, class sizes fell in the 1960s, as well as the 1970s. If smaller classes had conferred long-term benefits, 17-year-olds who entered school in 1968 should have outscored those who entered in 1960, but this did not occur.

Karen Hawley Miles, an independent education consultant, and **Linda Darling-Hammond,** from Teachers College, Columbia University, find little attention has been given to rethinking the use of existing instructional resources, especially teachers who are schools' most important and expensive resource. Miles and Darling-Hammond examine five schools that demonstrate that it is possible to support student achievement at extraordinarily high levels by reallocating instructional resources to maximize individual attention for students and learning time for teachers. They assert that it is unlikely that schools can find ways to create more individual time for students or more shared planning time for teachers without prohibitively raising costs, unless they rethink the existing reorganization of resources.

Miles and Darling-Hammond suggest that they focus primarily on the assignment and use of teaching staff because it is the most sizable and the most underexplored area for potential resource reallocation. They cite studies that demonstrate that few of new teaching staff were deployed to reduce class sizes for regular education students; most went to provide small classes to the growing number of special students, or for teacher release time. Only 43 percent of school staff are regularly engaged in classroom teaching, in comparison to 60 percent or more in European countries, enabling more time for collaborative planning and professional development.

Miles and Darling-Hammond describe six practices widely found in schools, and portray the impact of each on the use of teaching resources:

- Specialized programs conducted as add-ons;
- Isolated instruction-free time for teachers;
- Formula-driven student assignments;
- Fragmented high school schedules and curriculum;
- Large high schools; and
- Inflexible teacher work day and job definition.

The schools attempting to change these conditions used several strategies of resource reallocation. Reallocation involved reduction of specialized programs, more flexible student groupings, longer and varied blocks of instructional time to create more personalized environments, expanded common planning time for staff, and creative work schedules and staffing roles.

Only two of the five schools Miles and Darling-Hammond studied actually reallocated and restructured existing programs and staff, while the others were brand-new schools which did not suffer from the six standard practices found in schools. Three were elementary schools, and two were secondary schools. Traditional elementary schools served regular education students in age-graded, self-contained classrooms. Three-quarters of the school's teaching staff worked with regular education students, the remainder with Title 1 and special education students who were pulled out of their regular classes for such special instruction. Class composition and class size stayed the same all day, for all subjects, except for special instruction. The elementary classroom teacher instructed all subjects except specialties like art, music, and gym, which were taught by specialists during the classroom teacher's free period. Teachers had 45 minutes 3 to 5 times a week free from instruction for planning, uncoordinated with other teachers' free time.

The high-performing schools changed this organization, increasing the percentage of teachers who worked with heterogeneous groups of students to 90 percent. The anomalous elementary schools teachers' adapted instructional grouping to student needs. These atypical schools kept teachers with the same students for 3 years, usually with the same homeroom class. Some teachers received as few as nine new students a year. All the elementary eccentric schools created more common planning time, although only one made dramatic changes. These novel elementary schools created master-teacher and other instructional adults in the classroom. Similar changes occurred in the two high schools studied.

To accomplish these changes, the uncommon schools directly challenged policies, regulations, and collective bargaining agreements. Changing school organizations to better fit an instructional vision does require schools to confront a host of obstacles. However, the biggest constraint may be a lack of vision. The sample schools described by Miles and Darling-Hammond are intended to assist those who lack a vision of what may be done.

Joyce Ladner, a member of the District of Columbia Control Board, describes the condition of the District of Columbia public schools. The Control Board concluded that the D.C. public schools were in crisis, by every important educational and management measure. As seen by the Control Board, the District of Columbia Public Schools (DCPS) were simply failing in their mission to educate the children of the District of Columbia, by neither providing a quality education nor a safe environment in which to learn.

The Financial Authority was created by the U.S. Congress in 1995 to repair the District of Columbia's failing financial condition and to improve the management effectiveness of government agencies. In November 1997, the Authority removed the superintendent and stripped the Board of Education of most of its power to control the schools. In their place, the Authority appointed a new Chief Executive Officer and an Emergency Board of Trustees.

Ladner describes DCPS from a report entitled *Children in Crisis: A report on the failure of the D.C.'s*

Public Schools. The DCPS' students score significantly lower on standardized academic achievement tests than their peers in comparable districts around the nation. In 1994, only 22 percent of DCPS' fourth-grade students scored at or above the basic NAEP level—a decrease of 6 percent from 1992. Many students are dropping out or leaving DCPS (40 percent) for neighboring districts and private schools. DCPS teachers in NCES' School and Staffing Survey (SASS) believed that a variety of serious problems affected their schools, more than in other states. These problems included student unpreparedness; disrespect for teachers; absenteeism; and apathy. Compared with the national average, more DCPS teachers and students report being threatened with violence. The infrastructure of the District's public schools is collapsing: boilers bust, roofs leak, firedoors stick, bathrooms crumble, and poor security permits intruders. The system does not know how many students it has, with estimates varying from 65,000 to 81,000.

Comparisons of the District's school expenditures with other jurisdictions are difficult because of the above managerial irregularities, which extend into the fiscal area. However, the District's per-pupil expenditure exceeds the national average, and is substantially higher than many comparable urban school districts and neighboring districts (one exception is Newark, New Jersey, which spent $2,512 more per student than the $7,655 DCPS spent in 1994–95). DCPS employs 16 teachers for every central administrator employed, compared with its peers who employ 42 teachers for every central administrator. In 1996, DCPS allocated more toward its Office of the Superintendent than the Fairfax County, Montgomery County, and Baltimore City public school systems combined. It also spent more than twice as much on the Office of the Board of Education than peer and neighboring district averages.

Ladner concludes with some of the steps of the new management team, such as imposing a hiring freeze, closing 11 schools and replacing 50 roofs, increasing security with security guards and new metal detectors, establishing a teacher evaluation program, ending social promotions, terminating a large contracted school maintenance contractor, and initiating new contracted services for school breakfasts and lunches.

While the above studies described "real world" evidence of how school districts deploy staff resources, three educational researchers conducted more traditional production-function studies from large databases. These studies use econometric techniques to find relationships between educational outcomes to school resources, while statistically controlling for student background characteristics. In the first of these, **Corrine Taylor**, University of Wisconsin, argues that none of the previous studies adequately accounted for geographic cost variations nor the costs created by the proportion of students with special needs. She hypothesizes that a stronger relationship between student achievement and school expenditures will emerge once these costs are taken into account.

Taylor employs NCES data sources: the National Education Longitudinal Study of 1988 (NELS); the Common Core of Data (CCD); and a district-level teacher cost index (TCI). NELS contains a nationally-representative sample of students who were followed at the 8th, 10th, and 12th grade level in 1988, 1990, and 1992, who took cognitive tests, and who completed questionnaires about a wide variety of education issues and student's interest and effort in school. Their parents, teachers, and school administrators also completed questionnaires regarding SES and school conditions. The TCI is a geographic cost adjustment that is an index that reflects the cost of employing teachers in particular regions of the country, based upon job and location amenities. The SASS provided the data to apply a regression analysis of the factors that influence teacher salaries, including those that are under the control of school districts, such as teacher experience and degree status, and those that are not, such as cost of living and quality of life. The estimates of these characteristics on teacher salaries results in an index number, where 100 is the national norm, that can be used to estimate teacher costs, holding constant discretionary factors. TCI index numbers are available at the state, county, and school district levels (Chambers and Fowler, 1995).

The paper does not mention the extensive work that is required to combine the data sets, NELS, CCD, and TCI. That integrating the data sets is problematic may be deduced by Taylor's description of her sample of students, she includes only public school students who participated in all three waves of the National Education Longitudinal Study of 1988 (NELS) study (11,598 students), who never dropped out of school (11,503), and who attended the same high school in both 1990 and 1992 (11,167). These restrictions were necessary if one wishes to consider only those students who are consistently associated with school resources at particular schools.

Taylor is very comprehensive. She examines three expenditures per pupil: total current; core; instructional salaries. She then cost adjusts each for both geographic differences and student need differences. The geographic cost adjustment is accomplished by dividing each nominal per-pupil expenditure by the TCI (times 100). Student educational need is measured by including separate control variables for the proportion of students in special education, limited English proficiency, and compensatory education, based on information from the Common Core of Data (CCD) school district data.

Using the 1992 mathematics score as the dependent variable, she includes prior achievement, student and family characteristics, student interest and effort, student view of the school environment, peer's characteristics, special needs students, community characteristics, and school characteristics. Although she finds the consistently positive and statistically significant effects of per-pupil expenditures on high school students' academic achievement, the effects do not increase appreciably when per-pupil costs are adjusted for geographic or student need. She concludes that these results demonstrate that the lack of a strong relationship between student achievement and school expenditures cannot simply be attributed to mismeasurement of the school's fiscal resources.

In yet another production function study, **Harold Wenglinsky**, of ETS, uses a different NCES national sample of student achievement, the National Assessment of Educational Progress (NAEP). He applies structural equations and hierarchical linear modeling (HLM) in an attempt to address what he believes to be the shortcomings of standard production function studies. Reviewing some 30 years of production-function studies in education (approximately 400 studies), he restates the imperfections in most of the studies: the data were not nationally representative, but instead use data from a particular state or school district; the usual measure was per-pupil expenditure, rather than more discrete measures, such as administrative overhead, or per-pupil instructional expenditure; the process of schooling was not considered, which may mediate the relationship between expenditures and student outcomes; measures of student background were inadequate; differences in costs caused by geography were not considered; especially in the early studies, measures of outcomes were unsophisticated, such as graduation rates.

Wenglinsky attempts to remedy these problems in his study via differences in the nature of the data base he employs and the nature of the analyses undertaken. He uses a NCES data base, NAEP, that contains a nationally-representative sample of student and school information from 4th-, 8th-, and 12th-graders, and information from their teachers and principals. The subject areas tested vary, but have included at various times mathematics, reading, history, geography, and science. Wenglinsky uses the 1992 mathematics assessment of students attending fourth grade, which contains measures of mathematics achievement, school environment, teacher education levels, teacher-student ratios, and student and school-level SES. He combines this data set with the CCD fiscal data for school districts. Note that this yields the same expenditure for every child in a school district, regardless of the school that they attend. Wenglinsky then adjusts these expenditures using the state-level cost adjustment, the TCI, rather than the school-district-level TCI adjustment that Taylor uses. For states which experience large within-state differences, such as New York, or Illinois, use of the state geographic cost adjustment will be less precise than the use of a county or school-district measure.

Wenglinsky links these data bases to yield a district level and student level. The district level was produced by aggregating NAEP data to the district level and linking it to the district-level CCD. Since NAEP is a sample of students, only the 203 school districts that could be matched were included in the analysis. The district-level database was used for all analyses except a multi-level approach. Let us explore for a moment why some multi-level approach may be desirable.

A common dilemma of education researchers is that the data often come from hierarchical data structures. For example, expenditure data are from the school district level, the teachers are employed and conduct their classes at the school level, and students' achievements are at the student level. Since traditional statistical techniques for modeling hierarchy have been inadequate, the usual choice has been to ignore these differences, and to combine all characteristics at one level of aggregation. Assume this is done in work similar to the Wenglinsky paper. The resulting data set may be at the student level, with a similar expenditure for every child in the same school district, and a similar school environment measure for every child in the same school (or aggregated in the opposite direction, so that there is only a average student achievement score at the school district level). In reality, we know that every child in a school district receives different allocations of resources, so the per-pupil expenditure should vary for *every* student, and the environmental measures should vary for every classroom (if not for every student). Ignoring these limitations have resulted in a variety of statistical problems, which make such studies vulnerable to legitimate criticisms by other education researchers.

Recent developments (such as Hierarchical Linear Modeling (HLM)), however, have led to the development of several approaches to analyzing hierarchical data sets, in which the researcher may retain data at the appropriate level, and then run analyses that compare these attributes properly (Bryk and Raudenbush, 1992). As will be discussed later, NCES is exploring the possibility of collecting a student-level resource measure with other student-level data in its sample surveys.

Returning to a discussion of the Wenglinsky paper, he uses two statistical analyses: LISREL, in which the educational researcher specifies how he believes each variable effects others; as well as HLM. Wenglinsky hypothesizes that a student's academic achievement is modified by the school environment and the teacher's highest degree and the student-teacher ratio, as intervening variables between the school district's resource choices and student achievement. He also examines more discrete expenditures than the total per-pupil expenditure, utilizing the school district instructional per-pupil expenditure, the central administration per-pupil expenditure, and the school administration per-pupil expenditure. The HLM analysis consisted of student achievement as the dependent variable and two resources (teachers' highest degree and teacher-student ratios) as independent variables. Wenglinsky finds that expenditures on instruction and central office administration affect teacher-student ratios, which, in turn, affect student achievement. The relationships also persisted when subjected to multilevel analysis using HLM. Interestingly, unlike Taylor, Wenglinsky finds that the relationships were affected by modifying the expenditures for geographic cost differences.

Andrew Reschovsky and **Jennifer Imazeki** of the University of Wisconsin-Madison explore the quandary of developing a school finance formula that guarantees the provision of an adequate education to low-income students. Imazeki and Reschovsky recognize that the *cost* of education can be defined as the minimum amount of money that a school district *must* spend in order to achieve a given education outcome. In comparing two districts with equal spending per pupil, educational performance may be lower in one of the districts if the *costs* of providing any given level of education are higher in that district, or if that district is more *inefficient* in its use of resources.

They stress that the importance of costs in any discussion of equity in the financing of schools is that the achievement of equity (in outcomes) will require higher spending in districts facing high costs. The courts are moving from a focus on equity in spending to one of educational *adequacy*, where adequacy is defined in terms of minimum standards of student performance. Imazeki and Reschovsky believe a prerequisite for designing a outcome-equitable school finance system is knowledge about how much it will cost each school district to provide an adequate education for its students. In their paper, they review traditional school aid distribution formulas, as well as other cost measures, and then go on to develop their own cost index for school districts in Wisconsin, which takes into account student educational needs. They then develop a simulation of a school aid formula designed to achieve education adequacy.

The traditional way that states finance the education of students with special educational needs is by "weighting" them; that is, if a school district receives $1,000 for a regular student, a handicapped student might generate $2,300 in state aid for the school district, or 2.3 times as much. These weights typically have been derived from episodic studies of a few school districts or states where information exists regarding what some school districts spend for the education of such children. Other geographic cost indexes, such as McMahon or Chambers, do not consider student educational need in an explicit way. As such, they understate the costs of some school districts, since some school districts will have to hire more teachers (perhaps at a premium) and spend more on non-teacher resources (social workers, drug counselors) in order to achieve any specific educational goal. Indeed, even more sophisticated efforts (such as Duncombe, Ruggiero, and Yinger, 1996) that include student need typically measure the cost of purchasing a given set of inputs to be used in providing the education of students.

Imazeki and Reschovsky specify a regression equation where student outcomes are a function of school resources, the characteristics of students, and the family and neighborhood. They consider such student need variables as the percent of students with disabilities (and severe disabilities), and the percent of students eligible for free and reduced-price lunch. They also use a "value-added" measure of student achievement; that is, the change in test scores over time. Because of the complexities involved, Imazeki and Reschovsky decided not to include a measure of efficiency. As has been previously found, they find that there is a "U-shaped" relationship between spending per pupil and school district size, and, as expected, higher proportions of students from poor families and those with disabilities are associated with higher costs.

Setting the tenth-grade score at the average for all Wisconsin districts as the adequacy standard, Imazeki and Reschovsky construct a cost index by using the results of the regression to predict hypothetical spending for each district. These predictions are then compared to actual spending in a average district with average costs and average levels of educational outcomes. They then go on to develop a state-aid formula to fund the "adequate" level. Surprisingly, while per pupil aid remains substantially higher in low-property wealth districts as compared to high-property wealth districts, the largest percentage increases in aid go to high-wealth school districts.

Let us now turn to two presentations that did not involve finance and its relationship to student outcomes. These researchers were involved in what might be considered the "cutting edge" of education finance research by examining the effect on traditional fiscal equity measures of applying geographic cost adjustments and student need adjustments. **Lauri Peternick** and **Becky Smerdon**, American Institutes for Research, **William Fowler**, NCES, and **David H. Monk**, Cornell University, were struck by the dramatic differences in the coefficient of variation (CV) when geographic cost adjustments were applied to nominal per-pupil expenditures. Using New York State school districts' expenditures per pupil from the CCD, they examined financial equity within the state by conducting two sets of analyses, including and excluding New York City. One set of per-pupil expenditures were nominal, another were adjusted for student needs, a third used the geographic cost adjustment of the TCI, and a fourth used both student needs and geographic cost. Student needs used 2.3 weights for students with an "individual educational plan (IEP)," and 1.2 for students at-risk (in poverty) and limited English proficient (LEP). Four equity measures were examined: the CV; the Gini coefficient; the McLoone Coefficient; and the slope.

Peternick, Smerdon, Fowler, and Monk find the CV is greatest when measuring nominal per-pupil expenditures. Employing a geographic cost adjustment reduces the CV, as does the needs adjustment. Applying both adjustments almost halves the CV. The Gini is similarly affected. The McLoone Index, which measures equity for the lowest half of the distribution, however, demonstrates the largest inequity with a geographic cost adjustment. Including New York City in the analysis, the nominal data show increased equity. The opposite occurs when the per-pupil expenditures are adjusted for geographic cost differences or needs.

The slope demonstrates the relationship between median household income and per-pupil expenditures. Cost adjustments hinder the explanatory power of median household income (or housing value), while student needs adjustments serve to increase income's explanatory value. Median household income has a larger effect when New York City is included.

Peternick, Smerdon, Fowler, and Monk conclude that the results presented demonstrate the varying impact different adjustments may have on equity measures. In addition, the inclusion of a single large urban school district may have dramatic effects upon the results of equity measures.

The final article is from researchers seeking to address the demand for school-level resource data. Aside from the interest of parents and taxpayers about resource allocation and productivity of their school, there are questions of accountability and management, as well as equity and adequacy that are feeding the thirst for financial information at the school level. **Julia Isaacs** and **Michael Garet** of American Institutes for Research, and **Stephen Broughman**, NCES, are attempting to design a method to collect school-level financial data during the next administration of the NCES Schools and Staffing Survey (SASS), which is scheduled to be administered for the fourth time in 1999–2000. SASS provides nationally-representative and state-representative data about schools, and any financial data would permit baseline estimates of spending in the nation's schools. Presently, only a few states have accounting systems that extend to the school level, and the existence of more than 84,000 public schools in the country make it unlikely that any uniform reporting system would be quickly adopted. Most financial reporting is still contained at the school-district level.

Only two extant systems are in use when school-level fiscal resources are reported. The most commonly used is a simple extension of the existing accounting system to the school level. Coopers and Lybrand developed a software package of this type that the lay person could use to recode the school district budget to the school level. It contains algorithms to allocate expenditures from the school district level to the school for some functions (such as student transportation), using some information pertinent to the activity (such as numbers of student transported). Every state that has implemented school-level financial reporting has used the traditional accounting system extended to the school-level.

However, Isaacs, Garet, and Broughman describe an alternative approach from their AIR colleagues, called the Resource Cost Model (RCM). The RCM is a "bottom-up" approach to school resources, aggregating from the school-level the number of staff in certain assignments, and the time they spend in certain activities. Prices are then assigned to each person for each assignment. In this way, the "service delivery" system can be described, as can its cost. For example, two schools may give compensatory students additional instruction: one in a "pull-out" service delivery system; the other by having an aide assist the student. As one can imagine, the "pull-out" delivery system, where a student is sent to another class with another teacher, will be much more expensive than the simple assistance of an aide.

Isaacs, Garet, and Broughman have developed a proposal for collecting school-level financial data via a questionnaire to the school business official (who usually resides at the school district level). The business official would report school expenditures (if he has them), and expenditures at unspecified locations. These unspecified expenditures would then be prorated, using additional information needed for prorations.

A group of education finance experts convened by Isaacs, Garet, and Broughman in January, 1998, suggested that a synthesis of the two approaches be attempted. Work on refining the public school expenditure instrument is still underway.

One prospective note about a comment made earlier. Much of this volume revolved around the connection between per-pupil expenditures and student achievement, and the difficulties researchers encountered because financial data were at a higher level of aggregation than the rich student-level information that NCES obtains through its student-level surveys. The Education Finance Statistical Center (EFSC) within NCES is conducting work to see if a student-level resource measure can be developed in time to accompany a longitudinal study of students that will begin in 1999 for kindergartners, termed the Early Childhood Longitudinal Survey (ECLS). For the most up-to-date information on the work of the EFSC, NCES finance publications, finance graphics, and finance data sets, including those containing geographic cost adjustments, readers are urged to visit the web site http://**nces.ed.gov/edfin** where readers may also email finance questions, if they are not already answered in "frequently asked questions."

Does Money Matter for Minority and Disadvantaged Students? Assessing the New Empirical Evidence

David Grissmer

Ann Flanagan

Stephanie Williamson

RAND

About the Authors

David Grissmer is a senior management scientist at RAND. His education research has included work on teacher supply and demand and the analysis of national and regional test score trends. He is currently engaged in research analyzing state achievement scores, estimating the number and location of children at educational risk in the United States and assessing the federal research and development portfolio on children.

Ann Flanagan is a consultant at RAND. She is analyzing state test scores to determine the effects of family characteristics and educational policies and resources. Her previous research has included estimating the amount and kind of federal research and development on children as well as research on the incidence, effects, and prevention of drug use.

Stephanie Williamson is a research assistant at RAND. Her research areas have included analysis of national and regional test score trends and examination of the relationship between family characteristics and student academic achievement. Her work also includes analyzing reasons for the convergence of the black-white score gap. She is currently investigating academic skills at work, including the exploration of the relationship between various pregraduation experiences and a range of post high school outcomes.

Does Money Matter for Minority and Disadvantaged Students? Assessing the New Empirical Evidence

David Grissmer

Ann Flanagan

Stephanie Williamson

RAND

Until relatively recently, the consensus among social scientists was that providing schools additional resources would have little impact on student achievement—the so-called "money doesn't matter" thesis (Ladd, 1996). This counter-intuitive view actually dates from the "Coleman Report" which found family influence strong and little effects of school resources (Coleman et al., 1966). Influential reviews by Eric Hanushek (1989, 1994, 1996) also argued that evidence from over 300 empirical studies provided no consistent evidence that increased school resources raised achievement scores. While this view was consistently challenged by many educators, policymakers, and parts of the research community, the empirical evidence simply suggested otherwise.

This scholarly consensus began to crack in the early 1990s. Hedges and his colleagues conducted a formal meta-analysis of the studies that Hanushek had reviewed. They found that most of these studies lacked the statistical power to detect resource effects even when they were quite large. When Hedges and his colleagues pooled data from all available studies, the results indicated a positive, statistically significant effect and provided evidence that some programs may have large effects (Hedges et al., 1992; Hedges and Greenwald, 1996). Other work conducted with alternate methodologies like Hierarchial Linear Modeling rather than the "production function" framework used in the econometric community often showed positive effects of resources.[1]

Nevertheless, Hanushek made one argument that was hard to rebut. Measured in constant dollars, per-pupil expenditures (PPEs) doubled between the late 1960s and the early 1990s. Yet the National Assessment of Educational Progress Tests (NAEP) of representative samples of 9-, 13-, and 17-year-old children seemed to show little improvement during the period when resources rose so rapidly. The increases in reading and mathematics scores from the early 1970s to 1992 were between 0.1 and 0.2 standard deviation or about 4–5 percentile points.

However, accumulating evidence is now challenging both the NAEP evidence and the accuracy of previous empirical studies. The accumulative evidence is certainly sufficient to replace the "money doesn't matter" hypothesis with one that states that additional money matters for students from less advantaged backgrounds and minority students, but may not matter for students from more highly advantaged backgrounds. Several lines of research are converging toward this hypothesis. They include the following:

[1] For two recent examples see Gamoran, 1996 and Raudenbush, forthcoming.

- Re-analysis of experimental data on effect of class size
- Evidence that model specifications used in many previous studies involving non-experimental data have been flawed
- Evidence that from 1967–91, increases in available educational resources aimed at increasing regular students' achievement has been markedly overestimated
- Evidence that the more limited real resources available to increase achievement scores from the late 1960s to the early 1990s was disproportionately targeted at minority and lower income children
- Evidence that minority and less advantaged children made substantial gains in test scores in the 1970 to 1990 period, but more advantaged white students made only small gains
- Evidence that the timing of score gains of minority children seem to be related to both the civil rights and war on poverty efforts as well as declines in class size.

. . . evidence is now emerging which shows that disadvantaged students received the largest resource gains and that large score gains occurred among these students.

A more consistent set of evidence is now emerging which shows that disadvantaged students received the largest resource gains and that large score gains occurred among these students. We first discuss the evidence from NAEP scores and the companion findings concerning resource growth and targeting. We then discuss several hypothesis for large score gains among blacks in the 1970s and 1980s, and the correspondence with experimental data on the effects on class size. Finally we discuss why estimates on the effects of resources from non-experimental data are now being seriously challenged, and probably have to be discounted in favor of the experimental data.

Rising Resources and Rising NAEP Scores

The often-quoted evidence that real per-pupil resources doubled in education from the late 1960s to early 1990s while NAEP scores stagnated is flawed on four accounts. First, although mean NAEP scores did not rise much, this was partly because of rapid growth in the low-scoring Hispanic population. When disaggregated, scores for all racial-ethnic groups rose in reading and mathematics for all age groups. Non-Hispanic whites scores rose by smaller amounts, while scores for Hispanic and blacks rose dramatically. Second, the real increase in educational expenditures was far less than the CPI adjusted PPE data would indicate. Use of more appropriate indices for adjustment of educational expenditures due to their labor intensity provides much smaller estimates of real growth. (Rothstein and Miles, 1995; Ladd, 1996a) Third, a significant part of the smaller estimated increase went for students with learning disabilities, many of whom are not tested.[2] A significant part also went for other socially desirable objectives that are only indirectly related to academic achievement. Taking into account better cost indices and including only spending which would have been directed at increasing achievement scores, Rothstein and Miles (1995) concluded that the real increase in per pupil spending on regular students was closer to 30 percent than to 100.

Finally, the association of additional resources with increased test scores depends upon the distribution of the increased spending. The evidence in

[2] All sides agree that a disproportionate fraction of the expenditure increase during the NAEP period was directed toward special education (Lankford and Wyckoff, 1996; Hanushek and Rivkin, 1997). Hanushek and Rivkin estimate that about a third of the increase between 1980 and 1990 was related to special education. NAEP typically excludes about 5 percent of students who have serious learning disabilities. However, special education counts increased from about 8 percent of all students in 1976–77 to about 12 percent in 1993–94. These figures imply that 7 percent of students taking the NAEP tests were receiving special education resources in 1994, compared to 3 percent in 1976–77. This percentage is too small to have much effect on NAEP trends, but it should in principle have had some positive effect.

Rothstein and Miles (1995) shows that a disproportionate amount of resources was directed toward minority and lower income students.[3] Scores of minority students and lower scoring white students all showed large gains. The argument that additional resources did not matter is not applicable to these students. However, if significant additional resources were also directed toward advantaged students, the evidence would show much smaller gains, and the argument that "money doesn't matter" may apply to these students.

NAEP Data[4]

Trends. Figure 1 shows how black and white 17-year-olds' reading and mathematics scores changed between 1971 and 1996.[5] Figures 2 and 3 show the same data for 9- and 13-year-olds. Each score by race is relative to the earliest test score recorded, so a difference between black and white scores at a given year represents a change in the black-white score gap. The following points stand out:

- The black-white gap narrowed for all ages in both subjects due to substantial gains in black students' scores while white students registered smaller gains.
- The black-white gap narrowed the most for 13- and 17-year-olds due to dramatic increases in black scores from the late 1970s to the late 1980s when black gains were 0.6 to 0.7 standard deviation above white gains.

For 13- and 17-year-olds, the gap stabilized or widened in the 1990s due to significant declines in black reading scores and stable black mathematics scores, while white mathematics scores were increasing. By 1996 black students' gains were between 0.2 and 0.6 standard deviations greater than white students' gains. The black-white gap for 9-year-olds narrowed by 0.25 to 0.35 standard deviation by 1996. The pattern of gains among this group is quite different from that of black adolescents, and the pattern also differs somewhat for reading and mathematics. Black 9-year-olds gained more than older blacks during the 1970s and gained less than older blacks during the 1980s. Although reading scores among black 9-year-olds show declines after 1988, unlike adolescent reading scores, they have returned to 1988 levels. Additionally, mathematics scores continued increasing after 1988 among this cohort.

It is important to stress that even when black gains were largest, they never came close to eliminating the black-white gap. The largest reduction in the gap was for 17-year-olds' reading scores between 1971 and 1988. In 1971 the median black

[3] Rothstein's and Miles' data analyzed detailed data in only nine school districts. More national evidence is needed concerning the relative allocation of additional resources among different types of students. There is little doubt that many of the new programs which were initiated or expanded were directed toward minority or low-income children. These included compensatory education programs such as Title 1 and HEADSTART, efforts within states to change to more equitable funding formulas and desegregation initiatives. However, funding also may have increased for advantaged students. More direct evidence is needed from school-district-level analysis of funding trends for high and low income districts.

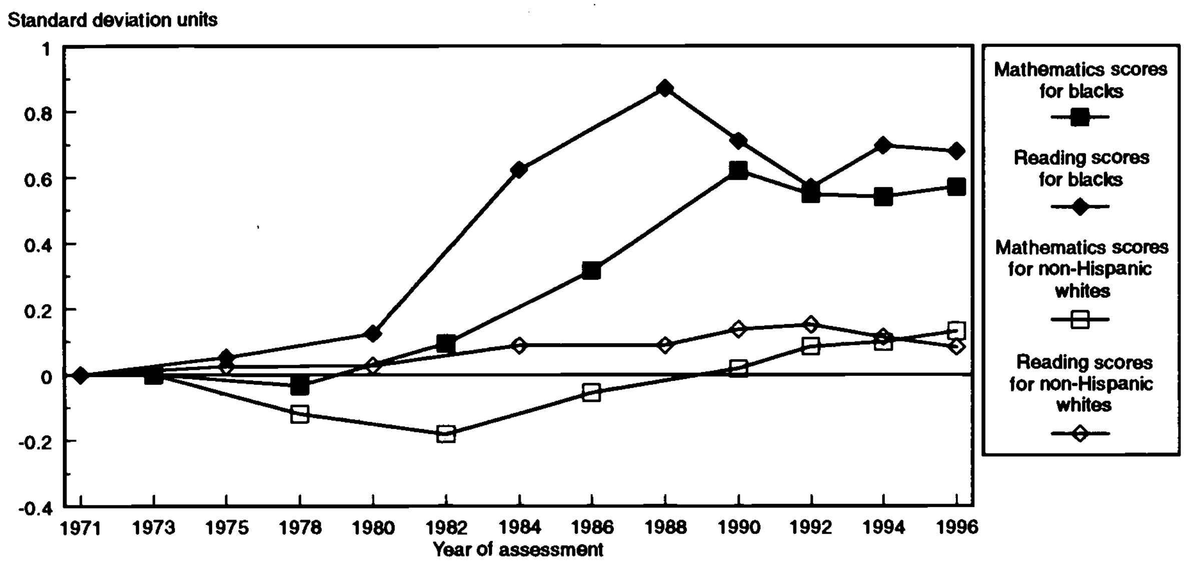
[4] See Cambell et al., 1994 and Cambell et al., 1996 for descriptions of the NAEP data and further references.

[5] The scores have been converted to relative scores by assuming the earliest test score for each race is zero. Thus, the difference in scores reflects changes in the black-white gap from the earliest test. The scores are converted to standard deviation units by taking the mean score difference from the earliest test and dividing by a metric that remains constant over the period—the standard deviations of all students for the earliest year. Another common practice is to measure the gap with respect to the standard deviation in the same year. Since the standard deviation for all students declines for mathematics scores, but increases for reading scores this method changes the metric over time and would result in a somewhat different measure of gap reduction.

The 1973 and 1971 scores for non-Hispanic white students were estimated because the only published scores are for combined Hispanic and non-Hispanic white students in those years. Tests after 1973 have separate data for Hispanic and non-Hispanic white students. We make a small correction in the 1971 and 1973 white data by determining the proportion by age group of students who were Hispanic and assuming that the difference between Hispanic and non-Hispanic white scores were the same in 1971 and 1973 as for the 1975 reading and 1978 mathematics tests.

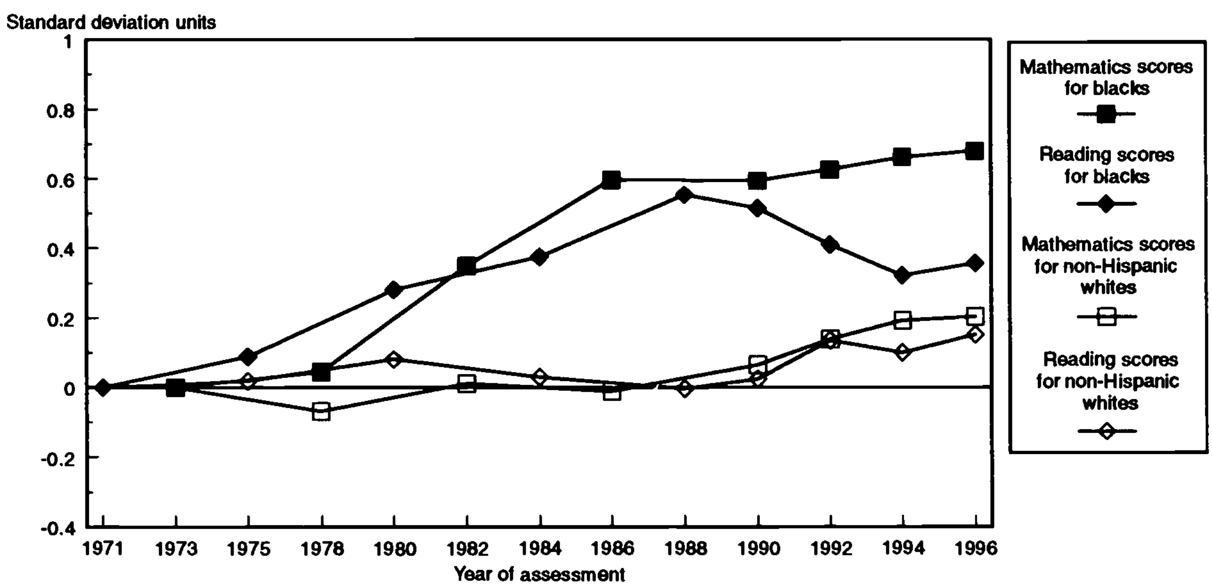
Also the 17-year-old NAEP scores reflect only students rather than all 17-year-olds. Consequently, the 17-year-old scores will be biased with respect to 9- and 13-year-old scores. We make a correction for 17-year-old scores using the proportion of 17-year-olds in school by race in 1971–73 and 1996. We assume that those not tested would have scored one-half standard deviation below the mean score for their respective race—probably a conservative assumption. School enrollment data from the October Current Population Survey (CPS) shows approximately 88 percent of white and 83 percent of black 17-year-olds were in school in 1970 versus 89 and 90 percent in 1988 (Cook and Evans, forthcoming).

Figure 1.—NAEP mathematics and reading scores for 17-year-old students, by race/ethnicity



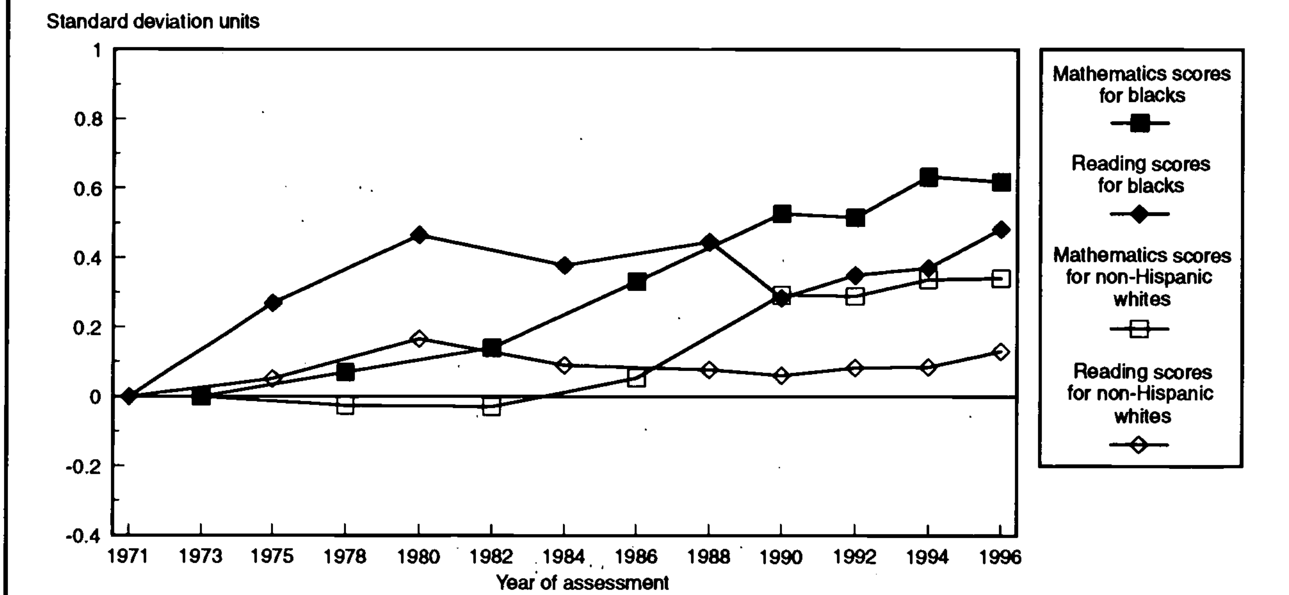
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).

Figure 2.—NAEP mathematics and reading scores for 13-year-old students, by race/ethnicity



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).

Figure 3.—NAEP mathematics and reading scores for 9-year-old students, by race/ethnicity



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).

scored between the 10th and 12 percentiles of the white distribution. By 1988 the median black scored between the 26th and 28th percentiles of the white distribution. For the other age groups, the gap remained even wider.

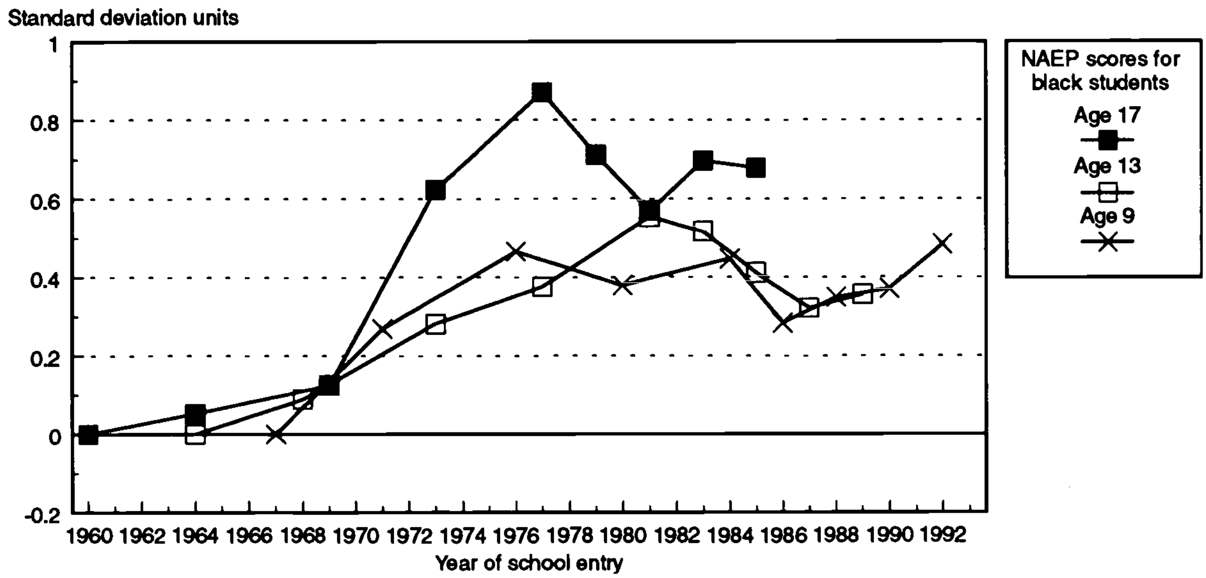
Cohorts. Reading and mathematics are ordinarily thought of as a cumulative process in which early gains are necessary before later gains take place. Charting scores by entering school cohorts tests for the presence this pattern. We characterize cohorts by the year in which they would normally have been in first grade, namely the year in which they were six years old. Each entering school cohort could have taken three tests in their school career—at age 9, 13, and 17. Figures 4 and 5 show each NAEP test score between 1971 and 1996 by the year of school entry. The scores for each age group are connected so that the pattern of increase by age within each entering school cohort can be more easily determined.

The following findings stand out:

- Black gains were small for cohorts entering school prior to 1968.
- The most significant black score gains occurred for cohorts entering school from 1968 through 1972 and 1976 through 1980.[6] After this period of rapid increase in both mathematics and reading scores, mathematics scores have stabilized while reading scores have declined.
- Except for 9-year-old mathematics scores, cohorts entering school after 1980 have registered no further score gains.
- Black reading gains precede mathematics gains. The data show small gains in reading for cohorts entering in the 1960s and dramatic gains for the 1968–72 cohorts. The mathematics data show no evidence of gains before the 1971 cohort.

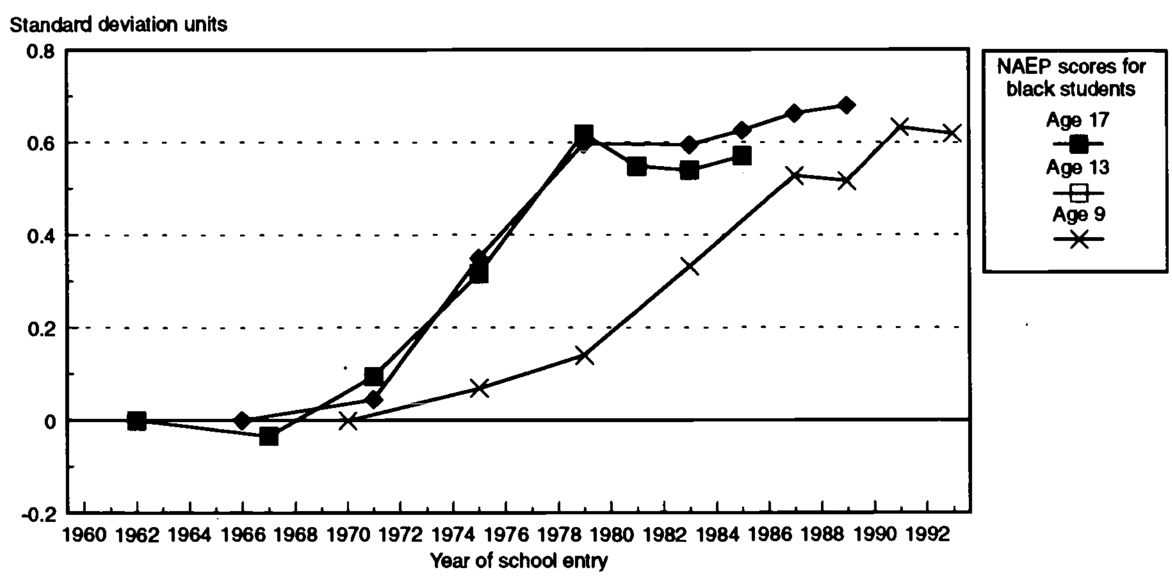
[6] For cohorts entering before 1968, we have no data on 9-year-olds. The position of the cohort curve is thus more uncertain for 9-year-olds than for the 13- or 17-year-olds—especially for reading. The large gain in 9-year-olds' reading between the cohorts entering in 1968 and 1972 may well indicate that there was also some gain at age nine between the 1964 and 1968 cohorts. If that were the case, and if we had the data, it would have the effect of raising all subsequent 9-year-old points in the cohort graphs and make the 9-year-old patterns closer to the pattern for older groups. For mathematics, however, 9-year-olds in the cohorts that entered in 1970 and 1975 scored at about the same level, making earlier gains appear less likely.

Figure 4.—NAEP reading scores for black students, by year of school entry



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).

Figure 5.—NAEP mathematics scores for black students, by year of school entry



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).

While the data does show strong cohort patterns, it also indicates that scores can increase at later ages above gains achieved at earlier ages.

Regions. The regional data shows that significant black gains and black-white gap reductions occurred in all regions for each age and subject, although some regional differences do exist (See figure 6 and 7).[7] Black score gains were somewhat larger in the south and west, although the reduction in the black-white gap was fairly uniform across regions.

Taken together, the NAEP data raise a number of questions:

- Why did both black and white scores rise for all ages in both reading and mathematics?
- Why did black scores rise substantially more than white scores at all ages and in all subjects?
- Why were black gains mainly concentrated for cohorts entering school between 1968–72 and 1976–80?
- Why did older black students gain and then lose more than younger black students?
- Why did black reading gains precede black mathematics gains?
- Why did significant black gains occur in all regions of the country with somewhat higher gains in the south and west?
- Why were black-white gap reductions fairly uniform across regions?
- Why did low-scoring students gain more in mathematics and less in reading than higher scoring students, regardless of race?

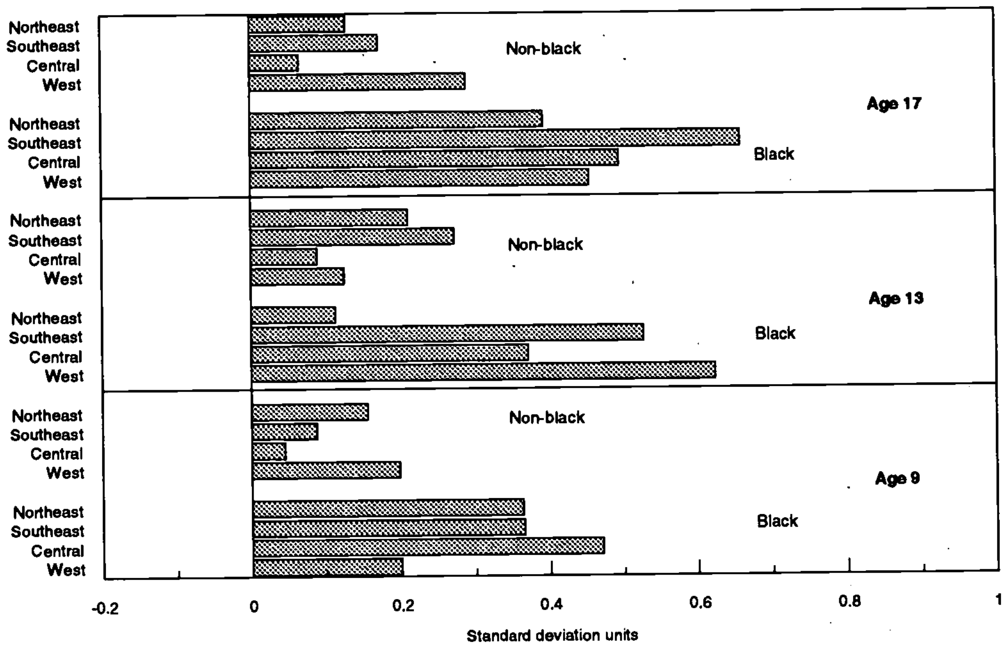
Early childhood interventions are widely thought to have the largest potential effect on academic achievement, partly because of their influence on brain development.

The most striking feature of the NAEP results for blacks is the size of adolescents' gains for cohorts entering from 1968–72 to 1976–1980. These gains were 0.6 standard deviation across subjects. Such large gains for very large national populations over such short time periods in tests similar to the NAEP are rare, if not unprecedented. Scores on IQ tests given to national populations seem to have increased gradually and persistently throughout the twentieth century, both in the United States and elsewhere (Flynn, 1987; Neisser, in press). While evidence exists for large gains on the RAVENS test which measures a narrower ability than tests like the NAEP, the gains on tests similar to the NAEP have averaged about 0.02 standard deviations per year—a fraction of the black rate in the 1980s. Neither these gradual, persistent gains in IQ scores cannot be explained, nor can it be explained whether the gains are larger for minority or other subgroupings of the population (Flynn, 1987). But no evidence exists in this data involving large populations showing gains of the magnitude made by black students over a 10-year period.

It is even unusual to obtain gains of this magnitude in intensive programs explicitly aimed at raising test scores. Early childhood interventions are widely thought to have the largest potential effect on academic achievement, partly because of their influence on brain development. Yet only a handful of "model" programs have reported gains as large as half a standard deviation (Barnett, 1995). These were very small-scale programs with intensive levels of intervention. Even when early childhood programs produce large initial gains, the effects usually diminish over time. Among blacks who entered school between 1968 and 1978 gains were very large among older students and were not confined to small samples, but occurred nationwide.

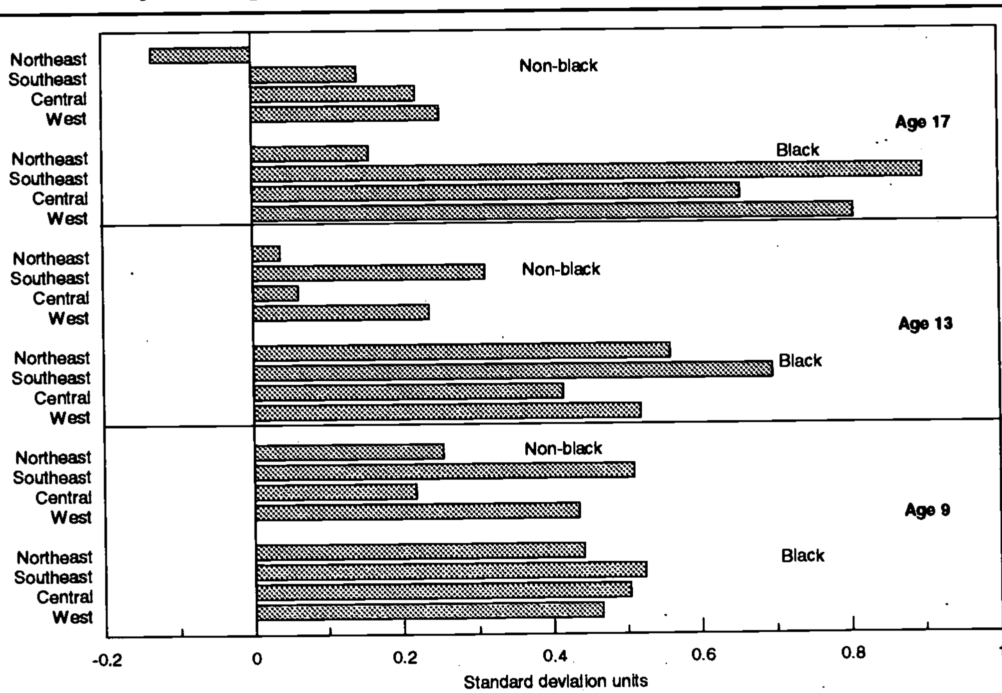
[7] The race x region data is unpublished and was provided by Michael Ross of the National Center for Education Statistics. We only have regional data for 1971–92. For the purpose of reporting NAEP data, the Department of Education places Texas in the West whereas, the U.S. Bureau of the Census places Texas in the South. This is important when interpreting black regional scores in the West, because a sizable proportion of these blacks are in Texas.

Figure 6.—Change in NAEP reading scores between 1971 and 1992, by region, race/ ethnicity, and age



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).

Figure 7.—Change in NAEP mathematics scores between 1971 and 1992, by region, race/ ethnicity, and age



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).

Large changes in scores of 0.5 standard deviation and more which are sustained through older ages have been observed when sustained interruptions in schooling occurs at younger ages (Ceci and Williams, 1997). Black students typically gain about 0.4 standard deviations per year on the NAEP tests between the ages of 9 and 13. In terms of "grade equivalents," black adolescent gains were equivalent to approximately 1.5 years of additional schooling. The large black gains sustained for older students suggests that there may have been a major change in the quality of blacks' school experience beginning in the late 1960s. This change in school experiences could reflect social and legal changes aimed at equalizing educational opportunity, additional educational resources that were especially helpful for black students, and the implementation of civil rights legislation creating new job opportunities for academically successful blacks, which may have made black students more eager to take advantage of any opportunities their schools provided.

However, before testing more specific hypothesis about changes in schools, we need to account for how changes in families may have affected test scores. Family characteristics account for the largest part of the variance in test scores in cross sectional models, and family characteristics changed significantly in this period. Thus, it is important to estimate how changing families would be expected to change achievement scores.

Family characteristics account for the largest part of the variance in test scores in cross sectional models . . .

Family Changes

The available evidence would indicate that changes in the family would be expected to have a positive effect on test scores from 1970 to 1990 (Grissmer, et al., 1994) (Cook and Evans, 1998). Higher parental education and smaller family size are the main factors leading to higher predicted test scores of approximately 0.2 standard deviation for black and white students. The sizes of the predicted effects are about the size of the white score gains, but much smaller than the score gains of blacks.[8] While these family gains can account for nearly all white score gains, they can explain only approximately one-third of black gains during the NAEP years.

Therefore, we must turn to events in the educational system: the growth of preschools and kindergartens, desegregation in the South, declines in class size, increases in teachers' age and experience and increases in the amount of teachers' education. Some of these factors would be expected to affect scores only at certain ages or for certain subjects or primarily in certain regions of the country, while others could potentially affect scores nationally at all ages in both subjects.

Changes in Schooling and Educational Resources

An assessment of the impacts of these factors on NAEP scores discounts many of them as substantial contributors to the overall black gains for all age groups (Grissmer et al., 1998).[9] In examining each

[8] Even these modest estimates of the gains attributable to improvement in family background may be too high. Consider the case of parental education. Parental education is correlated with children's test performance for two reasons. First, education changes parents in ways that make them more likely to provide their children with an environment conducive to learning. Second, education is a proxy for innate characteristics of parents that they pass along to their children. These innate characteristics also enhance children's test scores. When parents stay in school longer, their child-rearing practices probably change, but their innate characteristics do not. Keeping parents in school longer is, therefore, unlikely to raise children's test scores as much as we would expect on the basis of the cross-sectional estimates.

[9] This section presents a summary of much more detailed evidence provided in Grissmer, et al. (1998) for the size of expected effects from the changes in schooling and education cited in the rest of the article.

factor the evidence was assessed for how much a factor changed during the NAEP period, how many youth experienced the change and whether it changed more for blacks, how large the expected effects might be and whether they might be larger for black students, and how well the changes match the changes in NAEP scores.

Kindergarten attendance also increased during this period because of state mandates. A proxy measure of increasing attendance is the percentage of 5-year-olds in school (either kindergarten or first grade). About 66 percent of the entering class in 1960 were in K–1 at age 5 versus 89 percent for the entering class of 1990.[10] There is also a shift toward full- rather than half-day attendance. In 1970 only 12 percent of 5-year-olds attended for a full day, versus 41 percent in 1991.[11] Finally, black participation has increased somewhat faster than white participation. In 1969, 78 percent of white and 67 percent of black 5-year-olds were in K–1 while the percentages were almost equal in 1990.[12] Once again, strong differential effects by race would be required to affect the black-white gap.

The empirical evidence suggests that the large growth in pre-school participation and kindergarten may have a limited impact on 9-year-old scores, but would not significantly impact scores at ages 13 and 17 . . .

The empirical evidence suggests that the large growth in preschool participation and kindergarten may have a limited impact on 9-year-old scores, but would not significantly impact scores at ages 13 and 17 (Barnett, 1995) (Karweit, 1989). For pre-school the evidence is much stronger due to several methodologically strong studies. The evidence shows that small-scale, intensive interventions can have effects of 0.5 standard deviation or more in the short term, but its effect lessons for almost all studies measured at age 9 or older. Large-scale public programs show even weaker effects with similar fade-out.

The kindergarten evidence is based on differences between half-day and full-day attendance which show significant short-term effects, but similar fade-out effects. Larger short-term effects might be expected from the change from no attendance to half-day attendance, but it would counter the available evidence from both pre-school and full day kindergarten attendance for long-term effects to result from such a change. Similar to pre-school, more kindergarten participation might have residual effects to age 9, but no longer-term effects would be expected.[13]

The percentage of black high school graduates completing a minimum set of specified courses (4 years of English, 3 years of social science, 2 years of science and 2 years of mathematics) increased from 32 percent in 1982 to 76 percent in 1994. However, gains for white students were similar—from 33 to 76 percent. Changes in course work at the high school level in the 1980s may explain part of the score increases for older black and white students, but probably cannot explain much of the differential black score gains since course work changes were similar for black and white students.

Desegregation also appears to offer an explanation for a small part of the black score gains for all ages. Desegregation occurred primarily in the south over a short period in the late 1960s and early 1970s, but the regional NAEP data shows that black score gains occurred in all regions of the country. While the largest gains appear to have been made in the south, the extra southern gains accounts for less than 20 percent of overall black gains.

[10] Ideally, a measure of the percentage of those entering first grade who attended kindergarten is needed. By 1990, over 98 percent of children entering school attended kindergarten. However, such data is not available for earlier years. Our measure does not approach 98 percent because an increasing number of children in the 1980s delayed school entry for one year and attended kindergarten at age 6.

[11] Data from *Digest of Education Statistics*, 1992, table 47.

[12] *Digest of Education Statistics*, 1970, table 4 and *Digest of Education Statistics*, 1992, table 7.

[13] Similar to preschool also, some small-scale, specially designed kindergarten programs appear to have substantial short-term effects (Karweit, 1989). It is possible that kindergarten curriculum has shifted with some effects at age 9 as well.

Nationwide, three significant changes took place in schools during this period—lower class size, and better educated and more experienced teachers. Unlike many of the changes cited above, these changes, would have been experienced by nearly all students at all ages. Thus, if such changes would be expected to have effects on achievement, these changes would better explain NAEP score gains for all age groups and subjects.

Teachers' average level of experience declined in the early 1960s as substantial numbers of inexperienced teachers were hired to teach the baby boomers. As enrollments fell in the 1970s, the flow of new, inexperienced teachers was substantially reduced; the average experience level of teachers grew substantially from 1970 to 1990 (Grissmer and Kirby, 1997), and, by 1990, a significant number of teachers had 20 or more years of experience. We assume here that teachers as a group are most productive between 5 and 20 years of experience. Figure 8 shows the average changes in the percentage of teachers with 5–20 years of experience who would be teaching an entering cohort of age 9, 13, or 17 children.[14] The percentage of teachers in this experience range grew considerably for cohorts entering from the 1960s to the mid-1980s, then fell with the growth of teachers with 20 or more years of experience.

Nationwide, three significant changes took place in schools during this period—lower class size, and better educated and more experienced teachers.

Along with this gain in experience came more teachers with master's degrees. Figure 9 shows that the percentage of teachers with master's degrees experienced by an entering school cohort rose for all age groups. The education level of teachers grew steadily for entering school cohorts from the 1960s to the 1990s—although the period of fastest growth was from the 1960s to the mid-1980s. The size of classes was also reduced substantially during this period. Figure 10 shows the average pupil/teacher ratio for entering school cohorts up to age 9, 13, and 17. The pupil/teacher ratio—a measure of class size—also fell dramatically for cohorts entering school in the 1960s and 1970s, but slowed considerably in the 1980s. These changes occurred at both elementary and secondary levels although the timing was somewhat different. Part of the reason class sizes fell was also related to the baby boom. As enrollments dropped in the 1970s, rather than terminate teachers, the opportunity was used to reduce class size.

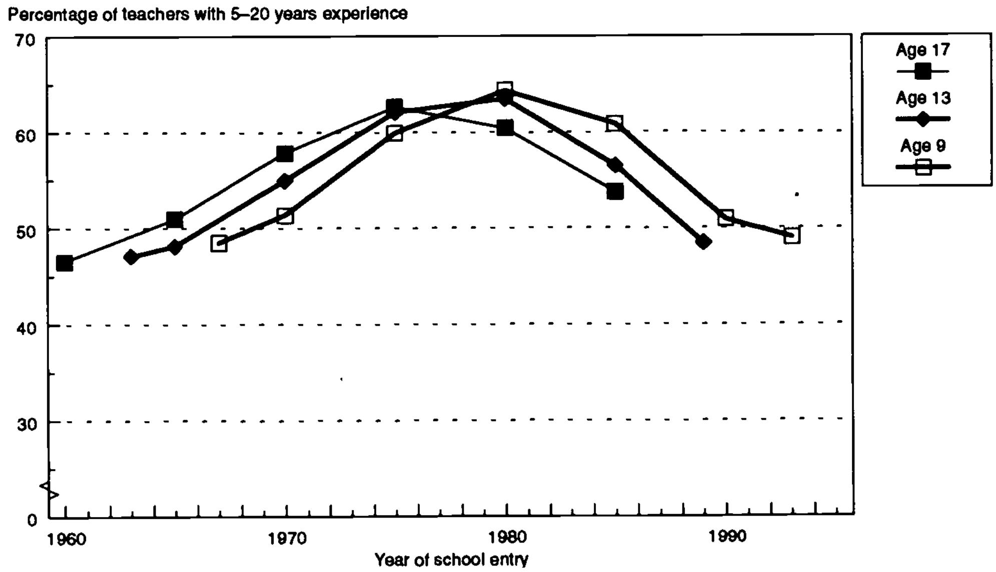
The empirical evidence on the effects of these three variables more greatly impacts class size than teacher experience or education for two reasons. First experimental evidence exists for the effects of class size. Second, the accuracy of all measurements using non-experimental data is now being questioned.

A large, multi-district study in Tennessee that randomly assigned students to classes of approximately 14 students instead of approximately 22, found that reducing class size between kindergarten and third grade had significant effects on achievement, and even greater effects for blacks (Krueger, 1997; Mosteller, 1994). The effects averaged about 0.20 standard deviations for whites and 0.30 standard deviations for blacks, with equal effects in reading and mathematics. Following the experiment, Tennessee also cut class sizes in 13 school districts with the lowest family income. Comparisons with other districts and test score changes within these districts showed gains of 0.35 to 0.5 standard deviations (see Mosteller, 1994).[15] The Tennessee data suggest that disadvantaged students may experience the most gains from class size reductions.

[14] We somewhat arbitrarily chose 5–20 years of experience as the period of peak productivity for teachers assuming an early learning curve for teachers and some average "burnout" effect after 20 years of service. The data here is the average percentage of teachers with 5–20 years of experience during the schooling experience of each age group from school entry. Thus 9-year-old students who entered school in 1970 would be estimated by taking the average percentage of teachers with 5–20 years of service between 1970 and 1973. For 13-year-old students entering in 1970, the average would be from 1970 to 1977.

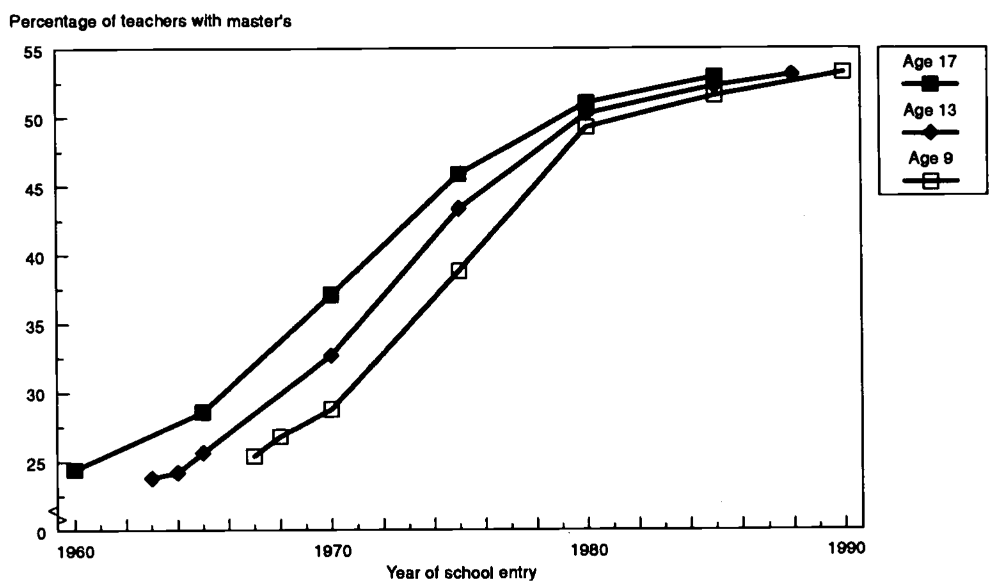
[15] Other research on school districts in Alabama shows similar overall effects when using models without prior year test score controls (Ferguson and Ladd, 1996).

Figure 8.—Average percentage of teachers with between 5–20 years of teaching experience in years of attendance



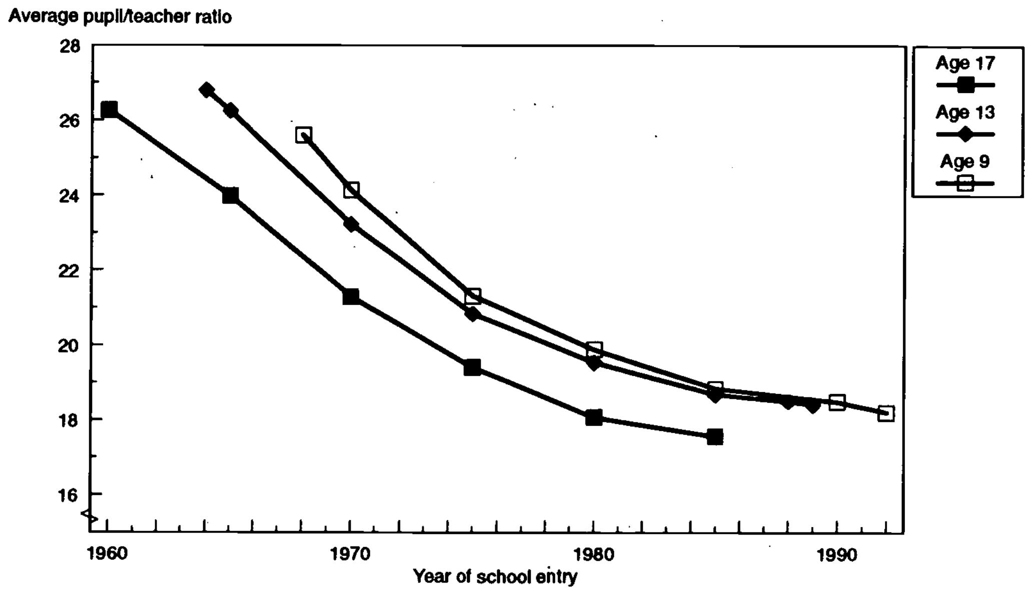
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).

Figure 9.—Average percentage of teachers with a master's degree or higher for years of attendance



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).

Figure 10.—Average cohort pupil/teacher ratio for years of attendance, by year of school entry



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).

It is likely that had a similar experiment been done using a national sample of K–3 students in the 1970s, even larger differences between black and white test scores would have been measured. This is because white students in Tennessee are poorer than average white students nationwide and black students in Tennessee probably were less disadvantaged in 1990 than blacks nationwide in the 1960s and 1970s. National reductions in class size in the 1960s and 1970s were approximately the same as the Tennessee experiment, thus effects ranging from 0.3 to 0.6 might be expected for black students and 0.00 to 0.20 for white students from the national class size reductions. Thus, reductions in class size may be a key factor in explaining large black gains and why black scores may have risen much more than white scores.

The Tennessee experimental evidence also leaves a number of questions unanswered. First, we do not know much about either the long-term effects of smaller elementary school classes or the cumulative effects of smaller classes from kindergarten through twelfth grade. In the Tennessee experiment, students were returned to large classes after third grade. By seventh grade the standardized benefits of smaller classes were only half as large as they had been at the end of third grade, and the benefits to black students were no larger than the benefits to whites (Mosteller, 1995). We do not know what would have happened if classes had remained small until students finished school. Second, the Tennessee measurements may only represent a short-term effect since only a single cohort was measured. Teachers and policymakers may be able to adapt their teaching and policies to take better advantage of smaller class sizes in the longer term.

Another problem with the hypothesis that class size reductions raised test scores is that class size fell in the 1960s as well as the 1970s. If smaller classes had conferred long-term benefits, 17-year-olds who entered school in 1968 should have outscored those who entered in 1960. This was not the case outside the South.[16] Further research is required to test the class size hypothesis as a strong contributing factor to black gains.

One important side effect of the Tennessee experiment is that it raised new doubts about non-experimental studies conducted in a "production function" framework.[17] These studies typically try to control standardized scores at Time One and then discern whether a resource like smaller classes affects gains between Time One and Time Two. In Tennessee, however, smaller classes exerted their entire effect on standardized scores in the first year. Thereafter, smaller classes simply served to sustain the initial standardized gains. Thus, the estimated effect on smaller classes in grades one through three would have been zero.

Current empirical measurements of the effects of teacher education and a master's degree show no consistently strong effects—but better specified models might change these results. It remains to be seen whether more teacher education and more experience raised achievement scores awaits stronger empirical evidence and determination of the current flaws in specifying estimation models.

Discussion

Recent research is undermining several of the assumptions and empirical evidence underlying the "money doesn't matter" conclusion. The validity of the empirical studies reviewed to arrive at this conclusion is being questioned due to the use of model specification which would not reproduce the experimental class size results. It is possible that fundamental flaws are present in nearly all non-experimental studies of the effects of school resources due to the methods of model specification. Second, experimental data which avoids the assumptions needed in models with non-experimental data indicates that reductions in class size—a key school resource parameter—have significant effects with larger effects for minority students. Third, NAEP data—which had previously been used together with the large perceived increases in school resources to support the "money doesn't matter" argument—now seems more supportive of a different conclusion. This evidence seems to support the thesis that money directed at minority and disadvantaged students brings higher achievement scores, but money directed toward more advantaged students may have much smaller or negligible effect. Moreover, the additional money available in the 1960s to 1990s was much less than previous estimates. Instead of doubling in real terms, the real increases directed toward achievement of regular students was closer to 30 percent during this period. These additional resources were also disproportionately directed toward minority and lower income students. Thus, a more consistent story is emerging from the empirical data which is more supportive of the thesis that additional money matters greatly for minority and disadvantaged students, but much less or little for advantaged students.

[16] Class size effects may depend on how teachers change their behavior when they have smaller classes (Murnane, 1996). The effects of smaller classes may take several years to appear, because teachers and students need time to adjust their behavior to smaller classes. However, effects appeared immediately in Tennessee and increased very little as a result of additional years in small classes. The Tennessee experiment did not address the important question of whether effects would grow at each grade level as teachers experienced smaller class sizes over many years.

[17] The Tennessee data shows that gains from smaller classes appear immediately and grew by small amounts over the first three grades (Krueger, 1997). This implies that production functions that utilize previous year's scores would not measure the class size effects evident in Tennessee. Two studies that have used some of the best data at the state level and had prior year's test scores as controls were considered to be among the strongest studies (Ferguson, 1991 and Ferguson and Ladd, 1996). However, these specifications now appear to provide biased results. In the latter study, results are also presented of cross-sectional estimates without controls, and these results may now have more credibility than those with prior year controls. Generally, the results of the model without prior score controls show stronger effects for most variables.

References

Campbell, Jay, Kristin Voelkl, and Patricia Donahue. *NAEP 1996 Trends in Academic Progress*, Washington, D.C.: National Center for Education Statistics/Educational Testing Service, NCES 97–985, September 1997.

Ceci, Stephen, and W. Williams. Schooling, Intelligence, and Income. *American Psychologist*, vol. 52: 1051–1058, 1997.

Cook, Michael, and William N. Evans. "Families or Schools? Explaining the Convergence in White and Black Academic Performance," draft copy.

Ferguson, Ronald F. "Paying for Public Education: New Evidence on How and Why Money Matters." *Harvard Journal on Legislation*, vol. 28: 465–97, 1991.

Ferguson, Ronald F., and Helen F. Ladd. "How and Why Money Matters: An Analysis of Alabama Schools." In Helen F. Ladd, ed. *Holding Schools Accountable*. Washington, D.C.: The Brookings Institution, 1996.

Flynn, J. R. "Massive IQ gains in 14 nations: What IQ tests really measure." *Psychological Bulletin*, vol. 101, no. 2: 171–191, 1987.

Gamoran, Adam. "Student Achievement in Public Magnet, Public Comprehensize, and Private City High Schools." *Educational Evaluation and Policy Analysis*, vol. 18, no. 1: 1–18, Spring 1996.

Grissmer, David, A. Flanagan, and S. Williamson. "Why Did Black Test Scores Rise Rapidly in the 1970s and 1980s." In *The Black-White Test Score Gap*, Christopher Jencks and Merideth Phillips, eds., Washington, D.C.: The Brookings Institution, Forthcoming.

Grissmer, David, and Sheila Nataraj Kirby. "Teacher Turnover and Teacher Quality," *Teacher's College Record*, vol. 99, no. 1: 45–56. Fall 1997.

Grissmer, David , Sheila Nataraj Kirby, Mark Berends, and Stephanie Williamson. "Exploring the Rapid Rise in Black Achievement Scores in the United States (1970–1990)." In *The Rising Curve: Long-Term Changes in IQ and Related Measures*, U. Neisser, ed. Washington, D.C.: American Psychological Association, 1998.

Grissmer, D. W., S. N. Kirby, M. Berends, and S. Williamson, 1994. Student Achievement and the Changing American Family, Santa Monica, CA: RAND.

Hanushek, Eric A. "School Resources and Student Performance." In *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, G. Burtless, ed. Washington, D.C.: The Brookings Institution, 1996.

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

Hanushek, E.A. "The Impact of Differential Expenditures on School Performance." *Educational Researcher*, vol. 18, no. 4: 45–51, 1989.

Hedges, Larry V., and Amy Nowell. "Group Differences in Mental Test Scores: Mean Differences, Variability, and Talent." In C.S. Jencks and M. Phillips, eds. *The Black-White Test Score Gap.* forthcoming 1998.

Hedges, Larry V., and Rob Greenwald. "Have Times Changed? The Relation between School Resources and Student Performance." In *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, G. Burtless, ed. Washington, D.C.: The Brookings Institution, 1996.

Hedges, Larry V., Richard D. Laine, and Rob Greenwald. "Does Monday Matter: Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes." *Educational Researcher*, vol. 23, no. 3: 5–14, 1992.

Karweit, Nancy. "Effective Kindergarten Programs and Practices for Students at Risk." In *Effective Programs for Students At Risk*, Slavin, Robert, Karweit, N. and Nancy Madden. Boston: Allyn and Bacon, 1989.

Krueger, Alan B., Experimental Estimates of Education Productions Functions, Princeton, NJ: Princeton University, Industrial Relations Section, Working Paper #379, May 1997.

Ladd, Helen F. *Holding Schools Accountable.* Washington, D.C.: The Brookings Institution, 1996a.

Ladd, Helen F. "Introduction." In *Holding Schools Accountable*, H.F. Ladd, ed. Washington, D.C.: The Brookings Institution, 1996b.

Lankford, Hamilton and James Wyckoff, "The Allocation of Resources to Special Education and Regular Instruction." In *Holding Schools Accountable*, Helen F. Ladd, ed. Washington, D.C.: The Brookings Institution, 1996.

Mosteller, F. "The Tennessee Study of Class Size in the Early School Grades." *The Future of Children*, vol. 5, no. 2: 113–127, 1995.

Murnane, Richard J., and Frank Levy. "Evidence from Fifteen Schools in Austin, Texas." In *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, G. Burtless, ed. Washington, D.C.: The Brookings Institution Press, 1996.

Neisser U., ed. *The Rising Curve: Long-Term Changes in IQ and Related Measures.* Washington, D.C.: American Psychological Association, 1998.

Rasinski, Kenneth, and Steven Ingels. "America's High School Sophomores: A Ten Year Comparison." NCES 93-087, Washington, D.C.: National Center for Education Statistics, June 1993.

Raudenbush, Stephen W., Randall P. Fotiu, and Yuk Fai Cheong. "Inequality of Access to Educational Resources: A National Report Card for Eighth Grade Math," forthcoming.

Rothstein, R., and K.H. Miles. *Where's the money gone? Changes in the level and composition of education spending.* Washington, D.C.: Economic Policy Institute, 1995.

Rethinking The Allocation Of Teaching Resources: Some Lessons From High Performing Schools

Karen Hawley Miles and Linda Darling-Hammond

About the Authors

Karen Hawley Miles is a nationally recognized and published student of strategic planning in public schools and of district and school resource allocation. Her focus is how to rethink the use of resources to improve instruction. She has spent the last 7 years working with school districts to rethink the use of resources and the organization of schools and she has worked to design school change and planning processes in several urban districts. She currently works with New American Schools Corporation to create and implement "break the mold" designs for public schools. She co-directed the recent study by the Economic Policy Institute, *Where Has the Money Gone?*, an analysis of the components of school spending over time. Prior to this, she worked at Bain and Company as a Strategy and Management consultant for hospitals and corporations. She has a B.A. in Economics from Yale University and a Doctorate in Education from Harvard University, specializing in school organization, change, and finance.

Linda Darling-Hammond is currently William F. Russel Professor in the Foundations of Education at Teachers College, Columbia University where she is also co-director of the National Center for Restructuring Education, Schools, and Teaching (NCREST) and Executive Director of the National Commission on Teaching and America's Future. She is actively engaged in research, teaching, and policy work on issues of school restructuring, teacher education reform, and the enhancement of educational equity.

This paper was originally published and supported by the Consortium for Policy Research in Education (CPRE). The work was funded by the National Institute on Educational Governance, Finance, Policymaking and Management, Office of Educational Research and Improvement, U.S. Department of Education, grant number R308A60003.

Rethinking The Allocation of Teaching Resources: Some Lessons From High Performing Schools

Karen Hawley Miles and Linda Darling-Hammond

Introduction

School reform proposals aimed at improving student achievement range from developing new standards and curriculum to personalizing student-teacher relationships and increasing time for teacher planning and learning. Research over the last 30 years has indicated that student achievement is higher, dropout rates are lower, and student affect and behavior are better in schools where students are well-known to their teachers (for reviews see Darling-Hammond, 1997; Braddock and McPartland, 1993; Lee, Bryk, and Smith, 1993). These findings have emerged from studies of school size, class size, and school organization. In terms of structural variables other than size, two critical conditions concern reduced tracking or curriculum differentiation and greater opportunity for students to work intensely with a smaller number of teachers over longer periods of time (Gottfredson and Daiger, 1979; Lee, Bryk, and Smith, 1993; Lee and Smith, 1994, 1995). Other studies have emerged that document problems with "pullout" models of instruction that serve students' needs by assigning them for brief periods of time to a variety of specialists, each of whom is supposed to treat one "problem" independent of the others (Commission on Chapter 1, 1992; Soo Hoo, 1990).

Meanwhile, reformers continually note that teachers need substantial time together in order to create new practices and engage in shared problem solving (see, for example, Darling-Hammond, 1997; Little, 1993; Sizer, 1992). Studies of teacher development in other countries, notably China and Japan, point out how teachers become more proficient from continually working on curriculum, demonstration lessons, and assessments together (National Commission on Teaching and America's Future, 1996; Stigler and Stevenson, 1991). The new curriculum reforms in the United States require substantial teacher learning which is, ideally, content-based and collaboratively pursued in tight connection to teachers' ongoing classroom work (Ball and Cohen, in press). In an era of belt-tightening and rising student enrollments, finding the resources for these reforms will require schools to reexamine the use of every dollar.

Much publicity has surrounded efforts to redirect dollars from administrative functions back to the classrooms. New Jersey's governor has even refused to reimburse district administrative costs above a fixed percentage that is lower than many districts currently spend. However, little attention has been given to rethinking the use of existing instructional resources, especially teachers—schools' most important and expensive resource.

On the surface it would seem that schools should have the needed resources to create more individual time for students and increase professional time for teachers. From 1960 to 1992, the number of pupils per teacher dropped from 26 to 17.6. Furthermore, schools employ 1 adult for every 9 students (NCES, 1994). Despite these generous ratios, class sizes exceed 25 for most students most of time, student loads are well over 100 for most teachers in secondary schools, and teacher planning time is both minimal and conducted in isolation from other teachers. The basic structure of schools has remained essentially the same across districts and over time, with new resources added largely around the regular classroom, rather than into it. And despite recent calls for "restructuring," a number of surveys suggest that public schools rarely engage in major reallocations of resources (Rettig and Canady, 1993).

How might existing resources be used to personalize student-teacher relationships and provide teachers time to plan and work more closely together? Can these goals be met in ways that also support student learning for all students, including those with special needs? With so few examples of public schools organized in different ways, there is little empirical research on these questions. Furthermore, researchers have not had common ways to describe and measure different models. This study aims to begin to fill these gaps by describing in detail how five schools have reallocated resources while supporting high levels of student learning. The sample is too small and the schools too unique to claim a causal connection between their organizational designs and their students' successes.[1] However, the schools do demonstrate that it is possible to support student achievement at extraordinarily high levels in contexts that manage instructional resources to maximize individual attention for students and learning time for teachers.

. . . we focus primarily on the assignment and use of teaching staff because it is the most sizable and the most underexplored area for potential resource reallocation.

We sought to identify the elements of organization and resource use that seem most important in these contexts and to quantify objectively the ways in which these schools use resources differently than traditional schools. We also aimed to learn from these examples about the conditions that facilitate or hinder resource restructuring. Although the schools look very different from one another, they share six principles of resource allocation implemented in different ways depending on their specific educational goals and strategies. In what follows, we describe their approaches, present a framework for examining resource allocation and use, and develop a methodology that may be used to measure the extent to which schools use their resources in focused ways to support teaching and learning.

Opportunities for Fundamental Reallocation of Resources

It is unlikely that schools can find ways to create more individual time for students or more shared planning time for teachers without prohibitively raising costs, unless they rethink the existing reorganization of resources. In this article, we focus primarily on the assignment and use of teaching staff because it is the most sizable and the most underexplored area for potential resource reallocation. As we noted above, pupil teacher ratios have dropped substantially and real expenditures have nearly doubled over the last 30 years. A recent analysis of staffing and spending patterns from 1967 to 1991 in nine very different districts across the country found that few of the new teaching staff were deployed to reduce class sizes for regular education students; most went to provide small classes to the growing number of students in special programs and to add a modest amount of time for teachers free from instruction during the school day (Miles, 1997a, 1997b; Rothstein and Miles, 1995).

As instructional staff have increased, the proportion of teachers has declined. Since 1950, the proportion of school staff who are classified as teach-

[1] Some large-scale quantitative studies do, however, suggest that student achievement is correlated with school designs that enable teachers to spend more extended time with smaller numbers of students and that allow teachers to make instructional decisions in teams (Gottfredson and Daiger, 1979; Lee, Bryk, and Smith, 1993; Lee and Smith, 1995; Darling-Hammond, 1997).

ers has dropped from 70 percent to 53 percent, not all of whom are classroom teachers. Overall, about 43 percent of staff are regularly engaged in classroom teaching (National Commission on Teaching and America's Future [NCTAF], 1996). By contrast, 60 percent to 80 percent of education staff in most European countries are classroom teachers, enabling much greater time for collaborative planning and professional development (OECD, 1995).

An analysis of the allocation of teaching resources in Boston public schools identified six reasons for the gap between potential and reality in U.S. schools (Miles, 1995). Many of these organizational practices are so widespread that Tyack (1994) describes them as the "grammar of schooling" while Sarason (1982) calls them "school regularities." These practices form the basis of our conceptual framework for understanding and quantifying the use of teaching resources in both traditional and nontraditional schools. Below we briefly describe the impact of each on the use of teaching resources.

An analysis of the allocation of teaching resources in Boston public schools identified six reasons for the gap between potential and reality in U.S. schools . . .

1. **Specialized programs conducted as add-ons.** In most school districts, a large portion of teachers work outside the regular classroom with special populations of students in categorical programs such as special education, Title 1 compensatory education, bilingual education, remedial education or gifted education. Such programs for special student populations absorbed 58 percent of the new dollars devoted to education from 1967 to 1991 (Rothstein and Miles, 1995). Most of these programs operate under federal, state, or district regulations and sometimes collective bargaining agreements that prescribe how teachers may be used and students may be grouped. Most districts operate these programs using a "pull-out" model in which students leave the regular classroom for portions of the day for instruction in small groups. In 1991, in Boston, teachers in specialized programs working outside the regular classroom represented over 40 percent of the teaching force.[2] Not only are pull-out strategies extremely costly, they also segregate students in sometimes stigmatizing ways and provide services that are often ineffective due to their fragmentation and lack of connection to the student's classroom experience. Schools rethinking resources will consider how remedial, special education, Title 1 and bilingual education resources might work together in an integrated plan to benefit all students in "regular education" settings.

2. **Isolated instruction-free time for teachers.** Currently, most school districts provide teachers with short periods of time free from instruction while using other classroom teachers to give instruction at these times. At the elementary level, teachers often have a 45 minute duty-free period four or five times a week which is covered by specialists in art, music, or physical education. In 1991, this represented 9 percent of Boston's elementary teaching resources. At the secondary level, a teacher might teach 5 of 7 instructional periods, using one for planning and the other for lunch or duties (monitoring hallways, the lunchroom, or study hall). Other teachers cover instruction during the portion of the student's instructional day in which the teacher is not teaching. Although secondary teachers have more preparation time than elementary teachers (about 4 or 5 hours per week as opposed to 3 hours), the short blocks of individual non-instructional time do not allow much substantive planning or collaboration. These activities would require longer blocks of uninterrupted time that is coordinated with other teachers. Schools rethinking their use of resources will consider ways of creating longer periods of time for teachers to plan and develop curriculum together.

[2] A similar analysis quantifies the impact of these practices in three other districts: Fall River, Massachusetts, Middletown, N.Y., and East Baton Rouge, Louisiana. See Miles 1997a.

3. **Formula-driven student assignment.** In search of efficiency and standardization, American schooling processes have been broken into segments, like grades and tracks, through which students are expected to move at an even, uniform rate. Districts use formulas to assign students to classrooms in a regularized fashion by age, subject, and program. These practices are costly, because the uneven allocation of teachers over grades, small programs and undersubscribed subjects contributes to unplanned differences in class size that do not reflect educational strategies. These practices may also preclude approaches such as multi-age or multi-ability grouping that can be more effective when teachers are prepared to engage in them (Anderson and Pavan, 1993; Slavin, 1990).

   Using formulas to allocate students to classrooms by age can create huge variation in class sizes. For example, Boston public schools cap elementary class sizes at 28. When the 29th student enters a school with only one class in that grade, a new teacher must be added. Thus, the class size average falls dramatically from 28 to 14.5. In 1991, regular elementary class sizes in Boston's 645 elementary classrooms varied from 15 to 31. Class sizes could vary by 8 or 9 students from one grade to another in the same school. The more separate programs and subjects a school has, and the more constrained by age grading or tracking practices, the more often this kind of unplanned variation in allocation of resources occurs. Schools looking to better match resources to student needs will consider assigning students to groups based on educational strategies rather than standard classifications. Schools may also strategically vary group sizes and the daily schedule for particular kinds of lessons or skill work.

4. **Fragmented high school schedules and curriculum.** The problems of age grading are compounded in high schools by tracking, schedules with large numbers of short periods (typically 45 to 50 minutes), and teacher and subject specialization. One unfortunate effect of this fragmented approach to schooling is its impact on teachers' pupil loads. With five classes of 25 to 30 students each, most secondary teachers work with 125 to 150 students per day. Reducing teaching loads without dramatically increasing costs can be accomplished by reducing the number of different groups teachers teach, either by combining traditionally separate subject areas, lengthening the duration of classes, or having smaller groups of students work intensively with teachers in smaller numbers of subjects, as in universities (Carroll, 1994).

5. **Large high schools.** Secondary schools average nearly twice the size of elementary schools (NCES, 1994), and frequently exceed 2,000 students. The conventional justification is that larger enrollments create economies of scale by distributing administrative and operating costs and offering a more diverse curriculum cost effectively. However, existing research suggests that high schools have created more internal specialization and departmentalization than can be scientifically justified. Most studies have found that, all else equal, above about 800 students, larger schools produce lower student achievement and lead to increased alienation, higher dropout rates, and larger numbers of administrative staff, thereby deflecting resources from classroom instruction. Furthermore, beyond about 400 students, there are few if any gains in curriculum quality (for reviews, see Lee, Bryk and Smith, 1993; Darling-Hammond, 1997).

Most studies have found that, all else equal, above about 800 students, larger schools produce lower student achievement and lead to increased alienation, higher dropout rates, and larger numbers of administrative staff . . .

6. **Inflexible teacher work day and job definition.** In Boston, like many other cities, the teacher contract specifies the required hours of work, starting and ending times, and how teachers can be assigned over the day. This makes it difficult to stagger starting times in order to make the best use of staff time or to meet student needs more fully. The contract prevented one high school from changing the work hours of its guidance staff to start and end later in the day so that students' meetings with counselors would not conflict with

their courses. The contract also requires that planning time be spread out over the day; this makes it difficult to combine instruction-free periods to create longer, shared blocks of time.

The use of part-time teachers is forbidden if they substitute for potential full-time positions. While intended to discourage management from substituting lower cost and potentially lower quality teachers for dedicated full-time ones, this regulation limits affordable solutions to coverage for teachers' noninstructional time. One way to create common planning time for groups of teachers, for example, is to schedule coverage for them by hiring a larger number of part-time specialist teachers during this time. Finally, rigid definitions of the work day exclude individuals who would be willing to work at times beyond the typical school hours. Schools looking to better match resources to student and staff needs may want to consider the use of highly-skilled staff in part-time positions and on varied job schedules.

Our observations in schools suggest, however, that altering any one practice alone may not free enough resources to significantly change the possibilities for student or teacher learning.

The strategies observed in schools that have sought to change these conditions can be described by six principles of resource reallocation:

1. Reduction of specialized programs to provide more individual time for all in heterogeneous groups,
2. More flexible student grouping by school professionals,
3. Structures that create more personalized environments,
4. Longer and varied blocks of instructional time,
5. More common planning time for staff, and
6. Creative definition of staff roles and work schedules.

These kinds of changes can create opportunities for realigning teaching resources to school goals (Miles, 1995). Our observations in schools suggest, however, that altering any one practice alone may not free enough resources to significantly change the possibilities for student or teacher learning.

Study Methods and Analytic Framework

Sample

To create a sample of schools that could offer insight into the possibilities and challenges involved in rethinking the allocation of instructional resources, the study sought elementary and secondary schools that:

1. Had engaged in a significant rethinking of resources touching on at least four of the resource principles listed above.
2. Used no significant extra resources above the school system average per pupil except startup or training grants.
3. Served a diverse student population in terms of income, ability, language background, and special needs.
4. Had used a new model of organization for at least 2 years.
5. Had evidence of strong and improving student performance.

To find such schools, we surveyed experts involved in reform networks nationwide. The five schools we selected represent different educational strategies. Three of the schools started from scratch with a new design in mind, and had considerable flexibility in hiring staff and creating programs. The other two schools restructured existing programs and staff. We selected three elementary schools and two secondary schools:

Quebec Heights Elementary School, Cincinnati, Ohio had, at the time of the study, 500 students in grades K–6, with 15 percent classified as having special education needs and 70 percent eligible for Title 1. Quebec Heights eliminated age- and program-based instructional grouping and put students in smaller, multi-aged, heterogeneous groups that remain together for 3 years. The school created reading groups of 8 or smaller each day. Teachers have common planning time each day and pursue professional development in the school's priority areas during the school day. Cohort analysis of student performance data shows that both special education and regular education students have improved faster than the Cincinnati average.

The Douglass Elementary School, Memphis, Tennessee had 475 students with 17 percent classified as requiring special education and 88 percent qualifying for Title 1 support. At the time of the study, the school was in its third year of implementing the "Success for All" program which restructured school resources to allow 90 minutes a day of reading plus daily individual tutoring for first and second graders who did not meet grade level standards. In addition, Douglass was working to integrate its special education students and teachers fully into the regular classroom. After the second year of implementing the program, the percent of second-graders (the only students with two years of the new model) above the median in language arts climbed from 17 to 59 percent. In addition, the school's evaluation of special education integration showed these students continuing to progress academically and socially.

The Mary C. Lyons Model Elementary School, Boston Massachusetts had 90 students in grades K–5, 60 of whom were classified as regular education and 30 of whom had severe emotional disturbances that previously required placement in highly restrictive settings. Over 80 percent of students qualified for Title 1. The school fully integrated all special education students into regular classes of 15 or smaller, each with a teacher and instructional assistant. Mary C. Lyons created extended school hours lasting from 7:00 am to 5:15 p.m. using outside contractors to provide instruction. The variety of different staffing arrangements included paraprofessionals, teacher interns, part-time workers, and staggered shifts. The school was 1 of 15 (out of 115) Boston schools to be over-chosen by every race for both special education and regular education slots for 3 years in a row. Achievement scores for both special education and regular education students improved faster than the Boston average, and 100 percent of the students were reading on grade level.

Central Park East Secondary School (CPESS), New York, New York served 450 students in grades 7 through 12, about 25 percent of whom qualified for special education and 60 percent for free or reduced-price lunch. All students were integrated into heterogeneous classrooms. The school restructured the typical secondary schedule to create two hour blocks of instructional time in both the humanities and math/science. Teachers had more than 7 hours each week of common planning time in addition to their daily individual preparation period. To reduce academic group sizes, CPESS allocates nearly all of its positions for teaching, rather than hiring guidance counselors and various administrative staff. All professional staff members serve as advisors to about 12 to 15 students each year. The school hires some teachers on a part-time consulting basis for electives like foreign languages. CPESS has been nationally heralded for its consistently exceptional outcomes: each year since its first graduating class, more than 90 percent of its students have graduated and more than 90 percent have been accepted to college.

International High, New York, New York is an alternative school of 475 recent immigrant students in grades 9 through 12. Only students who have been in the United States less than 4 years and who score below the 20th percentile on an English language proficiency exam are admitted. At the time of the study, over 75 percent of the students were eligible for free or reduced-price lunch. International High integrates all state-mandated subject matter in an interdisciplinary curriculum taught in multi-aged heterogeneous groups. Teachers work with no more than 75 students a term and spend 70 minutes or more with them each day. The teachers have nearly six hours each week of common planning and professional development time. All staff members lead a small advisory group which meets weekly to discuss issues of personal, academic, and social growth. Despite its "high risk" population, the school's dropout rate was less than 1 percent in 1993–94 as compared with 30 percent citywide, and both the graduation rate and college acceptance rates exceeded 95 percent. For more than a decade, these rates have exceeded 90 percent annually (IHS, 1995; Darling-Hammond, Ancess, and Falk, 1995).

Figure 1 summarizes the resource allocation strategies used in the five sample schools. It shows that each school implemented many strategies for allocating teachers and teaching time to address student needs and create planning time. Only the three alternative schools—Mary C. Lyons, CPESS, and International High—created new teaching roles by contracting with other providers for teaching or by restructuring some teaching positions.

Data Collection and Analytic Framework

To understand the resource allocation practices in each of the five schools, we collected information about school expenditures, staffing, and student scheduling. We collected comparable data for traditional schools, along with district level budget and staffing information. We conducted interviews with administrators and teachers and examined available written material at each school to understand the school's organization and its link to educational purposes and outcomes.

. . . this study created measures which allow comparison of resource allocation patterns between these school models and traditional schools.

In addition to describing the strategies each school used, this study created measures which allow comparison of resource allocation patterns between these school models and traditional schools. The measures were developed by taking each resource allocation principle, hypothesizing the quantifiable impact it might have on resources, and then testing this impact by using indices that are: (1) descriptive of practices in both traditional and nontraditional schools, (2) easy to understand, and (3) replicable.

Creating measures that accurately portray practices in the fluidly organized sample schools yet allow comparison to traditional schools creates a tension between the use of measures that are easily understood and calculated and those that can provide meaningful description. The subtleties involved can be seen through one example: the attempt to measure the impact of the principle "reduction of specialized programs to create more individual time for all." In a traditional school, regular class size provides a useful gauge of how much access to individual attention a student might have. But, regular class size does not reflect the student's experience in some innovative schools because it does not describe the way these schools organize by subject and over the course of the day. For example, the regular class sizes of 24 at Quebec Heights school do not reflect the fact that all students spend 90 minutes a day in groups of 8 for reading. In order to capture the additional individual time for all students, a measure of average instructional group size, rather than regular class size, is used. This measure relies on greater descriptive knowledge of a school, but it more accurately reflects student experience.

Figure 2 summarizes the measures used for each resource allocation principle. The first principle, "reduction of specialized programs to create more individual time for all in heterogeneous instructional groups" should lead to smaller average instructional groups for all regular education students and more even distribution of resources between regular and special program students. Three measures helped assess the extent to which innovative schools differed from traditional schools here.

1. **Students per teacher:** This number includes all teachers and students in the school from all programs. Our sample elementary schools had similar numbers of students per teacher. However, a school can reduce its functional student to teacher ratio by converting typical non-teaching slots to teaching roles as CPESS did. The index of students per teacher indicates only the opportunity to create small, flexible instructional groups. It does not reflect the actual size of the groups in which most students spend time.

Figure 1.—Resource reallocation strategies used by sample sites

Strategy	Mary C. Lyons Model Elementary	Quebec Heights Elementary	Douglass Elementary	International High	Central Park East Secondary
1. Reduction of specialized programs	X	X	X	X	X
2. More flexible student grouping	X	X	X	X	X
3. Structures to create more personal environments	X	X	X	X	X
4. Longer and varied blocks of instructional time	X	X	X	X	X
5. More common planning time	X	X	X	X	X
6. Creative definitions of staffing roles and work day	X			X	X

X=Sample school implements strategy.

SOURCE: Unpublished tabulations.

2. **Weight average group size:** This measure calculates the weight average size of the instructional group which a typical student experiences over the day for academic subjects. It incorporates the time spent in different group sizes over the day. For example, if students in a classroom of 24 spent 90 minutes a day (25 percent of their school day not including lunch) in reading groups of 8, then the weight average group size would be 20 (.75 times 24 plus .25 times 8). In a traditional school, the average group size and the regular class size would be the same. This measure may offer a clearer sense of how much access to individual attention most students in the school have.

3. **Percent of teachers in regular education instructional groups:** This figure divides the number of teachers who work with regular education students (including classroom teachers, subject specialists and other teachers who work all day instructing groups that include regular education students) by the total number of teachers in the school. The figure gives a sense of the extent to which a school has concentrated its resources on core classroom functions as opposed to pull-out programs of various kinds.

The second principle, more flexible student grouping, should allow educators to create instructional groupings that more closely match instructional needs. Formulas that mandate the size of groups and classrooms can create situations where group sizes vary for no educational reason. When teachers can create their own groups using criteria linked to educational strategies, they can reduce these unplanned variations and create a strategy which maximizes the use of limited resources. The *percent of regular education students in targeted group sizes* represents the extent to which a school has minimized random variation in class size. In schools where no group size target existed other than the contractually defined class size maximums, we measured how many students were in classes which were within 5 percent of the average size. More flexible student grouping also allows teachers to create smaller groups for target subject areas. The *average size of instructional groups in focus area* measures how schools focus resources to create more individualized attention in some subjects where they do so. If some regular education students spend time in much smaller instructional groups, this would be reflected in the average by calculating the percent of students receiving such support.

Four aspects of the third principle, structures to support more personal relationships between teachers and students, lend themselves to measurement. First, a primary indicator of a teacher's op-

Figure 2.—Measuring resource allocation patterns (staff allocation)

Resource allocation principles	Expected impact on resources	School measure
Reduction of specialized programs to provide more individual time for all in heterogeneous groups	Smaller sized regular education instructional groups	Students per teacher
	More even distribution of resources between regular and special program students	Average size of regular education instructional groups
		Percent of teachers in regular instructional groups
More flexible student grouping by school professionals	Smaller instructional groups in focus areas	Percent of students in target regular education size groups
	Less unplanned variation in class sizes	Average size of group in focus area
Structures to create more personalized environments	Lower teacher/student loads	Teacher/student loads per day
	More adults involved in instruction	Percent of adult instructors/ advisors
	Smaller teams of teachers and students	Size of teacher/student clusters
	Multi-year relationships between students and teachers	Length of student/teacher relationship
Longer and more varied blocks of instructional time	Longer instructional periods for academic subjects	Average length of instructional period for academic subjects
More common planning time for staff	More minutes of common planning	Common planning minutes/ week
	Longer periods of time for planning	Length of longest planning period
Creative definition of staff roles and work schedules	Use of part-time or contract staff Use of interns or paraprofessionals for instruction Staggered work schedules	Not applicable

SOURCE: Unpublished tabulations.

portunity to build relationships with each student is the academic teacher's *student load.* A second indicator of a school's effort to maximize personal relationships might be the *percent of professionals who serve as instructors or advisors* to regularly scheduled groups of students in an ongoing fashion. Thus, an assistant principal who worked with occasional discipline problems or a guidance counselor meeting once with each of 200 students to ensure compliance with graduation requirements would not be included. Although these singular contacts with students can be important, they do not aim to build long term, personal relationships between school professionals and students. The *average size of teacher and student teams or clusters* provides a third measure of the opportunity to create a more personal educational environment. For this measure, student-teacher teams had to be self-managing and self-contained. This means that virtually all instruction occurs within the cluster and that the cluster has primary responsibility for curriculum, grouping, discipline, and evaluation of its students. A final strategy schools might use to create personal relationships would be to keep teachers and students together for longer than the typical year. Thus, we include a measure of the *number of years teachers and students stay together.*

The extent to which sample schools created *longer blocks of instructional time* is measured by the average scheduled *length of instructional period* for academic subjects in secondary schools. In some of the schools studied, teachers vary the length of instruction from the schedule to suit a particular lesson. These variations were not calculated here.

Finally, two measures are used to understand how sample schools create more useful common planning time for teachers. The *number of minutes of common planning time* is defined as the amount of time each week that is shared with other teachers and spent on collaborative planning regarding curriculum, students, or school practices. A second indicator of the usefulness of the planning time is the *length of the longest planning period.* For some kinds of planning and development, teachers need time periods longer than the typical 40 to 50 minutes.

Adjusted for student mix, the schools in this sample used no greater resources than traditional schools on an ongoing basis.

Each innovative school is compared with a typical school in the same district with a similar student population. Meaningful comparisons must include an adjustment for the mix of students eligible for special services because schools typically receive additional resources to serve them. Adjusted for student mix, the schools in this sample used no greater resources than traditional schools on an ongoing basis. In two cases, no "traditional" school existed in the district which served the same mix of students as our sample sites. Mary C. Lyons Model Elementary School in Boston draws a large percentage of its population from special education students typically served by private schools. In this case, a hypothetical comparison was created, based on the assumption that these students were served in separate, self-contained classrooms of 4 each, the smallest existing class size, and that social services and other support staff were the same as at Mary C. Lyons.

International High in New York City serves a unique population of limited English speaking students. Traditional schools serve such students through bilingual programs and ESL courses offered separately from the rest of the high school curriculum, but do not typically have 100 percent of their population requiring such services. To create a comparison to International High, we used the New York City staffing allocation formula to determine the number of teachers the school would have received for these special needs students; we assumed the additional resources would be used to fund separate bilingual or ESL classes. Although this generous assumption about universal ESL services to limited English proficient students does not hold true in many of New York's traditional schools, it offers a best case scenario for the allocation of resources in a traditional model.

These calculations are intended to provoke discussion and to provide an objective way of comparing schools to each other. Obviously, other factors which are not incorporated in these measures contribute to the opportunity for individual attention and shared teacher time. For example, a teacher in a class of 24 may use sophisticated grouping practices that allow her or him to provide targeted individual or small group instruction to students throughout the day. These grouping strategies are not incorporated into this measurement scheme unless the entire school uses the strategy. The existence of planning and development time does not guarantee that it is used to improve teaching quality. Further, many schools find common planning time for teachers outside the school day on a volunteer basis. Thus, these measures are intended to be used in conjunction with a descriptive understanding of the way a school has organized itself to match resources to student needs and to provide opportunity for teacher growth.

Study Findings

We discuss our findings concerning elementary and secondary schools separately because they begin with such different organizational structures. With their relatively small teaching loads and self-contained multi-subject classrooms, elementary schools already allow more flexible, individual instruction. But their simple structures, with limited teacher time free from instruction, do not offer the same opportunities for freeing time and resources as secondary schools. Because of these simpler daily schedules, reducing the use of pull-out programs for special education, language and Title 1 instruction becomes a primary lever for creating smaller groups for all in elementary schools. In contrast, traditional secondary schools, with their short periods, multiple classes, large teaching loads, and greater amounts of non-teaching time offer more ways to reconfigure their resources.

Elementary Schools

Figure 3 presents the resource allocation measures for the three elementary schools. In the three urban districts studied, the traditional schools served regular education students in age-graded, self-contained classrooms. About 75 percent of the teachers worked with regular education students; the other 25 percent worked with Title 1 and special education students outside the regular classroom (figure 3, line e). Because all of these schools are in urban areas with high concentrations of students in poverty, even the traditional schools were using at least some of their Title 1 resources to add regular classroom teachers. Thus, their regular education class sizes averaged between 19 and 22 (line b). Class composition and class size stayed the same all day, for all subjects, except when students were pulled out for special education or Title 1 instruction. The elementary classroom teacher instructed all subjects except specialties like art, music, and gym which were taught by specialists during the classroom teacher's free period. Teachers had 45 minutes 3 to 5 times a week free from instruction in addition to short lunch periods. These times were not coordinated with other teachers in any systematic way.

Because all of these schools are in urban areas with high concentrations of students in poverty, even the traditional schools were using at least some of their Title 1 resources to add regular classroom teachers.

Reduction of Specialized Programs

In departing from this organization, all of the sample schools increased to above 90 percent the percentage of teachers who worked with heterogeneous groups of students (figure 3). The comparable percentages in the traditional schools ranged from 28 to 77 percent. The only teachers in the restructured schools not working with heterogeneous groups of students were teachers of special education students in separate classrooms at Quebec Heights.

Each elementary school used different levers for matching instructional resources to student needs, depending on its educational goals. First, Quebec Heights decided to use multi-age grouping to respond more effectively to diversity in student skill levels. Students were assigned to multi-age clusters, called "families," containing three to four teachers and 75 to 85 students in grades 1–3 or 4–6; the families remain together for 3 years. Students may work with any instructor within the family during the day but each has a homeroom teacher who has primary responsibility for a class of 22 students for the full year. Rather than divide the curriculum by age level, all students in the family study the same basic topics during the year, but at their own developmental level.

Second, Quebec Heights eliminated separate Title 1 programs and used these resources to reduce the size of reading groups for all students. Third, special education students and teachers were fully integrated into the families. In the primary grades, the special education resource teacher works as 1 of 4 teachers in a team responsible for a group of 85 students.

Figure 3.—High performing versus traditional elementary schools

Resource allocation principles	School measure	Quebec Heights Elementary		Douglass Elementary		Mary C. Lyons Model Elementary	
		Average	Traditional	Average	Traditional	Average	Traditional
Reduction of specialized programs to provide more individual time for all in heterogeneous groups	a. Students per teacher	15	15	16	16	11	7
	b. Average size of regular education instructional group	19	21	26	22	13	19
	c. Percent of teachers in regular education instructional groups	91%	77%	95%	76%	100%	28%
More flexible student grouping by school professionals	d. Percent students in target size instructional groupings	100%	65%	100%	60%	—	—
	e. Average size of instructional group in reading	7	21	20	22	—	—
Structures to create more personalized environments	f. Student loads for primary classroom teachers	22	21	24	19	13	19
	g. Length of time students stay with teacher (in years)	3	1	1	1	1	1
More common planning time for staff	h. Common planning minutes/week	325	100	135	0	405	45
	i. Length of longest planning period	45	45	45	45	105	45

— Not applicable.

SOURCE: Unpublished tabulations.

The Douglass Elementary School in Memphis used its Title 1 budget as the primary lever for rethinking resources to improve student performance. Because 97 percent of its students qualify for Title 1 assistance, Douglass has long been free to use Title 1 dollars across the school. At approximately $250,000 dollars per year, these resources represent nearly 20 percent of the school budget. Douglass restructured resources using an existing model for improving student performance, the "Success for All" program. Following this model, Douglass uses Title 1 dollars to hire reading teachers to work as one-on-one tutors with students who do not meet reading standards in the first and second grades. These teachers and all special education teachers combine with regular classroom teachers to reduce the size of instructional groups from 24 to about 17 for 90 minutes of reading a day for all students.

The Douglass example illustrates why simple measures of class size do not provide enough information about how resources follow instructional goals. Prior to implementing "Success for All," Douglass used Title 1 dollars for regular classroom teachers and had classes averaging 17 across the school. As Principal Myra Whitney commented, "We had slowly reduced all class sizes over the years with no plan for how anything in the classroom would change. It wasn't working, our students were still at the bottom in reading." To implement Success for All, Douglass raised class sizes for all other subjects in order to reduce group sizes for reading and provide targeted one-on-one tutoring assistance to ensure all students are reading by third grade. Douglass also redirected resources from grades 3–6 to the early grades. The decision to take resources away from some students and teachers to focus on others can produce tension. Douglass's use of a proven model that included clear staffing requirements minimized this friction. As one teacher put it, "Everything is specified by 'Success for All;' we didn't consider quarreling with it because research shows this works."

Douglass used "Success for All" as a catalyst for including its special education teachers and students in regular classrooms. By the third year of the program, all students and teachers from previously self-contained classrooms and resource rooms spent most of their time in heterogeneous groups. During the daily 90 minutes of reading time, special needs students worked in heterogeneous groups based on their reading skill levels. The integration of special education students was made easier by the fact that cooperative learning plays a large role in "Success for All" classrooms. Assigning special education teachers their own reading groups that included students from all programs further reduced the size of reading groups for all students. During most of the rest of the day, special education teachers team taught with regular education teachers. They used approximately one-quarter of their time for performing individual assessments and working with students needing more targeted help outside the regular classroom.

(E)ach elementary school used its resources from special programs to support its core design (and) . . . staff organization depended on the educational strategies and approach the school had adopted.

While Quebec Heights redesigned traditional age-grading practices and Douglass rethought its use of Title 1 resources, the Mary C. Lyons school used the reallocation of special education dollars as a redesign lever. By including special education students, who were previously educated in a private setting at a cost of over $30,000 each per year, with regular education students, Mary C. Lyons created a unique, individualized environment for students and teachers. Mary C. Lyons is open to all students from 7:15 a.m. to 5:00 p.m. daily. Each classroom had 15 or fewer students—10 "regular" education students and 5 students with severe emotional/behavioral issues—and was staffed by a teacher, a teacher intern, and an after-school teacher. The pairs of six classroom teachers and six teaching interns included three teachers with regular education certification and three with special education certification. This unusual integration of special education students and teachers is driven not by finances but by a belief that schools must meet children's academic and emotional needs at their level of development. The teaching staff is hired to ensure attitudes, skills, and expertise to meet a broad range of academic, social and behavioral considerations. They work closely as a team to analyze the effectiveness of their instructional efforts. Academic teachers have close to two hours daily of common planning time.

Virtually all teaching resources at Mary C. Lyons supported this design, including Title 1 funds and funds which would have paid for subject specialists in traditional schools. A typical Boston elementary school has four subject specialists (in art, music, physical education and computers) who supplement instruction and cover planning time for classroom teachers. Instead of this costly arrangement, Mary C. Lyons pooled these dollars to contract with outside teachers for the provision of art and music and part of the after school program.

In summary, each elementary school used its resources from special programs to support its core design. Quebec Heights and Douglass raised regular education class sizes and redirected funds to reduce reading group sizes. Mary C. Lyons used funds freed from eliminating separate programs to lower teacher student ratios dramatically all day. In each case, staff organization depended on the educational strategies and approach the school had adopted. The organization of resources and the educational goals in these schools were inextricably intertwined because the organization enabled the schools to implement new teaching strategies.

More Flexible Student Grouping

Perhaps the most striking difference between the sample elementary schools and traditional schools is the proactive, strategic way in which teachers adapted instructional grouping to student needs. In a traditional school, administrators assign students to programs and classrooms according to bureaucratic rules and categories that stay constant over time and subject. Teachers in sample schools used their knowledge of student needs, rather than a student's program classification or age, to assign each to a regular homeroom class and to manage their instruction throughout the day.

In sample schools, reading groups were significantly smaller than in traditional schools. Mary C. Lyons and Quebec Heights organized staff to allow groups of six and seven, respectively. Mary C. Lyons used the classroom teacher and teaching intern to create reading groups of six. Quebec Heights rotated Title 1 teachers and instructional assistants through regular classroom, so that each classroom had three instructors for 90 minutes of reading time per day. The primary classroom teacher determined the composition of the groups and content of lessons daily based on consultation with the expert reading teachers and review of students' progress in specific areas. Some lessons divided students into groups based on areas where they need further skill development, others grouped students heterogeneously to discuss reading content. This concentration of resources on the reading rotation meant that homeroom class sizes were one student larger on average than the traditional model.

Perhaps the most striking difference between the sample elementary schools and traditional schools is the proactive, strategic way in which teachers adapted instructional grouping to student needs.

At Douglass, all students spent 90 minutes per day in reading groups of 15–17, as compared with 22 in traditional schools. The composition of these reading groups varied each day and over the course of the year depending on the teachers' assessment of student needs. A team including the teachers, reading specialists, and the "Success for All" facilitator assigned students to skill-based reading groups across grades using formal assessments every 6 weeks. Since assignment to groups indicated skill level, as opposed to age or a static assignment of aptitude, the student moved on once he/she demonstrated these skills. Students not mastering skills by agreed upon times received one-on-one tutoring for 20 minutes each day from one of the three reading specialists. About 15 percent of first and second grade students received tutoring at any one time, but the students receiving tutoring varied over the year depending on who needed extra assistance in particular skill areas.

This continuous assessment and regrouping of students required significant time and joint effort. A full time "Instructional Facilitator" helped teachers to conduct, analyze, and act on the assessments. The facilitator received in-depth training for using "Success For All" reading assessment tools and worked with a district wide expert in "Success For All" who had further expertise. In pulling this facilitator from the classroom, Douglass once again traded general regular education class sizes for a strategic use of resources which supported their school design. In this case, the facilitator enabled a more careful matching of instruction to student needs as well as more effective use of joint planning time.

Traditional schools experience variations in class sizes driven by formulas and enrollment swings. Boston's school choice plan allowed Mary C. Lyons to cap the number of students by grade through the student assignment process. Mixed age grouping at Douglass and Quebec Heights allowed teachers to control group sizes. For example, if Douglass had used age-based grading, class sizes in the first and second grade would have been 24 and 26, respectively, with class sizes declining as the student moved toward sixth grade. Instead, the Douglass staff combined grades to create smaller groups of 23 in the first three grades and 26 in the intermediate grades. Thus, sample schools exerted more control over class sizes by combining ages and programs so that all students were in targeted class sizes rather than the 60 to 65 percent who would have been in such group sizes using traditional age grading.

Structures to support more personal relationships

Quebec Heights' family structure aimed to strengthen relationships between teachers and students by keeping teachers with the same family of 85 students for 3 years, usually with the same homeroom class. This meant that some teachers received as few as nine new students each year. As an intermediate teacher stated, "It's hard to overestimate how much time this saves us. We get started quickly in the new school year, students know the rules and boundaries, and I know what they can do."

The Mary C. Lyons school's small size and intense staffing ratios created a highly personal environment for all students. The staff also created time to discuss each student's progress as a team. All the professionals working with each group of students—the classroom teacher, intern, a special education evaluation specialist, the afterschool director, and a social worker—met weekly for 45 minutes. Together, they identified problems, discussed possible strategies, and shared successes and frustrations.

More Common Planning Time

All three schools created more common planning time, but the constraints of time and collective bargaining agreements meant that only Mary C. Lyons made dramatic changes (figure 3). Both Douglass and Quebec Heights increased common planning time using the conventional method of scheduling specialists to allow meeting time for small groups of teachers. Mary C. Lyons' academic teachers shared a 30 minute lunch period followed by 1 hour and 15 minutes of common planning time while students had lunch and recess and received instruction from their instructional intern and after-school teacher. In combination with the "student support" team meetings described above, teachers met together for a total of 405 minutes weekly, in contrast to no more than 60 minutes in a traditional school.

...Lyons school's small size and intense staffing ratios created a highly personal environment for all students. The staff also created time to discuss each student's progress as a team.

Creative Definition of Staff Roles and Work Day

The Mary C. Lyons school was able to create so much more planning time because it redefined teaching roles throughout the day. Whereas in a traditional school only the classroom teacher or subject specialists assume responsibility for classroom instruction, Mary C. Lyons had a master teacher and a highly trained and supervised "instructional assistant trainee" in each classroom. In contrast to often poorly trained paraprofessionals, the Mary C. Lyons "trainees" were college educated students working on their master's degrees in special education at Wheelock University. To do this, Mary C. Lyons negotiated with the Boston Teachers Union to convert paraprofessional slots to "instructional assistant trainee" positions. As part of their program, the Wheelock students work in schools for stipends of $10,000 per year and participate in intensive coursework over holidays and summer. Wheelock sends a faculty member every two weeks to observe and discuss the trainee's practice with the master teacher. The trainee's stipend costs less than the $18,000 in salary and benefits for a paraprofessional. The savings allowed the school to give each teacher an "instructional assistant trainee." Where possible, the new instructional assistants were recruited from the existing paraprofessional staff. While the trainee position represented a short term cut in pay, this position led to full-fledged certification as a special education teacher for these staff.

In addition, Mary C. Lyons contracted with teachers to cover "schoolwide" planning time. The "afterschool" teachers overlapped the regular school day by one hour. During this time, they managed the classroom along with the instructional assistant trainee. In addition to providing regular teachers with planning time, this overlap provided a chance for after school teachers to transition from the regular academic day along with someone who had been with the students all day, thus allowing more continuity and better care for the children. The eight afterschool teachers, who were provided through a contract with a private nonprofit organization, specialized in behavior management and brought a wide range of experience with emotionally disturbed as well as gifted students. The principal worked closely with the contractor to specify the qualifications of these teachers, and the contract was contingent on the hiring of such exceptional teachers.

Secondary Schools

The traditional high school, with its departmentalized instruction and highly-segmented school day, offers many more opportunities for rethinking resource allocations than do elementary schools. We examined a typical comprehensive high school in New York City serving about 3,300 students.[3] The school is considered a "good" school serving a largely middle- and working-class population of students. It had nearly as many special needs and Title 1 students as CPESS and was in the process of beginning to restructure its programs. At the time of the study, however, it used traditional staffing and scheduling practices.

As figure 4 shows, the sample high schools looked very different from the traditional high school on virtually every dimension measured. Although our analysis is focused on the use of instructional staff, it is worth noting that the traditional high school had many more non-teaching staff than the two restructured schools. Not including custodial and food service workers, more than 40 percent of its total staff had non-teaching assignments. These included 1 principal, 9 assistant principals, 13 secretaries, 10 school-based services specialists (social workers, psychologists, etc.), 3 librarians, 17 security guards, 22 non-teaching school aides, and 14 classroom-based paraprofessionals. In the restructured schools, just over 25 percent of staff had non-teaching assignments and most of them taught at least part-time (Darling-Hammond, 1997).

The traditional high school, with its departmentalized instruction and highly-segmented school day, offers many more opportunities for rethinking resource allocations than do elementary schools.

The traditional high school had one instructional staff person for every 14.7 students—and New York City staffing allocations would reduce the student load to 13 for a population of students like that at International High.[4] Because fewer than two-thirds of these instructional staff members taught full-time, however, class sizes averaged about 33. Special education, bilingual education, English as a Second Language, and Title 1 programs were administered separately, with generally smaller class sizes. The typical student attended school from 8:05 a.m. to 2:13 p.m. participating in seven different 42-minute classes with seven different teachers, plus one lunch period. Teachers taught five instructional periods a day, with two periods free from instruction: one used for planning and the other for rotating "building assignments" such as cafeteria duty or hall duty or other administrative or program responsibilities. Excluding these special duties, teachers routinely saw about 167 students per day.

By contrast, the two sample high schools began with resources roughly similar to the traditional school and ended with dramatically smaller group sizes and teacher loads. As noted, the first difference was allocating a greater share of resources to instructional staff rather than administrative and support staff. The second was assigning almost all instructional staff to work directly with students. As a consequence, all students experienced much smaller class sizes (18 at CPESS and 25 at International High), while their teachers also had much more planning and professional development time. Because teachers taught a smaller number of longer periods, pupil loads were also reduced: teachers at CPESS saw 36 students and those at International High saw 75 students in a given term. The schools achieved this by reducing specialization, reorganizing student groups and teaching structures, and investing heavily in professional development. The strategies the two

[3] In a related study (Darling-Hammond, 1997), Darling-Hammond and colleagues also examined a smaller, more affluent suburban high school of about 1,600 students. However, the large difference in student populations between this school and the redesigned schools in New York City made this school inappropriate for the current analysis.

[4] Because International High has a unique student population comprised of 90 percent Title 1 eligible and 100 percent limited English proficient students, an analogous traditional school could not be found for comparison. Instead, we used the New York City staffing guidelines, as outlined in the New York City publication *Comparative Analysis of the Organization of High Schools, 1992–93*, to estimate staffing for students identified for special needs programs.

Figure 4.—High performing versus traditional secondary school

Resource allocation principles	School measure	Central Park East Elementary	International High	Traditional
Reduction of specialized programs to provide more individual time for all in heterogeneous groups	a. Students per instructional staff member	10.2	10.2	14.7/13*
	b. Students per full-time teacher	13.3	15.8	23.6
	c. Average size of regular instructional group	18	25	33.4
	d. Percent teachers in regular instructional groups	89%	100%	70%
More flexible student grouping by school professionals	e. Percent students in target size grouping	100%	100%	60%
	f. Average size of advisory group	15	12	29 (home room)
Structures to create more personalized environments	g. Student loads per term	36	75	167
	h. Percent professional staff serving as instructors/advisors	100%	100%	65%
Longer and more varied blocks of instructional time	i. Average length of instructional period (in minutes)	120	70	42
More common planning time for staff	j. Common planning minutes/week	450	350	0
	k. Length of longest planning period (in minutes)	120	140	42

* A traditional high school that had a 100 percent Limited English Proficient student population like that at International would receive additional staff to reduce its student/teacher ratio for those students to 13:1.

SOURCE: Unpublished tabulations.

schools used reflected their instructional purposes and philosophies.

Reduced Specialization

CPESS reduced specialization in a host of ways in order to create smaller teacher-student loads and focus resources on academic subjects. CPESS follows the principles embraced by the Coalition of Essential Schools, one of which is that "less is more." Instead of aiming for broad coverage of content, CPESS has organized its curriculum around five "Habits of Mind" which encompass the abilities to weigh evidence, take varying viewpoints into account, see connections and relationships, speculate about possibilities, and assess value. These goals are reinforced in every course and in a comprehensive portfolio assessment system. The school concentrates its resources on a common core curriculum in grades 7–10, and uses college and community resources to expand curriculum options in the upper grades.

At the time of the study, all students took academic subjects in heterogeneous groups of about 18. Students in divisions I and II (grades 7–10) took two, two-hour academic courses each day, humanities and math/science plus a foreign language offered in a shorter period before school. All full-time teachers in these grades, with the exception of two special education resource room teachers, taught one of the two interdisciplinary courses. The resource room teachers helped students with their regular classroom work, thereby reinforcing rather than fragmenting students' learning. Students and teachers told us that the extraordinary outcomes achieved at CPESS were made possible by these small class sizes and the continuity provided by (a) having a team of teachers stay with the same students for two years (b) providing extra help to special needs students connected to their regular classroom work. In addition, teachers were grateful for opportunities to extend their disciplinary knowledge by planning together in subject area teams and to extend their knowledge of students by also working together in "house" teams organized around the teaching of shared groups of students (Darling-Hammond, Ancess, and Falk, 1994).

In the Senior Institute (grades 11–12), the school reduced its need for specialization by working out advanced course-taking opportunities for students at local colleges to supplement required courses and electives offered in the school. All students took at least two college courses during their last two years of high school, along with an internship in a local business or community organization. These strategies expanded students' academic and vocational opportunities while freeing staff time for the one-on-one advisement needed to support students in completing extensive research projects and the other portfolio entries they would defend before their graduation committees.

Language instruction and some electives were provided through outside contracts. Although students could opt to take more advanced courses in specific subjects, there was no tracking, no separate Title 1 programs, and no separate bilingual program. Instead of hiring guidance counselors, each teacher was responsible for counseling 12 students during scheduled advisory periods. The small size, personalization, and team organization also eliminated the need for attendance officers, deans of discipline, assistant principals, and supervisors, roles that deflect resources away from teaching positions in traditional high schools.

The organization of resources at International High also followed its educational mission—the education of recent immigrants—and its philosophy, which includes the following principles:

- Language skills are most effectively learned in context and when embedded in a content area.
- Successful educational programs emphasize rigorous standards coupled with effective support systems.
- Attempts to group students homogeneously preclude the way in which adolescents learn best (i.e., from each other).
- The carefully planned use of multiple learning contexts in addition to the classroom (e.g., learning centers, career internship sites, field trips) facilitates language acquisition and content area mastery.

The existence of clear school goals and a consensus about strategies enhanced International High's ability to design a coherent, carefully configured organization. The school reorganized its programmatic resources around 12 interdisciplinary themes. Six self-managing instructional teams, called "clusters," were composed of four to six teachers plus guidance and paraprofessional staff who developed two thematically based courses of study (e.g., Motion, Visibility), which integrated four subject areas (e.g., literature, global studies, mathematics, and physics). The team took responsibility for the total educational experience of about 75 students during a 13-week course of study. Students chose three of these thematic courses each year.

The existence of clear school goals and a consensus about strategies enhanced International High's ability to design a coherent, carefully configured organization.

All teachers taught heterogeneous groups of students that included all native languages and all grades, economic levels and ability levels. The faculty integrated English as a Second Language (ESL) techniques into their content-area courses while providing students with opportunities to further develop their language skills with instructors outside the core curriculum and in learning contexts such as internships outside the school. The success of this strategy is illustrated by the fact that virtually all International High students pass the New York

State competency tests in English as well as other subjects, and virtually all are accepted to college prior to graduation. Teachers designed the cluster schedule because they felt that planning and teaching in interdisciplinary teams would enable them to have more influence on student learning. They found that the school's first cluster experiment produced much greater levels of student success than had independent classes, because teachers' efforts were jointly planned and cumulative, and teachers could deal with students' needs and problems in a concerted fashion (Darling-Hammond, Ancess, and Falk, 1995).

In both schools, the integration of previously specialized resources, along with the investment of greater resources into teaching positions rather than nonteaching positions, translated into much lower pupil loads and more opportunity for individual student attention than in the traditional school. In addition to focusing resources on instructional positions, the sample schools used most of their teaching resources in one core academic program in which all students participated, rather than using special program resources for add-on remedial or special education programs. CPESS used 89 percent of teaching resources in its core instructional program, while International High used all staff in the core program, as compared with roughly 70 percent of teachers working in regular instruction in the traditional high school.

Smaller class sizes were also achieved by creating a broader role for professional staff in the restructured schools, rather than using specialists to perform "non-classroom" functions. Staff acknowledged this tradeoff in a set of "understandings that underlie professional staff work at CPESS" which includes the following statement:

> In return for smaller class sizes (maximum 20) and smaller total student rolls, teachers will work with students for a total of 22 hours per week in classes, advisories, or tutorials; conducting seminars; overseeing projects; giving lectures; or advising and coaching individual students (CPESS, 1991).

More Flexible Student Grouping

Reducing the number of programs, courses, and levels made it easier for the sample schools to control the size of instructional groups. As table 1 shows, although 64 percent of all classes in the traditional high school had 29 to 34 students, 21 percent of classes were smaller than 25. Class sizes were higher in regular education academic classes than in nonacademic classes.[5] In contrast, CPESS and International High place all of their students in target size groups.

Even more flexible grouping strategies were found in CPESS's Senior Institute (grades 11–12), where teachers and students focused substantial attention on preparing the graduation portfolio and applying to colleges. Time was allocated to allow teachers to provide coaching and support for independent study. A typical teacher would teach two classes over about 12 hours per week. He or she would also spend 4 to 5 hours per week supervising independent projects, another 4 to 5 hours in

Table 1.—Class sizes in the traditional high school

Class size	Academic	All classes
0–19	6 percent	8 percent
20–24	7 percent	13 percent
25–28	13 percent	13 percent
29–34	72 percent	64 percent
Over 34	3 percent	3 percent

SOURCE: Unpublished tabulations.

[5] New York City Schools, *Comparative Analysis of the Organization of High Schools, 1992–93*, p. 82–92.

advisory working with his or her 12 advisees on academic and personal concerns, and another 3 1/2 hours per week for one-on-one help to students. Class periods varied in length depending on their purpose. While students took courses and completed internships outside the school, teachers got time to work and plan together.

Structures to Create Personal Relationships

The sample schools were organized more like elementary schools than most secondary schools, featuring small teaching units and closer, more sustained relationships between teacher and student. In addition to the smaller class sizes, both sample schools used "advisory groups" as a key strategy for maintaining ongoing relationships with students. Each professional staff member worked with a group of 12 to 15 students and their families. The use of all professional staff, not just teachers, allowed advisory groups to be smaller than average class sizes. Advisory groups met for approximately 4 hours a week at CPESS, and for about 2 hours weekly at International High. Teachers and advisors used the time for individual study; to discuss health, social and ethical issues; and for individual and group advising and counseling. The advisor served as the "expert" on the student, meeting regularly with the family and other teachers to discuss the student's needs and progress, coordinating parent conferences and the preparation of narrative assessments of student work.

. . . both restructured high schools created longer periods and more flexible schedules to accommodate more ambitious kinds of work and to allow more time for teachers' to support student learning.

Whereas all professionals in the two restructured schools worked on a regularly scheduled basis with groups of students, only 65 percent of the professional staff at the traditional high school had regularly scheduled contact with a continuing group of students. While guidance counselors and other support personnel worked intensively with some students, they did so on a reactive, usually sporadic basis that was not designed to create close, long term relationships.

Longer and More Varied Blocks of Instructional Time

In contrast to the traditional high school's seven 42-minute periods each day, both restructured high schools created longer periods and more flexible schedules to accommodate more ambitious kinds of work and to allow more time for teachers' to support student learning. At CPESS, students in grades 7–10 had two 2-hour blocks of humanities and math/science each day. Since these two teachers worked together as a team, they could vary the split of time between the two to accommodate daily lesson plans. In addition, one morning per week students spent 2 1/2 hours in a community service project while their teachers were engaged in curriculum planning. Other coursework, such as language instruction, took place in smaller (usually 1 hour) blocks of time. In the Senior Institute, classes varied in length from 1 to 2 hours, while advisement sessions, internships, and independent work time were scheduled for longer blocks of time to allow students to undertake extended work with adequate coaching and time for research.

At International High, students typically had four courses each of which met for 70 minutes four times per week. They also had a 2 hour internship and an hour long seminar each week. Because each cluster of four teachers controlled their shared students' entire time schedule during a 13-week cycle, they could vary time across classes each day as needed for the work in which students were engaged.

More Common Planning Time

Both sample high schools created structures that demand and allow much more common planning time. In addition to individual "prep" time, CPESS teachers spent on average 7.5 hours per week in scheduled common planning time. CPESS used four strategies to create this time. First, teachers met with their disciplinary teams for 2 1/2 hours of weekly curriculum planning while students were in community service placements. Second, while teaching fellows and other professionals provided coverage, teachers had from 1 1/2 to 3 hours each week to meet with fellow "house" teachers and with students individually. Third, students' hours were increased during the week so they could be dismissed at 1:00 p.m. on Fridays to create time for a weekly 2 hour staff meeting. Finally, as its governance plan states, "the full staff agrees to meet during hours when the students are not in attendance to complete necessary business." In addition to the Friday meeting, teachers attended a regular Monday meeting from 3:00 p.m. to 4:30 p.m.

At International High, teachers had 140 minutes each week to plan with their cluster while students participated in college courses and other activities. During a half day each week set aside for club activities for students, teachers staff-initiated professional development. In addition, teachers had a 70 minute individual planning period each day which often coincided with that of other members of their team. These models offer stark contrast to the traditional high school model in which teachers had one or two separate 42-minute periods free from instruction, one often devoted to nonacademic duties and the other organized as an individual preparation period.

> *Both of the sample schools . . . focused teaching resources on core academic subjects by contracting with outside providers for electives and non-academic subjects.*

Creative Definition of Staff Roles and Work Day

Both of the sample schools have made many changes to the roles of teachers and the typical organization of the teacher work day in ways that enable greater personalization for students. They focused teaching resources on core academic subjects by contracting with outside providers for electives and non-academic subjects. CPESS increased resources for teachers by incorporating counseling and advising into the teaching role. In both schools, teachers are involved in curriculum and assessment development, hiring and evaluation of staff, schoolwide decisionmaking, and staff development. These broader professional roles not only reduce the need for nonteaching specialists assigned to manage and oversee teaching, they also enrich teachers' knowledge and skill by giving teachers continuous opportunities to reflect on their teaching and learn from one another, thus expanding the expertise available in the school (Darling-Hammond, 1997).

Policies, Regulations, and Contractual Issues

To accomplish these things, the sample schools directly challenged policies, regulations, and collective bargaining agreements. First, most of the schools changed the contractually defined teacher work day and contractual rules for such matters as seniority transfers. Second, in breaking down barriers between programs, age groupings, and subjects, they confronted staffing formulas, program administration rules, and, sometimes, teacher licensing categories. Third, many of these schools redefined both teaching and non-teaching positions to create new jobs which do not fit neatly into existing contractually defined categories.

Collective bargaining agreements in most districts clearly define the teacher work day, outlining the hours teachers are required to work and limiting the number of required afternoon and evening meetings. Most go further to specify the number of minutes of time teachers must have free for lunch and planning activities, and some limit the number of hours in a row teachers can be involved in instruction, making it difficult to create connected blocks of planning time. It is easy to understand the reasons for these provisions, but it is also clear that new strategies are needed for schools in which teachers jointly develop curriculum and manage their own and students' time.

As teaching jobs are broadened, schools can run into state, district, and collective bargaining restrictions. Using teachers across subjects or programs can require waivers. Mary C. Lyons use of three special education teachers and three regular education teachers to teach integrated classrooms of special needs and regular education students required waivers from the Boston teachers contract and Massachusetts state certification laws. The principal argued that she knew how to identify individuals with experience and disposition to handle both special education and regular education students. She developed a plan to create a team structure which took advantage of teachers' diverse skills and a professional development plan for each teacher and for the entire school, so that they would develop the skills they needed.

Schools also can run into certification problems in moving to interdisciplinary instruction, because many collective bargaining agreements and state regulations require teachers to hold certification in more than one subject to teach humanities or math/science in middle or high schools. Finding individuals with the subject and pedagogical knowledge to combine these subjects effectively is obviously critical to successful interdisciplinary instruction. Although certification in both fields is one indicator of this ability, it is not the only means for developing expertise in a second field. At CPESS, teachers with a background in one field plan in curriculum teams (a math/science team and a humanities team) that provide the additional disciplinary expertise they need to handle the breadth the core courses require.

Finding individuals with the subject and pedagogical knowledge to combine these subjects effectively is obviously critical to successful interdisciplinary instruction.

New job positions and hiring arrangements also confront some collective bargaining agreements and traditional allocation guidelines. For example, Mary C. Lyons and CPESS created a different kind of Instructional Assistant by using teaching interns—graduate students who are preparing to become teachers—instead of paraprofessionals or untrained support staff. In addition, three of the sample schools received waivers from collective bargaining agreements to use outside contractors for specific kinds of instruction.

Selection and retention of teachers with the required qualities and experience to match these school designs is critical to their success, yet many districts treat staff as interchangeable when they make assignments and move staff on seniority transfers. Some districts have solved this problem by creating alternative personnel tracks for specially designated schools. Cincinnati has done this for Paidea and Montessori schools. In Boston, schools negotiate control over the hiring process on a position by position basis. In New York, recent contract negotiations have allowed for teams of teachers, including principals and union representatives, to select their new colleagues in the growing number of schools that have a distinctive missions. With the recent creation of over 100 new small schools in the city joining the substantial number of longer standing alternative schools, this provision paves the way for widespread use of new staffing models.

Finally, teacher contracts, district policies and state regulations often define class size maximums by program, grade level, and sometimes subject. State guidelines specify the size of classroom for students at each level of special education classification. Schools can depart from these regulations if parents, teachers and special education professionals agree to an individual education plan that educates the student in a larger, more inclusive setting. Designs like that used at the Mary C. Lyons school require intense communication with students and parents to create understanding of the new approach and strategies to insure appropriate additional support for the students. They also demand that state and district officials work with schools to allow educationally sound designs.

District student and teacher assignment policies can frustrate attempts to use teachers differently. In the sample districts, schools moving students from more restrictive special education settings into the regular classroom sometimes faced a potential loss of teachers because special education staff were allocated based upon the number of students requiring separate education. If schools integrate students back into the regular classroom and resources are therefore reduced, the regular teacher in whose class the special education student now spends most of his time receives no extra resources and no reduced student load. In these cases, schools can find that regular education classrooms grow more unruly and crowded while the case loads of special education teachers decline. Over time, schools should find ways of shifting resources back into the classroom without losing special education expertise. To respond to this problem, Boston has adjusted its staffing formula to allow schools to use the resources for special needs students in inclusive settings.

Conclusion

Although these five high performing schools look very different from one another, they have all redesigned the way they allocate teaching resources to meet student needs and to create the time teachers need to implement a new vision of schooling. They demonstrate how schools can reallocate resources to implement new designs. The framework presented here aims to provide researchers and practitioners with a way to examine systematically the possibilities for reallocation and to measure their impact. Changing school organizations to better fit an instructional vision will require schools to confront long traditions and a host of state, district, and union policies and practices that conflict with many of the changes outlined here. These barriers can loom large. But, the biggest constraint may be lack of vision about the concrete changes in school organization that can create a more professional organization and improve student achievement. This paper aims to provide some clear, detailed examples schools might use to develop such a vision, including goals for student achievement along with educational strategies and an organization to accomplish these goals.

Changing school organizations to better fit an instructional vision will require schools to confront long traditions and a host of state, district, and union policies and practices that conflict with many of the changes outlined here.

The variety of models presented here suggests that resource reallocation and the design of an instructional vision and strategy are intertwined. There is little rationale for restructuring resources without an underlying educational design. At the same time, none of these models could have accomplished their goals without making changes in the use of resources. As these models and others are tested against evidence of improved student performance, one could imagine states and districts working with schools to adopt proven designs, through a conscious process of changing resource allocations, practices, and regulations at each level. As part of the process of choosing an appropriate design, schools might undertake a comprehensive review of how their practices, resources, knowledge and skills would need to change to implement a new model. Principles of resource allocation and indicators of their use could form the basis for tools which help schools and districts understand their progress. Districts could then organize their work to support these plans and develop strategies for helping schools make changes. This would include changes in state and district policies that may produce obstacles to alternative forms of organization.

The schools studied here have only touched the surface of opportunities for rethinking the way school resources are used; they have largely worked within existing salary structures and have not much explored the use of technology in the classroom. Nevertheless, they foreshadow the many ways schools may rethink existing resources to create more personalized education for students and more professional responsibility and growth for teachers.

References

Anderson, R. H. and Pavan, B.N. 1993. *Nongradedness: Helping It to Happen.* Lancaster, PA: Technomic Publishing.

Ball, D., and Cohen, D. In press. "Developing Practice, Developing Practitioners: Toward A Practice-Based Theory of Professional Education." *The Heart of the Matter: Teaching as the Learning Profession,* ed. L. Darling-Hammond and G. Sykes (San Francisco: Jossey Bass).

Braddock, J.H., and McPartland, J.M. 1993. "The education of early adolescents", *Review of Research in Education,* ed. L. Darling-Hammond, Washington, D.C.: American Educational Research Association. 19: 135–170.

Carroll, J.M. 1994. "Time for the Copernican Plan," *Phi Delta Kappan, 76*: 104–114.

Central Park East Secondary School. August 1991. Understandings That Underlie Professional Staff Work at CPESS. Unpublished manuscript.

Central Park East Secondary School. May 7, 1990. Basic Governance Plan. Unpublished manuscript.

Commission on Chapter 1. 1992. *High Performance Schools: No Exceptions, No Excuses.* Washington, D.C.

Darling-Hammond, L. 1997. *The Right to Learn: A Blueprint for Creating Schools that Work.* San Francisco: Jossey Bass.

Darling-Hammond, L. 1996. "Restructuring Schools for High Performance," ed. Susan Fuhrman and Jennifer O'Day. *Rewards and Reform,* (San Francisco: Jossey Bass).

Darling-Hammond, L., Ancess, J., and Falk, B. 1995. *Authentic Assessment in Action: Studies of Schools and Students at Work.* NY: Teachers College Press.

Gottfredson, G.D. and Daiger, D.C. 1979. *Disruption in 600 Schools.* Baltimore, MD: The Johns Hopkins University, Center for Social Organization of Schools.

International High School (IHS). 1995. *International High School 1993–94 End of Year Evaluation Report.* NY.

Lee, V. E. and Smith, J.B. Fall 1994. "High School Restructuring And Student Achievement: A New Study Finds Strong Links." *Issues in Restructuring Schools, 7,* (1–5, 16). Madison, WI: Center on Organization and Restructuring of Schools.

Lee, V.E. and Smith, J.B. 1995. *Effects Of High School Restructuring and Size On Gains in Achievement and Engagement For Early Secondary School Students.* Madison, WI: Center for Education Research, University of Wisconsin.

Lee, V. E., Bryk, A., and Smith, J. B. 1993. "The Organization of Effective Secondary Schools," ed. L. Darling-Hammond, *Review of Research in Education, 1993.* Washington, D.C.: American Educational Research Association.

Little, J. W. 1993. "Teacher Development in a Climate of Educational Reform." *Educational Evaluation and Policy Analysis*, 15, (2): 129–51.

Miles, K. H. 1995. "Freeing Resources for Improving Schools: A Case Study of Teacher Allocation in Boston Public Schools," *Educational Evaluation and Policy Analysis*, 17: 476–493

Miles, K. H. 1997a. *Spending More at the Edges: Understanding the Growth in Public School Spending from 1967 to 1991*. UMI Press.

Miles, K. H. 1997b. "Finding the Dollars to Pay for 21st Century Schools: Taking Advantage of the Times," *School Business Affairs*, June.

National Center for Education Statistics. 1994. *The Digest of Education Statistics*. Washington, D.C.: Government Printing Office.

National Commission on Teaching and America's Future. 1996. *What Matters Most: Teaching for America's Future*. NY.

New York City Board of Education. 1993. *High Schools Allocation Information, Guidelines and Instructions, 1994–95*. NY.

New York City Board of Education. 1993. *Comparative Analysis of the Organization of High Schools, 1992–93*. NY.

Organization for Economic Co-operation and Development (OECD). 1995. *Education Indicators*. Paris.

Rettig, M.D., and Canady, R.L. 1993. Unlocking the Lockstep High School Schedule. *Phi Delta Kappan*, 70, 310–314.

Rothstein, R. with Karen Hawley Miles. 1995. *Where's the Money Gone?* Washington, D.C.: Economic Policy Institute.

Sarason, S.B. 1982. *The Culture of School and The Problem of Change*. Boston: Allyn and Bacon.

Sizer, T. 1992. *Horace's School: Redesigning the American High School*. Boston: Houghton Mifflin.

Slavin, R.E. 1990. Achievement Effects of Ability Grouping in Secondary Schools: A Best Evidence Synthesis. *Review of Educational Research*, 60(3), 471–500.

Soo Hoo, S. 1990. School Renewal: Taking Responsibility for Providing an Education of Value. *Access to Knowledge: An Agenda for Our Nation's Schools*, edited by J.L. Goodlad and P. Keating, 205–222. New York: The College Entrance Examination Board.

Stigler, J.W. and Stevenson, H.W. 1991. "How Asian Teachers Polish Each Lesson to Perfection," American Educator, 12–47.

Tyack, D. and Tobin, W. 1994. "The 'Grammar' of Schooling: Why Has it Been So Hard to Change?" *American Educational Research Journal*, 31, (3): 453–479.

NOTE: We thank the staff at the five schools we studied for sharing with us their successes and frustrations. Special thanks to Patrick McNeeley, principal at Quebec Heights Elementary School in Cincinnati, Mary Nash at the Mary C. Lyons Model Elementary School in Boston, Myra Whitney, principal of Douglass Elementary School in Memphis, Paul Schwarz at Central Park East Secondary School, and Eric Nadelstern and Ruth Ellen Weiner at International High School. We also thank Lori Chajet and Peter Robinson who helped collect much of the secondary school data as part of their research with the National Center for Restructuring Education, Schools, and Teaching.

Financing Education in the District of Columbia from the Perspective of the Financial Authority

Dr. Joyce Ladner

District of Columbia Control Board

About the Author

Joyce A. Ladner is a Senior Fellow at the Brookings Institution in Washington, DC. She is also a Professor of Sociology (on leave) in the Howard University School of Social Work. She served as Interim President of Howard from 1994 to 1995 and Vice President for Academic Affairs from 1990 to 1994.

In 1995, she was appointed by President Bill Clinton to the District of Columbia Financial Control Board that was created to provide oversight and guidance in restructuring the city's finances and management. Ladner is the control board member with oversight for the public school system.

A nationally known scholar, Dr. Ladner has received the Doctor of Humane Letters from Howard University and Tougaloo College as well as numerous other academic and professional awards. She has published extensively in the areas of diversity, higher education, urban issues, public policy, family and gender problems. She has edited, authored, or co-authored four books, *Tomorrow's Tomorrow: The Black Woman, The Death of White Sociology, Mixed Families: Adopting Across Racial Boundaries*, and *Lives of Promise, Lives of Pain: Young Mothers After New Chance*. She is currently writing her memoirs on growing up in Mississippi during the civil rights era of the sixties, and a book titled *The Ties That Bind: Timeless Values for African American Families* to be published in 1999. Dr. Ladner has also appeared frequently on television news programs and National Public Radio, where she comments on social issues.

Dr. Ladner received her B.A. in sociology from Tougaloo College (1964), her M.A. (1966) and Ph. D. (1968) in sociology from Washington University in St. Louis. Her memberships include the Council of Foreign Relations, the American Sociological Association, and member of the boards of the Congressional Economic Leadership Institute, the U.S. Attorney General's Advisory Council on Violence Against Women, the Center for National Policy (vice-chair), the Washington Urban League, the Washington Women's Forum, the Coalition of 100 Black Women, and Suited for Change (an organization that helps poor women enter the job market).

Financing Education in the District of Columbia from the Perspective of the Financial Authority

Dr. Joyce Ladner

District of Columbia Control Board

Introduction

The District of Columbia Financial Authority was created by the U.S. Congress in 1995 to repair the District of Columbia's failing financial condition and to improve the management effectiveness of government agencies. The Authority has identified public education, along with public safety and some activities in public works, as priority areas.

In November 1996, the Authority overhauled the governance and administration of the District of Columbia Public Schools (DCPS). After issuing a scathing report on the quality and management of the District's schools, the Authority removed the superintendent and stripped the Board of Education of most of its power to control the schools. In their place, the Authority appointed a new Chief Executive Officer and an Emergency Board of Trustees to run and oversee the schools.

This article describes the condition of the District's public schools that caused the Authority to take such a bold action. Next, it outlines the unique nature of the financing of the District's public schools, and then briefly discusses higher education in the District and the Authority's efforts to address problems at the University of the District of Columbia.

The Failure of the District of Columbia Public Schools

In fall 1996, after extensive study by the Authority's staff and consultants and after conducting several public hearings, the District of Columbia issued a report entitled *Children in Crisis: A Report on the Failure of the D.C's Public Schools.* The full text, as well as other D.C. Financial Authority information and reports are available on the Internet at www.dcfra.gov.

In that report, the Authority concluded that the deplorable record of the District's public schools by every important educational and management measure had left the system in a state of crisis. DCPS was simply failing in its mission to educate the children of the District of Columbia. In virtually every area, and for every grade level, the system failed to provide the District's children with a quality education and safe environment in which to learn.

Abysmal Education Outcomes

DCPS fails to teach its pupils even the basics of education. As a result, the system's students score significantly lower on standardized academic achievement tests than their peers in comparable districts around the nation.

Figures 1 and 2 benchmark DCPS's performance against other school systems and figure 3 measures DCPS's performance by ward over time. Figure 1 shows that the average SAT scores of students taking the test in school districts surrounding D.C. and peer school districts across the country exceeded those of students attending D.C. public schools.

Figure 2 shows that the performance of D.C. public school students on the National Assessment of Educational Progress (NAEP) lagged behind both the average scores of students at the national level and those of students in the northeastern states. In 1994, only 22 percent of DCPS's fourth-grade students scored at or above the basic[1] level—a decrease of 6 percent from 1992.

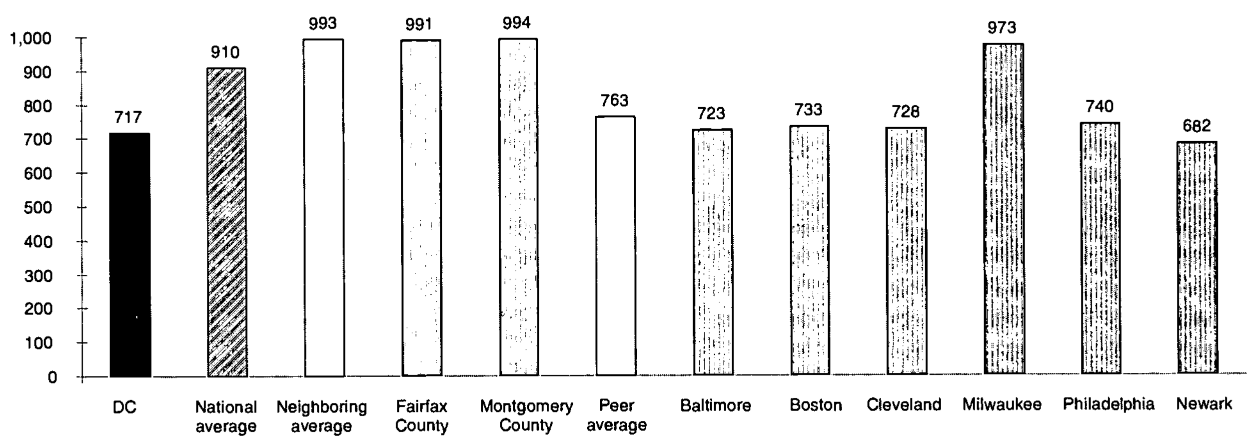
The results from Stanford Achievement Tests administered in May, 1997, indicate that there has been little improvement since 1994. Thirty-three percent of third-graders were below basic levels in both reading and mathematics. Twenty-nine percent of eighth-graders were below basic levels in reading, and 72 percent of eighth-graders were below basic levels in mathematics.

Performance varies extensively among sections of the District of Columbia (wards); the least affluent wards have experienced the greatest decline in test scores over the last 5 years. Figure 3 shows that while test scores in the more affluent sections of the District have remained the same or improved slightly, scores in the poorer wards have declined dramatically. Several schools in wards seven and eight have seen startling declines in test scores of 15 to 20 percentage points or more. The low achievement levels attest to the fact that thousands of children, especially those in the less affluent sections of the District, are not being taught the fundamental skills necessary to succeed after they leave DCPS.

Drop Out Rates Are Too High

As academic performance continues to decline, many students are dropping out or leaving DCPS for neighboring districts and private schools. Figure 4 illustrates the dramatic migration of students out of the public school system—40 percent of high school students either dropped out or left the District's public schools between 1989 and 1995.

Figure 1.—Average composite SAT scores: 1995

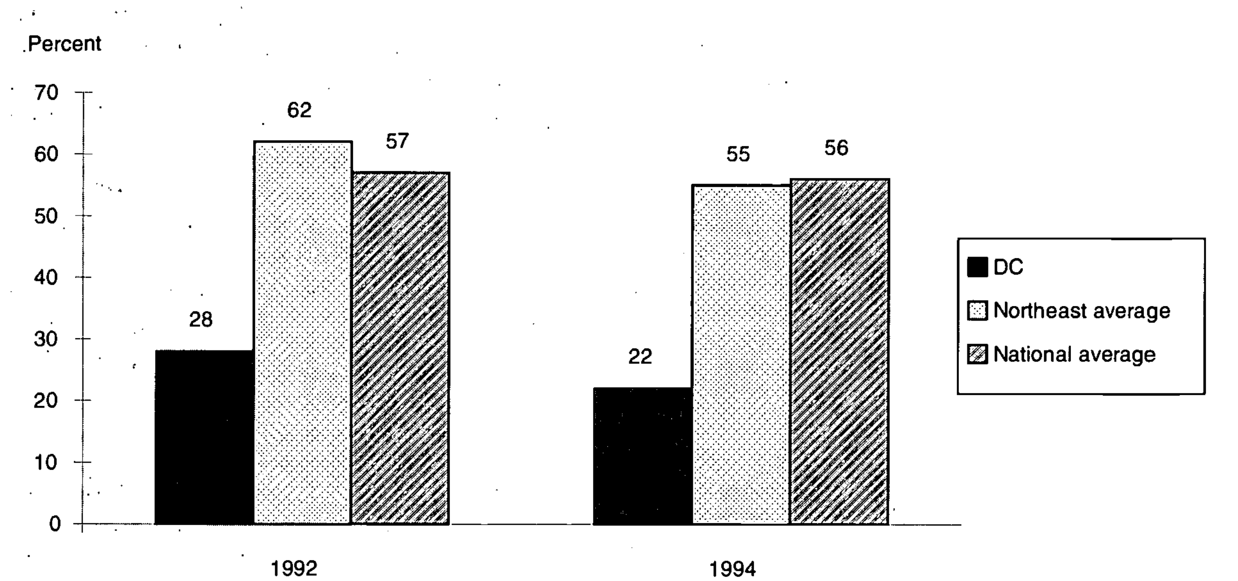


NOTE: Information is not provided for Chicago, New Orleans, and Detroit. The SAT is not widely taken in Chicago and New Orleans because the local colleges and universities do not require it. Information on SAT scores in Detroit was unavailable.

SOURCE: Information reported by individual districts.

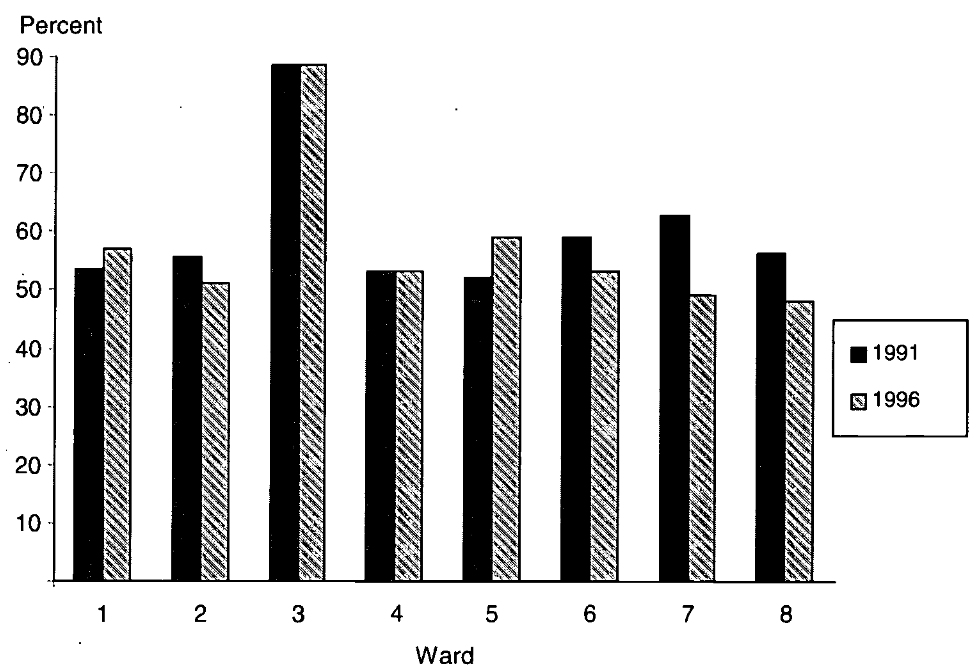
[1] "Basic" is defined as a partial mastery of knowledge and skills that are fundamental for satisfactory work at this grade level. "Below basic" is defined as little or no mastery of fundamental knowledge and skills for this grade level.

Figure 2.—Percentage of fourth-grade students scoring at or above the basic level on the NAEP trial state assessment of reading achievement: 1992 and 1994



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress.

Figure 3.—Changes in sixth-grade test scores, by District of Columbia wards: 1991 to 1996



NOTE: Average of reading and mathematics percentiles from the sixth grade.

SOURCE: Office of Educational Accountability, District of Columbia Public Schools.

An increasing number of high school students left DCPS between 1993 and 1995. Figure 5 shows the annual attrition for the classes of 1993, 1994, and 1995. On average, 24 percent of students left their class in 10th grade, 23 percent in 11th grade, and 6 percent in 12th grade. Drop-out rates are even greater in the transition from elementary to the junior high level as parents who can afford it send their children to private schools.

Unsafe Environments Disrupt Learning

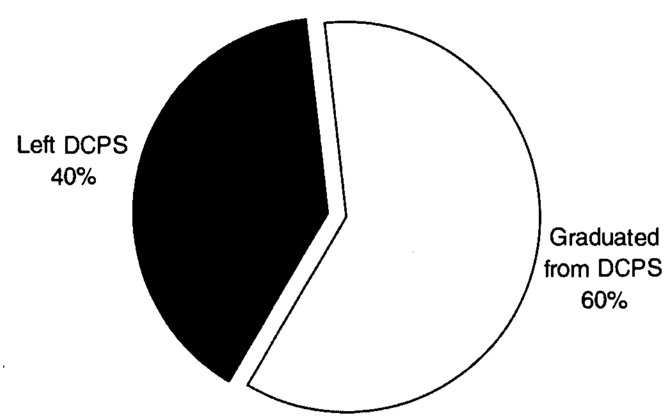
The District's schools, like other urban environments, have serious problems in terms of both security and deteriorating facilities.

A National Center for Education Statistics (NCES) survey of teachers revealed that more instructors from DCPS than instructors from other states believed that a variety of serious problems affected their schools. Problems included: lack of parent involvement, unpreparedness to learn, tardiness, disrespect for teachers, student absenteeism, student apathy, and verbal abuse of teachers (figure 6).

Compared with the national average, more DCPS high school teachers and students report being threatened with violence. Figure 7 shows that according to the 1995 National Education Goals Report, 26 percent of DCPS teachers reported that they were threatened, injured, or physically attacked by a student in the past 12 months, compared with a national average of 14 percent. Other state-level violence statistics are also bleak.

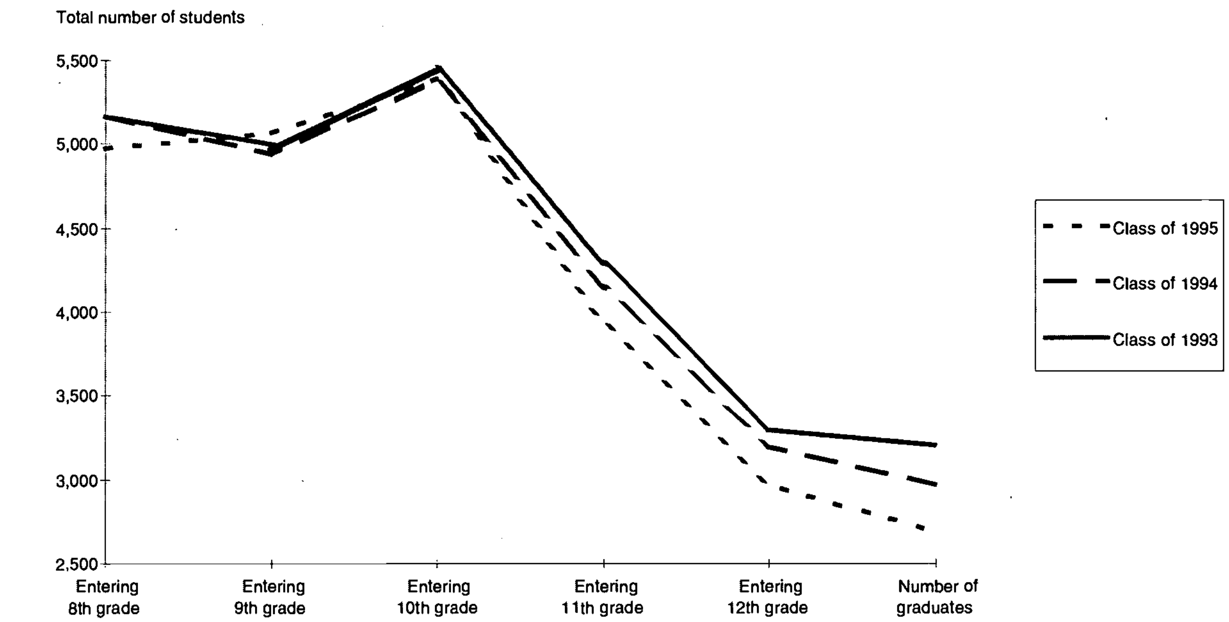
The infrastructure of the District's public schools is collapsing. The alarming condition of facilities leaves students exposed to discomfort and even to potential harm—boilers burst, roofs leak, firedoors stick, bathrooms crumble, and poor security permits unauthorized individuals to gain access, threatening the safety of students. Such conditions make it almost impossible to focus on the primary mission of educating the children. Figure 8 compares the condition of the District's schools to other schools in the nation.

Figure 4.—Percentage of students beginning ninth grade who eventually left DCPS or who and graduated from DCPS: 1989–95



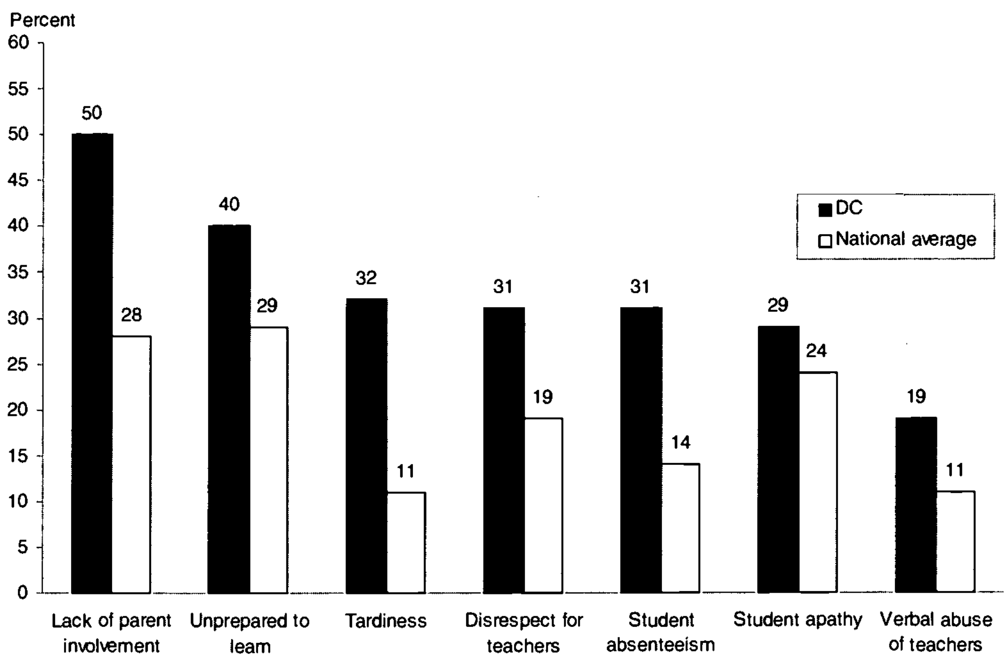
SOURCE: District of Columbia Public Schools (DCPS) Dropout and Migration Statistics 1991–95.

Figure 5.—Enrollment changes for students in grades 8 through 12 in District of Columbia public schools (classes of 1993–95)



SOURCE: District of Columbia Public Schools Dropout and Migration Statistics (1991–95).

Figure 6.—Percentage of public school teachers perceiving issues as serious problems affecting their schools



SOURCE: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey, 1993–94.

Mismanagement Undermines Learning

The inability of the District's schools to effectively implement long-term education and operational plans leaves students without teachers or classrooms, textbooks unordered or lost in warehouses, teachers untrained and uncertified, and students who are disabled without access. Additionally, poor resource allocation distorts priorities, ensuring that educational needs go unmet even when funds are available.

Two critical areas that are mismanaged are personnel and procurement. DCPS' personnel operations are in disarray. Every aspect is problematic:

- inability to identify how many employees work in the schools;
- lack of adequate administrative controls;
- incomplete planning, and
- inability to understand or relate the impact of personnel on the educational mission.

Procurement is just as bad. Poor contract management has left an indelible mark on the District's children—who, among other things, have been forced to eat cold cereal for lunch and have been subjected to unqualified individuals operating school facilities.

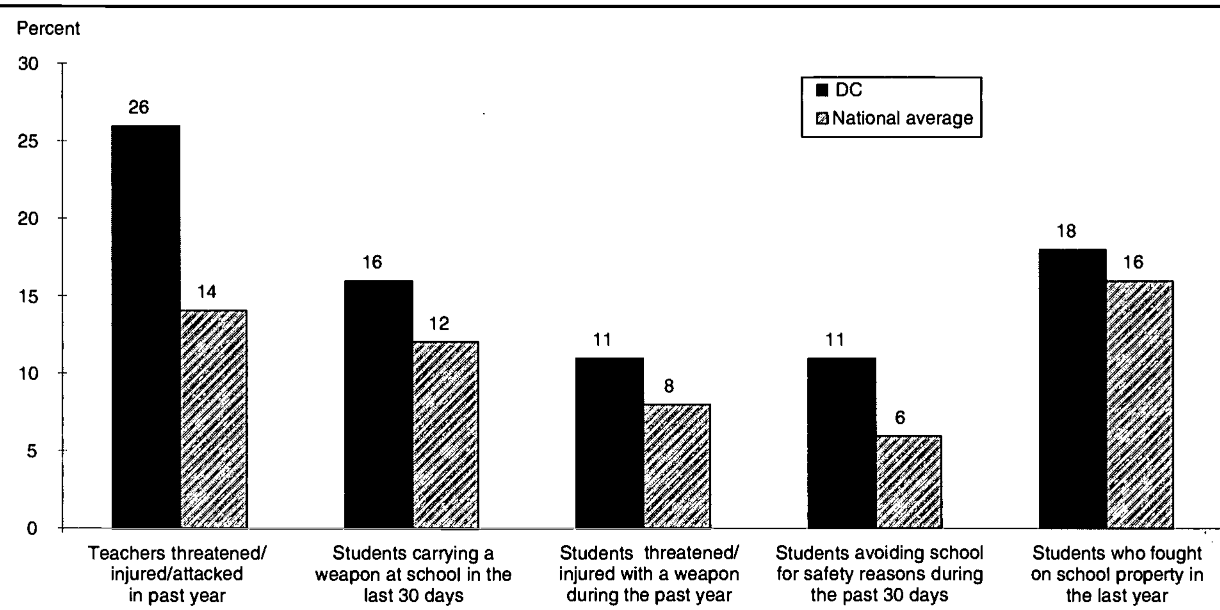
Basic data are also not available. The most critical is that the District public school system does not know how many students it has. Estimates vary between 65,000 and 81,000 students; the discrepancy is alarming. NCES found a discrepancy of 20.6 percent between the 1990 census and the number of students reported by DCPS (figure 9). This discrepancy is significantly greater than any of the other major jurisdictions in the Greater Washington Metropolitan area.

Lack of basic data on students and employees makes budgeting next to impossible.

Financing Education in the District of Columbia Public Schools

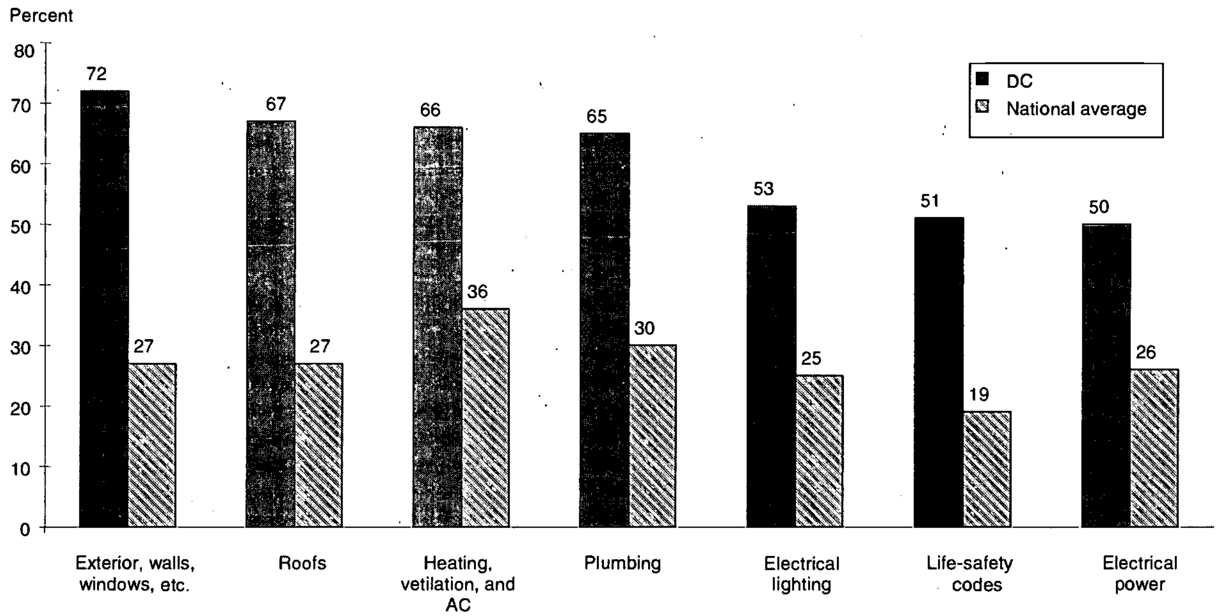
The District has a unique financing arrangement. Not only must the District carry out the roles of a city, a county, and a state, but it's budget must be directly approved by Congress. The result is confusion and, in terms of Congress, a level of meddling in local affairs that no other jurisdiction in the coun-

Figure 7.—Percentage of high school teachers and students reporting violence and safety issues



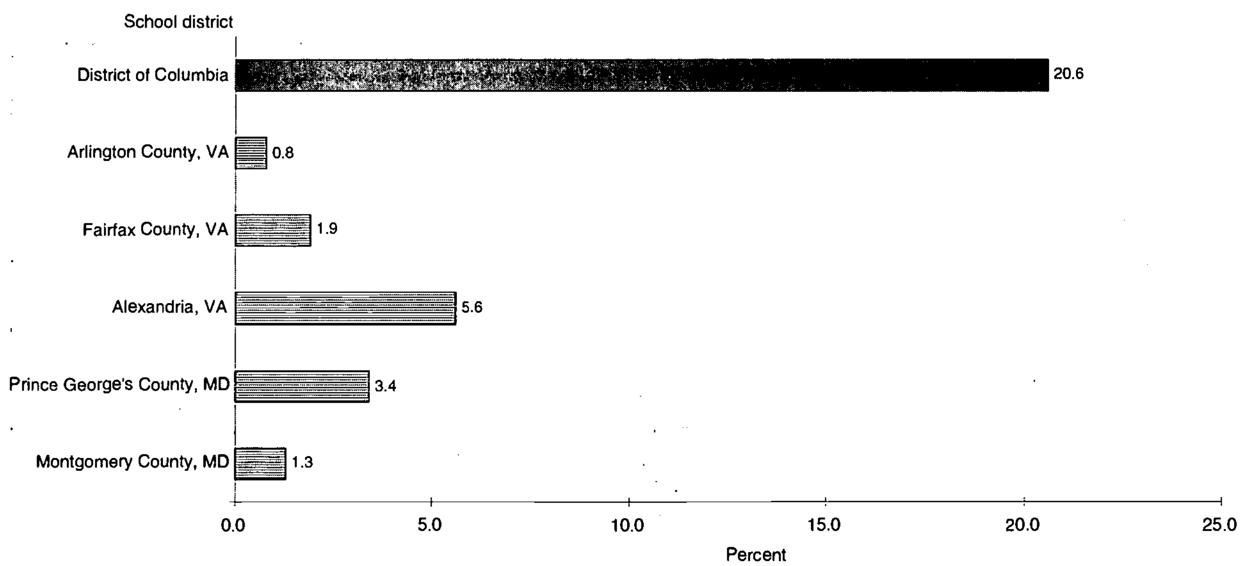
SOURCE: Data Volume for the National Education Goals Report (Volume Two: State Data).

Figure 8.—Percentage of schools with inadequate features



SOURCE: GAO Reports—School Facilities: Profiles of School Conditions by State; Condition of America's Schools.

Figure 9.—Discrepancy between public school enrollments as reported by school systems and by 1990 census



SOURCE: National Center for Education Statistics.

try would permit. Frequently, the congressional members attach all types of stipulations to the District's budget. In no area of the budget has this involvement been more than in education. Congress has mandated charter schools, English literacy requirements, public charter schools, a world class schools task force, per capita school funding, and the establishment of a commission on consensus reform. Public education budgeting in this environment is clearly a challenge.

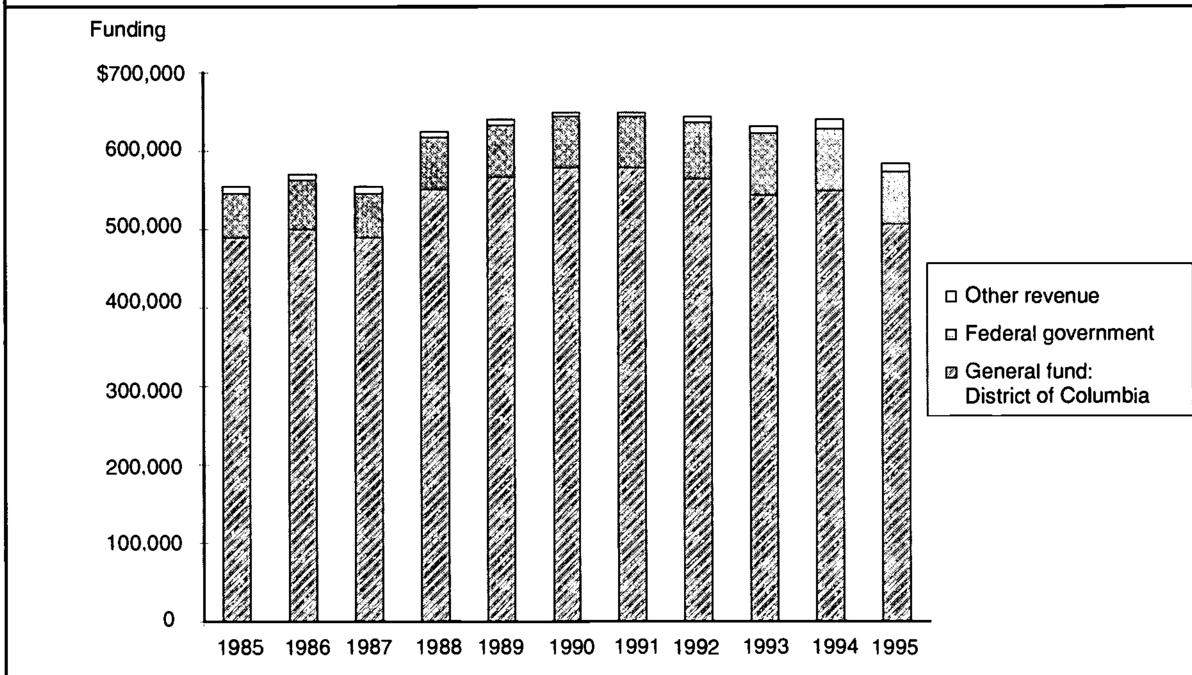
Although the former Superintendent and the former President of the Board of Education argued that more dollars are needed to support the District's schools, it remains to be determined whether or not additional operational funding is needed in light of an assessment of per capita spending. Funding for the District's public schools comes from local funds, and federal and private grants. Funding peaked in Fiscal Year (FY) 1994 and has declined slightly since then. Figure 10 shows funding over the past 10 years in constant 1996 dollars.

Comparisons of the District's school expenditures with other jurisdictions are difficult because of the many ways in which expenditures are reported. However, the District's average cost per student clearly exceeds the national average, and it is also substantially higher than many comparable urban school districts and neighboring districts. Figure 11 depicts these comparisons.

Even though spending per student in the District is high, much of that funding is not reaching the classroom. The 1992 report of the DC Committee on Public Education (COPE) found that DCPS spends "less than other districts on instruction and more on central administration and overhead." Four years later, DCPS expenditures toward classroom instruction continued to lag behind that of its neighbors.

Figure 12 compares DCPS' instructional spending to the Washington suburban jurisdictions of Fairfax and Montgomery counties.

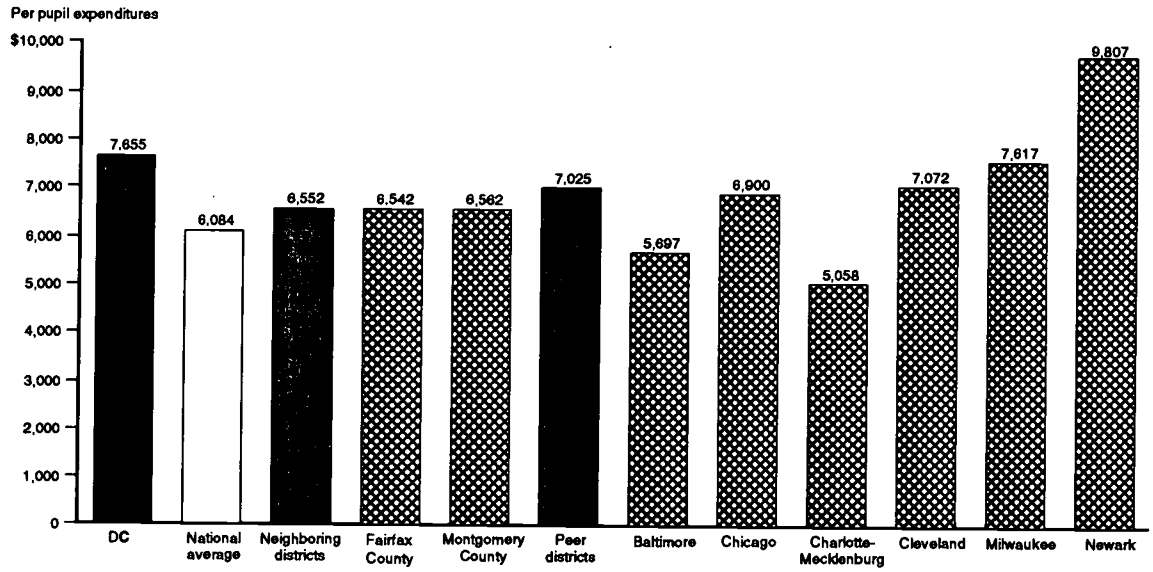
Figure 10.—District of Columbia public schools' funding from local, federal, and other funds (in constant 1996 dollars): FY 1985–95



NOTE: All funds are adjusted to 1996 dollars based on consumer price indexes for major expenditure classes.

SOURCE: District of Columbia Public Schools Comprehensive Annual Financial Report, 1995; DCPS Component Unit Financial Report, 1994.

Figure 11.—Average per-pupil expenditures: 1994–95*

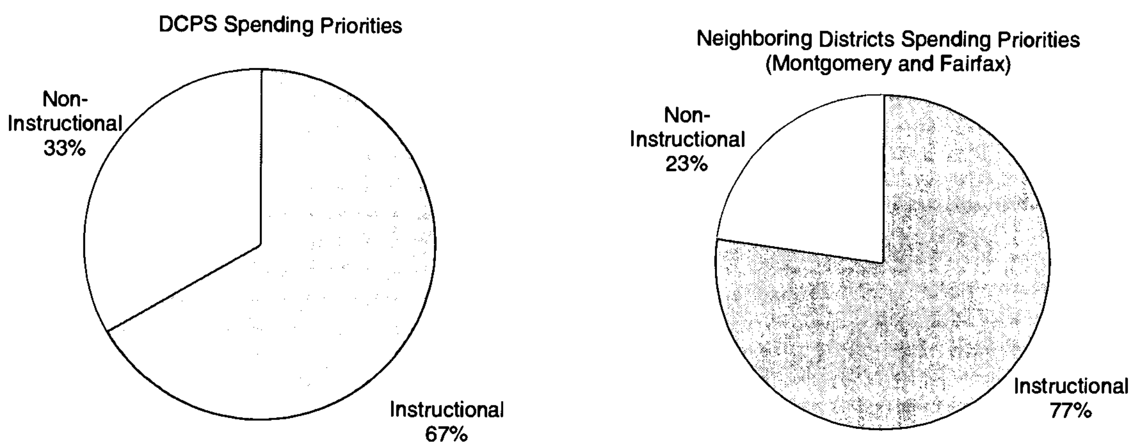


*These expenditures for the District do not include any cost of financing capital expenditures or the costs of the teacher pension program. It is unclear whether other jurisdictions' average cost per student includes such costs. However, these costs would add another $1,250 per student in the District.

NOTE: DC expenditures per pupil assume 80,450 students and total expenditures of $615.8 million and includes food services, continuing education, and other Federal education costs.

SOURCE: FY 1996 Budgets; National Center for Education Statistics.

Figure 12.—Comparison of District of Columbia public schools' (DCPS) instructional and non-instructional spending to Washington suburban jurisdictions



BEST COPY AVAILABLE

SOURCE: FY 1997 Budgets.

DCPS' large central administration takes away from instructional spending. DCPS' teacher to central administrator ratio highlights the size of its administration. Figure 13 depicts the fact that DCPS employs 16 teachers for every central administrator employed, compared with its peers who employ 42 teachers for every central administrator employed.

In 1996, DCPS allocated more toward its Office of the Superintendent than the Fairfax County, Montgomery County, and Baltimore City public school systems combined. DCPS exceeded its neighbors' average allocation by $4.55 million and its peers' by $3.38 million (figure 14).

DCPS spent more than twice as much on its Office of the Board of Education than peer and neighboring district average. In FY 1996, DCPS allocated over $1.4 million to the Office of the Board of Education—more than three times the average $454,000 of neighboring Fairfax and Montgomery counties (figure 15). DCPS even allocated $203,000 more than Chicago Public Schools (CPS), despite the fact that the CPS board was responsible for overseeing a district with 400,000 students—five times as many as DCPS.

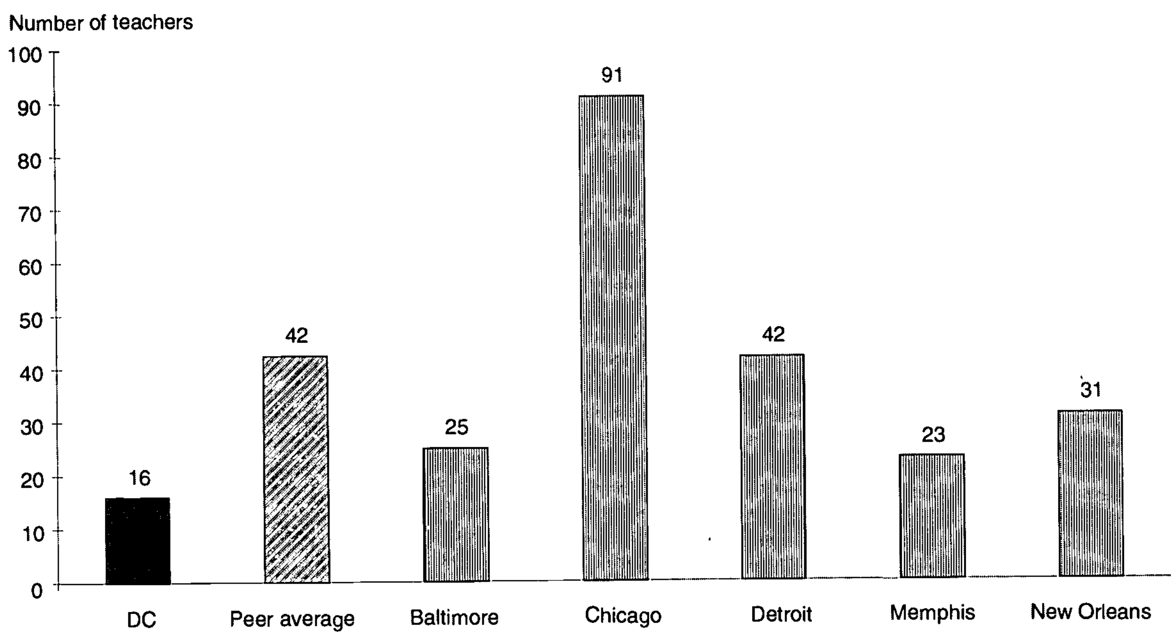
Authority Actions and Progress

With the deplorable conditions outlined in our report, the Authority believed that strong action needed to be taken. Therefore on November 15, 1996, the Authority, through use of powers granted by Congress, ordered the removal of the Superintendent and reduced the powers of the elected Board of Education. In their place, the Authority installed a retired Army Lieutenant General, Julius Becton, and established an Emergency Transitional Education Board of Trustees made up of experienced leaders with the ability to instill a vision and lead a large organization in the direction of that vision until June 2000.

The Authority also ordered the Trustees and the CEO to develop specific measures that will be used to gauge their success in meeting the following goals:

- enhancing the quality of education at all District schools;
- improving student participation, performance, and outcomes;

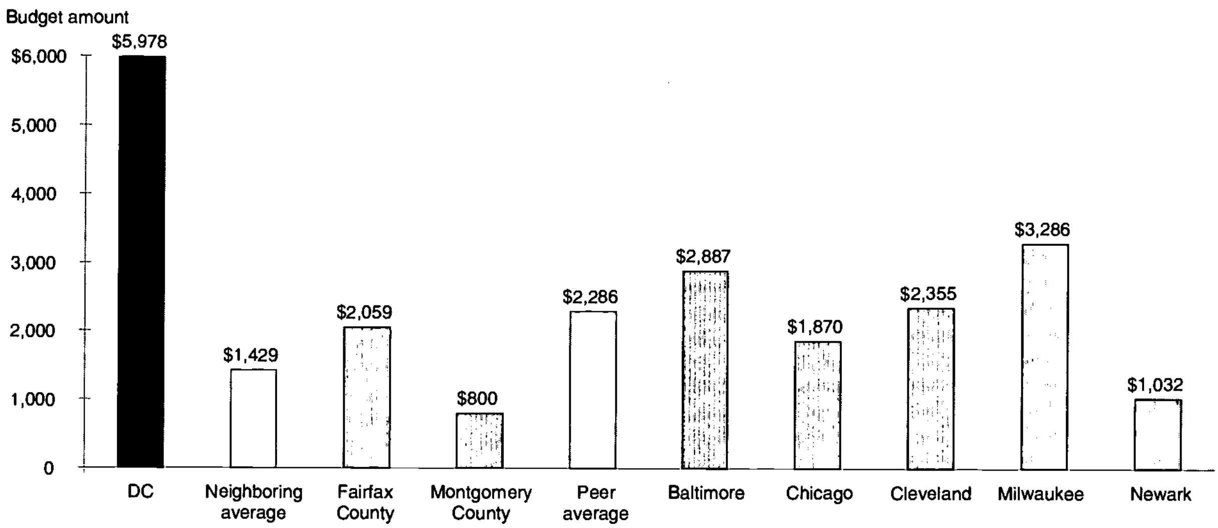
Figure 13.—Total number of teachers per central administrator: 1992–93



NOTE: Peer district data are not available for Boston, Cleveland, Milwaukee, and Newark.

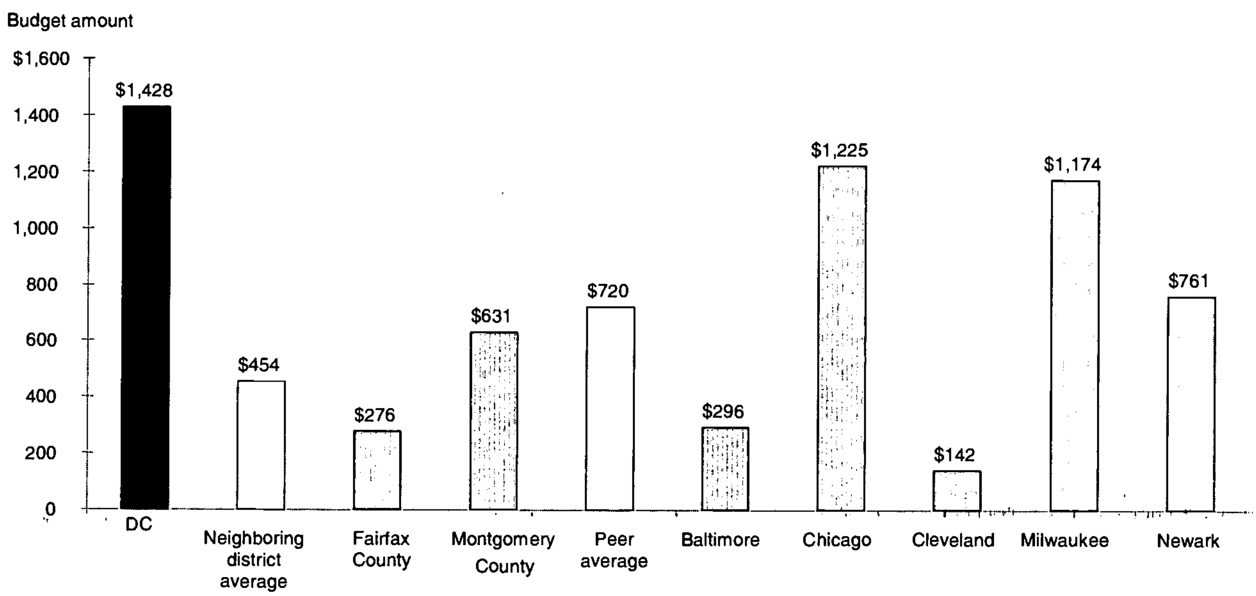
SOURCE: National Urban Education Goals: 1992–93 Indicators Report (Council of Great City Schools).

Figure 14.—Office of the Superintendent (or equivalent) adopted budget (in thousands of dollars): 1996



SOURCE: School District Budgets or Comprehensive Annual Financial Reports, FY 1996.

Figure 15.—Board of education (or equivalent) adopted budget (in thousands of dollars): 1996



SOURCE: School District Budgets or Comprehensive Annual Financial Reports, FY 1996.

- improving the quality, effectiveness, and efficiency of school system management;
- creating internal accountability and focus on customer satisfaction;
- improving the physical conditions of school properties; and
- reducing incidences of crime, especially offenses that threaten student and staff safety.

Within months of instituting the new structure, General Becton and the Board of Trustees found that as bad as the Authority's report painted the District's school system, the true situation was even worse:

- stacks of unpaid bills piled in boxes;
- personnel actions that had not been processed for months, even years;
- a special education program that was completely broken;
- numerous teachers who were not certified to teach;
- several personnel on the payroll who had died;
- no plans for facilities improvements; and
- a maintenance contractor who, according to the District's Inspector General, had overcharged the struggling school system by more than $6 million over the past 3 years.

To correct these problems, the new management team has:

- hired an academic officer, an operating officer, a human resource officer, and a chief financial officer;
- fired budgeting and finance personnel;
- updated all personnel actions and imposed a hiring freeze;
- developed a performance measurement system;
- developed a facilities plan and abated more than 1,600 fire code violations;
- closed 11 schools and replaced approximately 50 roofs;
- painted and spruced up over 40 schools (volunteers);
- increased security with new metal detectors and increasing training for security guards;
- established a teacher evaluation program;
- proposed new academic standards that end social promotions and require passing standardized reading and mathematics tests before progressing to certain key grades and a C average before graduating from high school;
- terminated a large school maintenance contract;
- awarded new contracts to provide better school breakfast and lunches; and
- provided closer coordination and interface with city personnel and financial management systems to record the school's employees and budget expenditures more accurately.

These are obviously just a beginning, and there have been some slips along the way such as the 3-week delay in school opening to finish roof replacement. General Becton and his chief operating officer have recently resigned. The chief academic officer replaced General Becton. But true reform, particularly in the academic area, is a long-term task and cannot be achieved overnight. The Authority will be closely monitoring the progress on the new management team to ensure that significant progress is made.

Higher Education in the District of Columbia

Public higher education in the District is principally provided by the University of the District of Columbia (UDC). UDC is unique in that it is the only urban land grant institution of higher education in the United States. UDC has an open enrollment policy and about 4,500 full-time equivalent students.

As a part of the Authority's efforts to revitalize public education in the District, we have undertaken a comprehensive review of higher education services in the District of Columbia, specifically UDC. UDC has had a difficult time adjusting to reductions in public funds, which currently are about 50 percent of the UDC budget. Our review of UDC found:

- UDC has never had a clear mission;
- the role of the faculty in the shared governance model has never been sufficiently detailed;
- operations and management functions have failed to provide efficient or effective services in support of the academic, community, or research mission of the institution;
- academic programs are not focused in meeting student or market needs; and
- UDC's allocation of resources is misdirected.

In summary, the Authority has recognized that if the District's quality of life and government is truly going to get better, we must fully invest our resources and efforts toward providing a better-than-adequate education for students in the District of Columbia school system. In making an investment in the world class education our students deserve, we will find what other systems have found—a better functioning school system will offer our youth hope for the future, a better opportunities for lifetime employment, and increased awareness of life's opportunities. In doing so we will find, just as other cities have found, that as we invest in education, the need for human service dollars will decrease as our students are better educated and are better prepared for a productive life.

Does Money Matter? An Empirical Study Introducing Resource Costs and Student Needs to Educational Production Function Analysis

Corrine Taylor

Department of Economics

University of Wisconsin-Madison

About the Author

Corrine Taylor is a doctoral candidate in economics at the University of Wisconsin-Madison. She will join the faculty at Wellesley College as an assistant professor of economics in the fall of 1998. Ms. Taylor teaches microeconomics, state and local public finance, and statistics. Her research focuses on the economics of education, in particular, elementary and secondary school finance.

As a graduate student, Ms. Taylor served as a researcher on the Teacher Compensation Project of the Consortium for Policy Research in Education. She was awarded a Spencer Dissertation Fellowship for Research Related to Education (1996–97). Ms. Taylor is a member of the American Economic Association and the American Educational Research Association.

Does Money Matter? An Empirical Study Introducing Resource Costs and Student Needs to Educational Production Function Analysis

Corrine Taylor

Department of Economics

University of Wisconsin-Madison

Introduction

Do expenditures on school resources have a positive effect on student outcomes? This question is important to many audiences: parents of school-aged children; citizens concerned about the effectiveness of their tax dollars; educators trying to improve student outcomes; and state policymakers charged with developing fair school finance formulas. Despite thirty years of research by economists, sociologists, and educational researchers, beginning with the Coleman Report (1966), this question still has no definitive answer.

Most economic analyses take an "educational production function" approach. These studies use econometric techniques to relate educational outcomes (e.g., students' academic achievement) to school inputs while controlling for other contributions such as those of the students themselves, their families, peers, and communities. Within this broad framework, educational production function studies exhibit a wide range of empirical approaches.[1] They vary in their choice and measurement of educational outcomes, explanatory variables of interest, and control variables. They also differ in their geographical scope and their unit of analysis.

Findings from these studies are as mixed as their empirical approaches are varied. Some studies estimate large, positive effects of school inputs on student outcomes; others find little or no effect; still others conclude that additional school resources are inversely related to student outcomes. The most well-known result of this vast literature is Hanushek's (1986, 1989) conclusion of "no strong or systematic relationship between school expenditures and student performance." Hanushek's finding is based on his syntheses of more than thirty separate educational production function studies.[2] A more recent synthesis by Hedges, Laine, and Greenwald (1994) challenges the validity of the analytical method of "vote counting," employed by Hanushek. Using the same primary studies as Hanushek's 1989 analysis, but a more sophisticated synthesis methodology known as "meta-analysis," Hedges, Laine, and Greenwald reach the opposite

[1] The many approaches of educational production function studies are reviewed by Hanushek (1979, 1986), Cohne and Geske (1990), and Monk (1992).

[2] Hanushek's famous 1986 analysis in the *Journal of Economic Literature* includes 147 regressions from 33 separate education production function studies. His updated 1989 study in *Educational Researcher* includes 187 regressions from 38 primary studies. He reports the exact same conclusion in the two synthesis studies.

conclusion.[3] They find a statistically significant and economically substantial, positive relationship between school inputs and student outcomes.

The relevance of the findings from these syntheses depends not only on the quality of their methodological approaches but, more importantly, on the quality of the primary research studies. In reviewing the primary studies considered in these syntheses, I find that none of the primary studies adequately accounts for across-district variations both in the resource costs of educational services (notably teacher compensation) and in the proportion of students with special needs, who require additional, more costly services.

These variations in resource costs and student needs are significant. The power of school districts to purchase a standard "market basket" of educational resources varies by twenty to forty percent within states and as much as forty percent across states (Chambers, 1981; McMahon, 1995). Student needs vary widely across districts as well, with the proportion of special-needs students approaching fifty percent in some large urban school districts (Odden & Picus, 1992). I expect that a stronger relationship between student achievement and school expenditures will emerge after accounting for these resource-cost and student-need differentials.

To test this hypothesis, I use a unique data set merged from three high quality, national data sources: the National Education Longitudinal Study of 1988, the Common Core of Data, and a district-level teacher cost index.[4] I specify and estimate a value-added student achievement model for which my explanatory variable of interest is per-pupil expenditures. I find that the estimated effects of per-pupil expenditures on high school students' academic achievement are consistently positive and statistically significant. However, these effects do not increase appreciably when the measure of expenditures is corrected to account for resource-cost differentials or when differences in the proportions of special-needs students are taken into account.

The remainder of this paper is organized as follows: I first present my conceptual model and describe the data sources, sample, and variables used in my empirical analysis. Next, I explain how I conducted my estimations and present and discuss the results. Lastly, I summarize my findings and presents suggestions for future research.

Conceptual Framework

Educational Production Function Studies

My conceptual model is the basic value-added, reduced-form specification of the educational production function presented in Hanushek's (1979, 1986) reviews. The educational outcome of interest is academic achievement. An individual student's achievement at time *t* (A_t), is modeled as a function of the student's prior achievement (A_{t^*}), other student characteristics and effort *(I)*, and the influences of the student's family *(F)*, peers *(P)*, school *(S)*, and community *(C)* during the period between t^* and *t*. That is,

$$A_t = f(A_{t^*}, I_{t-t^*}, F_{t-t^*}, P_{t-t^*}, S_{t-t^*}, C_{t-t^*}).$$

The effects of the school inputs on achievement are of primary interest in educational production function analyses. The types of school inputs considered in these analyses depend on the policy questions being addressed. Studies that focus on *how* schools allocate their funds typically consider teacher/pupil ratios, and teachers' education levels and years of experience as the school inputs. My policy interests involve the equity of school finance

[3] Hanushek's analytical method of "vote counting" examines only the sign and level of statistical significance of the estimated effects of the seven different school inputs on student performance. He gives one "vote" to each estimated effect with a positive sign. Whether he considers only those effects that are statistically significant or he ignores statistical significance, Hanushek concludes that the proportion of positive effects is too small to indicate a strong relationship between school inputs and student performance.

Hedges, Laine, and Greenwald's "meta-analysis" considers not only the signs but also the magnitudes of the estimated effects of school inputs on student outcomes. Additionally, their more sophisticated methodology accounts for dependence among regressions estimated within the same study using slightly different empirical specifications and among regressions in different studies that used the same data sources.

[4] The teacher cost index was developed by Jay Chambers of the American Institutes for Research, and like the other data sources, was released by the U.S. Department of Education's National Center for Education Statistics.

formulas; hence, I consider schools' fiscal resources as the school input of interest.

The efforts of states to provide more equitable educational opportunities and student outcomes by reducing across-district disparities in schools' fiscal resources inspired my two primary research questions: 1) Is there a positive, systematic relationship between student performance and schools' fiscal resources? and 2) How does the strength of that relationship depend on the precise measure of fiscal resources? Specifically, is the relationship between student achievement and per-pupil expenditures (PPEs) stronger when the PPE measure reflects the costs of educational services and the population of special-needs students? If this is the case, then states would be more likely to achieve their student equity objectives by attempting to equalize not nominal per-pupil expenditures, but rather per-pupil expenditures adjusted for costs and student needs.

Variations in Costs

One problem in educational production function studies that link schools' fiscal resources to student outcomes is that the costs of equivalent educational services vary widely across districts. Researchers estimate that these costs vary by twenty to forty percent within states and up to forty percent across states (Chambers, 1981; McMahon, 1995). In studies that ignore such differential resource costs, disparate outcomes for districts with identical expenditure levels seemingly lend support to the notion that money does not matter. In fact, higher student achievement should be *expected* in low cost districts which, for the same nominal expenditure level, can purchase more or higher quality real resources than high cost districts can afford, all else being equal.

. . . disparate outcomes for districts with identical expenditure levels seemingly lend support to the notion that money does not matter.

One recent production function study does attempt to account for variations in education costs by location. William Sander (1993) adjusts his expenditure and income variables by a cost-of-living index developed by Walter McMahon (1988), and finds that teacher-related spending is positively related to ACT scores in Illinois. Although Sander's study represents an improvement over the prior literature, cost-of-living adjustments do not adequately account for educational price differentials.

The cost of living is but one factor affecting the attractiveness of a school district as a place to live and work. Other characteristics—including the size of the school district, the types of students served, the crime rate, the level of pollution, the climate, access to medical facilities, availability of recreational opportunities, and consumption opportunities—also affect the attractiveness of districts, and ultimately affect the salaries that are required to attract and retain individuals with specific professional characteristics (Chambers, 1981). A cost-of-living adjustment fails to adequately account for variations in salaries of school personnel due to differences in job and regional characteristics. Since personnel costs comprise at least 80 percent of school expenditures and since variations in personnel costs dominate the pattern of cost differences across districts it is important to account for them (Chambers and Fowler, 1995).[5]

While a number of approaches have been taken in efforts to develop an index for personnel costs (see Chambers, 1981, pp. 45–52), Chambers argues that the most appealing approach is based on the hedonic wage model. The theoretical framework, established by Lucas (1972), maintains that through a simultaneous process of matching the attributes of individual employees and the working conditions offered by employers, differential wages are determined. In its application to the market for school personnel, hedonic wage theory recognizes that differences in the characteristics of school districts require different salary levels to attract the types of

[5] Transportation and energy costs vary widely across districts as well, but account for a much smaller portion of schools' expenditures.

personnel needed to provide a given level and quality of educational services across districts.

The personnel index indicates the relative cost of employing workers with similar skills and jobs in different environments. The different environments are characterized by district and regional factors that are beyond the control of local school decision-makers (Chambers, 1981, p. 63).[6] The types of district and regional factors considered reflect the overall quality of the environment within which the individual works and lives as well as the condition of the labor market in which prevailing wages and employment levels are determined. Thus, a personnel cost index accounts for variations in district and regional characteristics, controlling for personal and job assignment characteristics.

Adjusting expenditures by a personnel cost index allows for more meaningful comparisons of PPE levels across districts that face different resource costs. We would expect that cost-adjusted expenditures are better at capturing the quantity and quality of the educational services purchased, and that such "real" measures should be more closely related to student performance than the typically considered "nominal" measures.

Just as nominal expenditure levels make for poor comparisons . . . with different resource costs, simple PPEs make for poor comparisons across districts with different proportions of special-needs students.

Variations in Student Needs

In educational production function analyses for which the observations are individual students, the ideal measure of a school's fiscal inputs would be the dollars (adjusted to reflect resource costs) spent on each individual student. However, school expenditures are most accurately measured (and often only available) at the district level and are difficult to accurately allocate to schools, classrooms, or individual students. Hence, whether the unit of analysis is individual students, schools, or districts, most analyses that focus on fiscal resources simply use district-level PPEs—total district expenditures divided by the total number of students in the district—as the measure for school inputs. Just as nominal expenditure levels make for poor comparisons across districts with different resource costs, simple PPEs make for poor comparisons across districts with different proportions of special-needs students.

The distribution of special-needs students—including special education, compensatory education, and limited English proficiency (LEP) students—is not uniform across school districts. The incidence of students with physical and mental handicaps varies widely across states and districts. Large, urban districts and small, rural districts tend to have higher proportions of students for whom English is not the primary language. Urban and rural areas also tend to serve a higher proportion of students living in poverty (Odden and Picus, 1992). The costs of providing services to these special-needs students vary depending on such factors as the number and types of students with special needs, the size of the school, and the kinds of services provided. In general, though, studies estimate that special education programs are about 2.3 times as costly as regular programs (Kakalik et al., 1981; Moore, Strang, Schwartz, and Braddock, 1988; Chaikind, Danielson, and Brauen, 1993), and compensatory and LEP programs are at least 20 percent more costly (Odden and Picus, 1992; Parrish, Matsumoto, and Fowler, 1995).

A variety of federal and state aid programs are designed to help districts offset the additional costs of providing extra services for special-needs students. Under Chapter 1 of the Elementary and Secondary Education Act (ESEA), the federal and state governments provide extra funds to districts for compensatory education. Title VII of the ESEA makes funds available for bilingual education programs. The federal Education for All Handicapped Children Act mandates and helps fund special edu-

6 It is essential that adjustments for differential costs of education be based only on factors which are *beyond* the control of district decision-makers, so that inefficient spending practices are not encouraged.

cation programs. Analyses of expenditures that include these additional funds should also reflect the size of the special needs population for whom these funds are provided.

Because the distribution of special-needs students varies widely among school districts, simple comparisons of PPEs across districts fail to reflect differences in school resources available for the average student. Districts with smaller proportions of the more costly special-needs students, in effect, have more money to spend on the average student than do schools with higher proportions of these special-needs students, ceteris paribus. Hence, in educational production function studies relating school expenditures to student achievement, control variables for the proportion of special-needs students in each district need to be included in the regressions.

Hypothesis

Figures 1–3, show how I expect these variations in resource costs and students' needs to affect the relationship between student achievement and school expenditures. Figure 1 is a stylized representation of Hanushek's conclusion that there is no relationship between student achievement and school expenditures. Figure 2 illustrates my hypothesis. I expect that districts with higher levels of student achievement and lower nominal expenditures (upper left portion of graph) face lower costs of education and have relatively fewer special-needs students. Under these conditions, the adjusted measure of PPEs would be higher than the nominal measure. (The arrows represent the change in PPE measure from nominal to adjusted.) Similarly, I expect that districts with lower levels of student achievement and higher nominal expenditures (lower right portion of graph) face higher costs of education and serve a higher proportion of special-needs students. For these districts, the adjusted measure of PPEs would be lower than the nominal measure. If my expectations are correct, then a (larger) positive relationship between student achievement and school expenditures should emerge as the measure of expenditures is adjusted to account for these differences in resource costs and student needs (see figure 3).

Districts with smaller proportions of . . . special-needs students, in effect, have more money to spend on the average student than do schools with higher proportions of these special-needs students . . .

Empirical Model

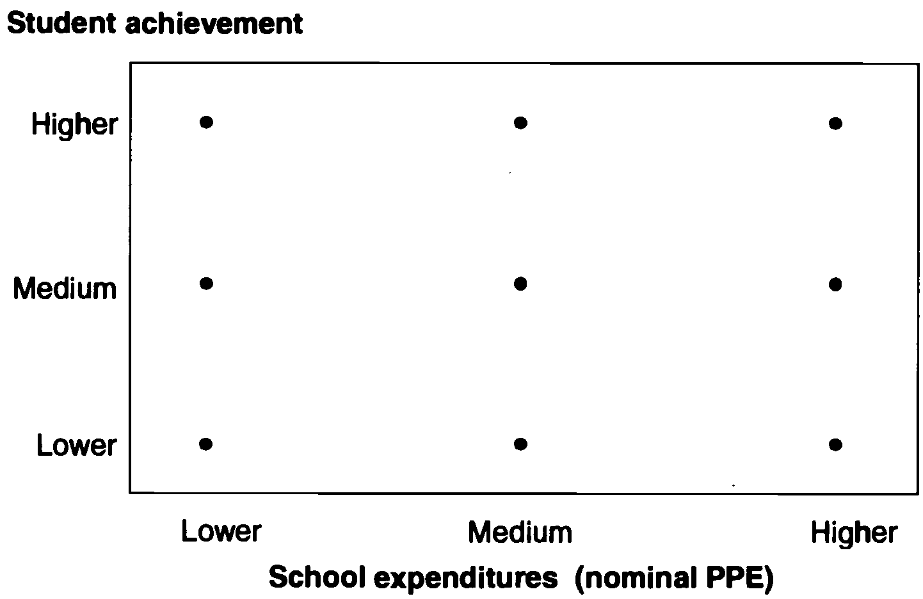
Data Sources

This study uses data merged from two large data sets and a smaller data file, each released by the National Center for Education Statistics. The first source is the restricted-use version of the National Education Longitudinal Study of 1988 (NELS), a general-purpose panel study that surveyed and tested eighth graders from about 1,000 public and private middle schools in the spring of 1988 and followed these students through high school. The first three waves of NELS include scores on cognitive tests administered to students in 1988, 1990, and 1992 as well as information from questionnaires administered to students, their parents, teachers and school administrators over the same time period (Ingels et al., 1994).

The second source is the Common Core of Data (CCD), an annual, comprehensive database containing descriptive data on all public elementary and secondary schools and school districts in the United States. The CCD also contains enhanced financial data at the district level for fiscal years 1990, 1991, and 1992. Additionally, the CCD contains demographic indicators derived from special tabulations for school districts from the 1990 Census (National Center for Education Statistics, 1995).

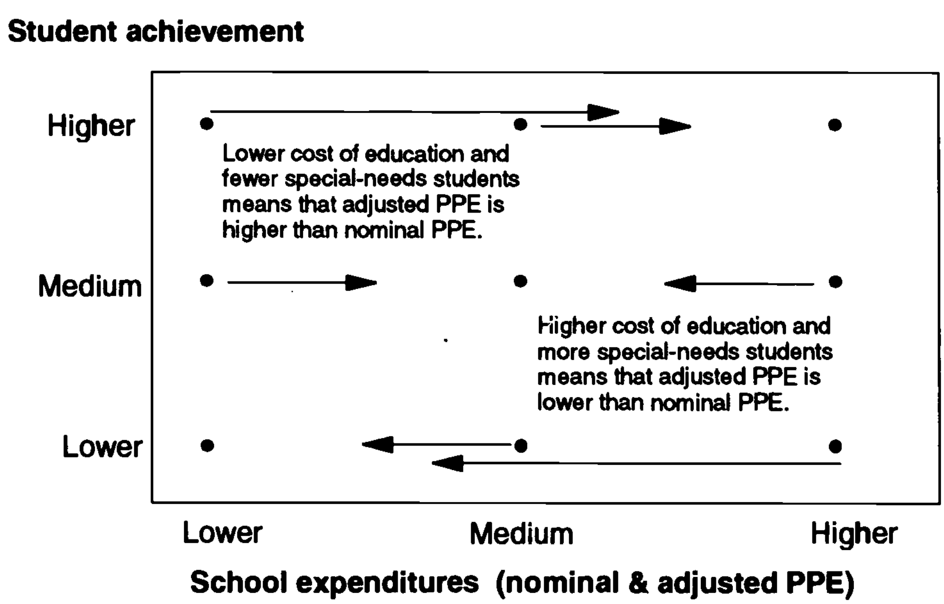
The third, smaller data source is a national, district-level teacher cost index (TCI) developed by Jay Chambers of the American Institutes for Research. Chambers' TCI reflects across-district variations in non-discretionary resource costs of teacher services. Based on a hedonic wage model, the TCI was created using survey data from over 40,000 public school teachers who participated in the NCES's Schools and Staffing Survey for school year 1990–1991. Chambers' TCI is the only nationwide, dis-

Figure 1.—With the traditional measure of per-pupil expenditure (PPE), no relationship between school expenditures and student achievement is evident



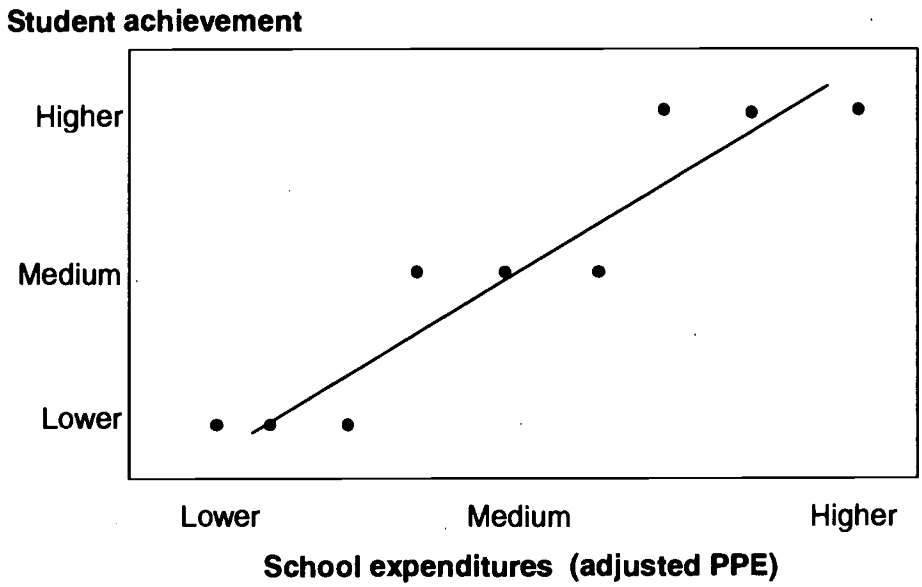
SOURCE: Author's illustration.

Figure 2.—Adjusting expenditures to account for the cost of education and special-needs students may bring a new picture into focus



SOURCE: Author's illustration.

Figure 3.—A positive relationship between school expenditures, measured by adjusted per-pupil expenditure (PPE), and student achievement is expected to emerge



SOURCE: Author's illustration.

trict-level index available that takes into account both the factors that underlie differences in the cost of living and variations in other teacher and school attributes that are within local control (Chambers and Fowler, 1995). Appendix A describes the construction of the TCI.

Sample

My sample is drawn from those students who participated in all the first three waves of the NELS panel study (16,489 students). I consider only students attending public schools (11,598) because they are the only ones to whom I can assign reliable, comparable expenditure data from the CCD.[7] I further refine my sample to include only students who never dropped out of school (11,503) and who attended the same high school in both 1990 and 1992 (11,167).

These restrictions are imposed because I want to consider only those students who are consistently associated with school resources at particular schools. The disadvantage is that these students constitute a more stable student body than is reflected in the total student population. To the extent that dropout rates, transfer rates, or participation in all three waves of the NELS survey are systematically related to PPE levels, my findings are not generalizable to the entire student population; rather, they must be qualified to apply to this more stable group of students.

I further eliminate observations with missing data in three critical areas: test scores, special-needs students, and TCI values. I lose a substantial number of observations by considering only students with complete test score data in both 1988 and 1992; this restriction leaves 7,854 students.[8] Eliminating observations lacking CCD data on the number of special-needs students and observations with miss-

[7] NELS oversampled students in private schools; hence, the large proportion of students who are eliminated given that I consider only students who attend public schools.

[8] In this paper I do not tackle the potential "pretest to post-test selection problem" discussed by Becker and Walstad, 1990.

ing TCI values leaves a sample size of 6,990. Missing values for some control variables reduce the number of observations used in the regression computations to 5,955.[9]

Variables

The dependent variable in my regression equations is the student's 1992 (senior year for most of the students) score on the NELS mathematics test. The specific measure I use for mathematics achievement is the item response theory (IRT) theta score, which is standardized to a mean of 50 and a standard deviation of 10. To eliminate floor and ceiling effects, three forms of the mathematics tests were administered to the students in 1992, depending on their prior achievement. Students who performed in the highest quartile on the 1990 test were given the most difficult version of the 1992 exam; those in the lowest quartile in 1990 received the easiest version of the 1992 exam; and the rest of the students received the test of medium difficulty in 1992. Item response theory was used to calculate scores that could be compared across test forms that differed across the years and across the students in a given year. The theta score, which is standardized across the three waves of testing is the best score to use when assessing gains in cognitive skills. (See Ingels et al., 1994 for more information about NELS testing and IRT scoring.)

. . . is the estimated effect of PPEs on student achievement strengthened by accounting for across-district variations in resource costs and student needs?

The independent variables include controls for achievement in eighth grade, in order to analyze the *gain* in cognitive outcomes during the high school years. I include both the 1988 mathematics IRT theta score and the average 1988 IRT theta score on the other three NELS tests—science, reading, and social studies—as control variables.[10] I use the average of the other test scores as an additional control to reduce bias from unmeasured pre-existing differences among students (see Gamoran, 1996; Gamoran and Mare, 1989; and Jencks, 1985). I expect to find strong, positive relationships between these measures of prior achievement and the measure of achievement on the mathematics test in 1992.

Other control variables included in my empirical analysis capture student and family characteristics, the student's interest and effort in mathematics and in school, and characteristics of the student's peers, school, and community. Descriptive statistics for these control variables are reported in table 1.[11] Definitions and sources for all the variables are provided in appendix B.

Methodological Approach

Recall that two primary questions are addressed in this study. First, do these high quality, nationwide data reveal a positive relationship between student achievement and PPEs? Second, is the estimated effect of PPEs on student achievement strengthened by accounting for across-district variations in resource costs and student needs? Addressing the first question is a straightforward matter of examining the statistical significance and substantive magnitude of the coefficient estimates on the PPE variables. Addressing the second question is more involved.

[9] Other fields with missing data include: the percentage of students in the district living in single-parent homes; the percentage of students in the district in minority families; historical dropout rates in the high school; and enrollment in the twelfth grade. In future studies, I intend to impute values for missing data in these fields.

[10] I use the 1992 math score as the dependent variable and include the 1988 math score as a control variable, rather than using the gain in score as the dependent variable, because the former specification is less restrictive. In particular, the gain score specification implicitly assumes that the coefficient on the 1988 math score should be one. Typically, the coefficient estimate on prior achievement in the same subject is in the range of 0.70 to 0.80.

[11] The means and standard deviations are weighted to account for the oversampling of certain populations in the NELS three-wave panel. The weight used in computing these descriptive statistics is the relative weight, F2PNLWTi / mean (F2PNLWT).

Table 1.—Descriptive statistics*

Variable	Mean	Standard deviation	Minimum	Maximum
Dependent variable				
Math score, 1992	54.19	10.03	27.07	80.6
Explanatory variables				
Prior achievement				
Math score, 1988	45.71	8.36	24.89	67.
Average of other scores, 1988	46.22	7.54	25.89	66.2
Student and family characteristics				
Minority	0.23	0.42	0.00	1.0
Female	0.50	0.50	0.00	
Single-parent family	0.14	0.35	0.00	1.0
Socioeconomic status	0.01	0.76	-2.43	1.9
Student interest and effort				
Interest and effort in math	2.56	1.34	0.00	4.00
Time spent on homework	6.64	3.34	0.00	16.00
Class attendance	3.29	1.22	0.00	5.00
Student's view of school environment				
Perceives disruptive environment	0.85	1.01	0.00	4.00
Experiences disruptive environment	0.96	1.37	0.00	7.00
Peers' characteristics				
Peers from single-parent homes	2.61	0.78	1	5
Percent minority students	23.67	29.51	0	100
Peers' absenteeism	0.47	0.50	0	1
Peers' dropout rates	2.02	1.53	0	6
Special-needs students				
Percent special education	9.65	4.17	0	23.16
Percent with limited English proficiency	1.95	3.15	0	25.20
Percent below poverty level	16.99	11.27	0.40	66.20
Community characteristics				
Percent adults w/ at least some college	45.86	15.02	10.80	92.00
Median income for households w/ kids	36,907	13,083	11,337	114,544
School characteristics				
Size				
Twelfth grade enrollment	275	168	12	1110
Problems in school				
Composite of minor to serious problems	8.09	4.44	0	15
Type				
Comprehensive school	0.91	0.29	0	1
Magnet school	0.10	0.29	0	1
Public school of choice	0.34	0.47	0	
Year-round school	0.04	0.18	0	1
Vocational-technical school	0.09	0.28	0	1
Region				
Midwest	0.33	0.47	0	1
Northeast	0.13	0.33	0	
South	0.34	0.47	0	1
West	0.20	0.41	0	1
Urbanicity				
Suburban	0.45	0.50	0	
Urban	0.20	0.40	0	
Rural	0.35	0.48	0	1

* Weighted to reflect population means.

SOURCE: Author's calculations using the National Education Longitudinal Study (NELS) and Common Core of Data (CCD) data.

85

Coefficient Comparisons Across Regressions

To address the second question I run four main regressions then compare the coefficient estimates on the PPE variables across these regressions. The four regressions differ only in their measure of PPE and in their controls for special-needs students. I consider two measures of PPE: nominal and cost-adjusted. "Nominal PPE" is calculated by simply dividing the district's expenditures by the number of pupils in the district. "Cost-adjusted PPE" divides the nominal PPE value by the teacher cost index (TCI) times 100. (The TCI is centered at 100 in the population rather than at one; hence the need to multiply by 100.) Additionally, I consider two alternative specifications of the model: in the first specification I do not control for the proportion of special-needs students; in the second specification, I do. In the second specification I include separate control variables indicating the proportion of students in each of the following special needs categories: special education, limited English proficiency, and compensatory education. The combination of the two alternative PPE measures and the two alternative specifications produce the four distinct regressions.

. . . student achievement on the 1992 NELS mathematics test is positively related to per-pupil expenditures.

To examine the robustness of the results, I consider three alternative categories of expenditures. The three expenditure categories are: 1) total district expenditures; 2) core current expenditures; and 3) expenditures on instructional salaries. The first category encompasses all current operation and capital outlay expenditures. The second includes just three key types of current operation expenditures: instructional expenditures (salaries and benefits for teachers and aides, contracted services, and supplies), pupil support services, and instructional staff support. The third category is the narrowest of all: only instruction-related salaries for teachers and aides are considered. Table 2 reports descriptive statistics for the nominal and the cost-adjusted PPE measures in each of these three expenditure categories.

To meaningfully compare the coefficient estimates across regressions, the nominal and cost-adjusted PPE measures used in the regressions need to be on a common scale. Therefore, I create a new variable, called "comparable cost-adjusted PPE," by multiplying each observation of the "cost-adjusted PPE" by a constant factor. The factor equals the ratio of the mean nominal PPE to the mean cost-adjusted PPE. The factor differs slightly across the three expenditure categories, but in all cases is approximately 0.987. (Descriptive statistics for the "comparable cost-adjusted PPE" measure are also presented in table 2. Note that the means for the nominal and comparable cost-adjusted PPE variables are identical by design.) It is the "nominal PPE" and the "comparable cost-adjusted PPE" variables that are included in the regressions, thus allowing for meaningful across-regression comparisons of the coefficient estimates on the PPE variables within each expenditure category.

Within each expenditure category, I expect to find that the magnitude of the coefficient on the PPE measures increases: 1) as the measure changes from "nominal PPE" to "comparable cost-adjusted PPE"; 2) when the regressions control for special-needs students; and 3) as both cost and student needs are taken into account (i.e., we move from nominal PPE and no controls to cost-adjusted PPE and special-needs controls).

Estimation Results

The results confirm that student achievement on the 1992 NELS mathematics test is positively related to per-pupil expenditures. This result holds for all three expenditure categories, whether the PPE measure is nominal or cost-adjusted, and whether or not control variables for special-needs students are included in the regression. Table 3 summarizes the estimated effects of the various expenditure measures on achievement for both model specifications. The coefficient estimate is consistently positive and statistically different from zero, though it is substan-

Table 2.—Descriptive statistics, alternative measures of expenditures

	Mean	Standard deviation	Minimum	Maximum
Measure 1: Total district expenditures				
Nominal per-pupil expenditure (PPE)	5,577	1,871	2,895	14,918
Cost-adjusted PPE	5,655	1,621	2,957	15,346
Comparable cost-adjusted PPE	5,577	1,599	2,912	15,134
(Comparability factor: 0.9862)				
Measure 2: Core current expenditures				
Nominal PPE	3,394	1,176	1,819	9,277
Cost-adjusted PPE	3,434	953	1,746	8,496
Comparable cost-adjusted PPE	3,394	942	1,726	8,398
(Comparability factor: 0.9884)				
Measure 3: Instructional salaries				
Nominal PPE	2,245	724	1,086	5,934
Cost-adjusted PPE	2,274	580	1,014	5,500
Comparable cost-adjusted PPE	2,245	573	1,001	5,428
(Comparability factor: 0.9870)				

SOURCE: Author's calculations using the National Education Longitudinal Study (NELS) and Common Core of Data (CCD) data.

tively small.[12] For example, the coefficient on nominal core PPE in the regression that controls for special-needs students is 0.381. This coefficient means that for an additional $1,000 in per-pupil expenditures, the math score is expected to increase by 0.381 points over the four years of high school. Given that typical gain in math score is about 8.5 points, the extra $1,000 per pupil raises test scores by only 4 percent of what is already expected.

The results lend mild support for the hypothesis that accounting for differential resource costs and student needs would reveal a stronger positive relationship between student achievement and school expenditures. In table 3, I use a solid arrow to indicate changes in the magnitude of the coefficient that are in the expected direction; broken arrows indicate changes in the unexpected direction. While the direction of change is as expected in 13 of 15 cases, the magnitude of the change is minuscule compared to the standard errors. Indeed, the confidence intervals for the coefficients within each of the three expenditure categories almost entirely overlap.

Although not of primary interest in this study, it is interesting to examine the effects of the other explanatory variables included in the model. These other effects may shed light on the weak effects of the fiscal resources. Table 4 presents all the estimated effects from the regressions that use (comparable) cost-adjusted core expenditures per pupil as the explanatory variable of interest. Performance on the 1992 mathematics test is positively and statistically significantly related to prior achievement in both math and other subjects. Higher math achievement

[12] Because the NELS observations do not come from a random sample, the reported OLS estimates of the standard errors may be understated. Using a hierarchical linear modeling technique to account for the clustering of students within schools, I found that the HLM standard errors were virtually identical to the OLS standard errors. This result is not surprising, since there were only ten students, on average, in each school in 1992, and the magnitude of the bias for the standard errors increases with the average group size. (See Moulton, 1990, p. 335.) Other departures from random sampling (e.g., oversampling minorities) may also require the imposition of higher standards in judging statistical significance. (See Ingels, et al., 1994, pp. 42–53.) The root design effect for the full panel, when using the mathematics IRT score as the dependent variable, is 2.273. Multiplying the OLS standard errors by 2.273 will give a conservative standard error to use in judging statistical significance. Even imposing this most stringent standard for the standard errors, all the coefficients of the expenditure variables are statistically greater than zero at the 5 percent level of significance.

Table 3.—Comparison of effects of expenditures on 1992 math score

Coefficients from regressions differing in per-pupil expenditure (PPE) measure and special-needs controls	Model 1 No special needs controls	Model 2 With special needs controls
Measure 1: Total district expenditures		
Nominal PPE	0.221	0.214
	(.051)	(.051)
Comparable cost-adjusted PPE	0.226	0.231
	(.049)	(.050)
Measure 2: Core current expenditures		
Nominal PPE	0.374	0.381
	(.093)	(.094)
Comparable cost-adjusted PPE	0.406	0.444
	(.096)	(.097)
Measure 3: Instructional salaries		
Nominal PPE	0.633	0.630
	(.159)	(.161)
Comparable cost-adjusted PPE	0.649	0.700
	(.163)	(.166)

NOTE: Standard errors are in parentheses. Solid arrows indicate that the coefficient change is in the predicted direction. Broken arrows indicate that the coefficient change is opposite the predicted direction.

SOURCE: Author's calculations using the National Education Longitudinal Study (NELS) and Common Core of Data (CCD) data.

is also positively and significantly related to higher socioeconomic status. Females' performance on the math tests is worse than males', and minorities' performance is worse than non-minorities'. Students from single-parent homes perform worse than those from two-parent households, but not significantly so. All three separate measures of student effort are positive and statistically significant. Students who experience multiple disruptions at school perform worse than those in less disruptive learning environments. The signs on most of the other non-expenditure-related explanatory variables are generally as expected. The most notable unexpected result is the negative coefficient on the median income for households with children. The effects of the PPE variable were highly sensitive to the inclusion or exclusion of this income variable, even though the correlation coefficient is only about 0.5. The positive coefficient on the percent of LEP students in the regressions that used control variables indicates that limited English proficiency may not be a substantial handicap on math tests. Indeed, international studies consistently rank U.S. school children among the lowest in math performance. Perhaps in schools with higher proportions of LEP students, the students are able to draw more from their prior mathematics knowledge. In future analyses, I will consider performance in the other NELS subjects as well. I expect, for example, that the coefficient on LEP students will be negative on the reading test.

Conclusions and Directions for Future Research

This paper contributes to the understanding of the effects of school expenditures on student achievement by drawing on three nationwide data sets which are merged to create a rich sample for the empirical analysis. I expected to find (1) that the relationship between student achievement and nominal expenditures would be weak, and (2) that the relationship between achievement and cost-ad-

Table 4.—Regression estimates of effects on 1992 math score

Explanatory variable of interest is cost-adjusted core current per-pupil expenditure (PPE)

Explanatory variable	Model 1 No special-needs controls Coefficient	Model 1 Std. error	Model 2 With special-needs controls Coefficient	Model 2 Std. error
Intercept	5.497[3]	0.738	5.467[3]	0.840
Prior achievement				
Math score, 1988	0.753[3]	0.013	0.750[3]	0.013
Average of other scores, 1988	0.230[3]	0.014	0.231[3]	0.014
Student and family characteristics				
Minority	-0.732[2]	0.219	-0.754[2]	0.218
Female	-1.359[3]	0.136	-1.350[3]	0.135
Single-parent family	-0.341	0.194	-0.312	0.194
Socioeconomic status	0.913[3]	0.107	0.926[3]	0.107
Student interest and effort				
Interest and effort in math	0.354[3]	0.052	0.360[3]	0.052
Time spent on homework	0.180[3]	0.021	0.182[3]	0.021
Class attendance	0.480[3]	0.059	0.471[3]	0.059
Student's view of school environment				
Perceives disruptive environment	-0.201[1]	0.073	-0.205[1]	0.073
Experiences disruptive environment	-0.382[3]	0.055	-0.379[3]	0.055
Peers' characteristics				
Peers from single-parent homes	0.135	0.092	0.179	0.093
Percent minority students	0.014[2]	0.004	0.007	0.004
Peers' absenteeism	-0.179	0.136	-0.136	0.136
Peers' dropout rates	-0.107	0.047	-0.110[1]	0.047
Community characteristics				
Percent adults w/ at least some college	0.020[1]	0.008	0.028[2]	0.009
Median income, hholds w/ kids (000s)	-0.029[1]	0.010	-0.035[2]	0.011
School characteristics				
Size				
Twelfth grade enrollment (00s)	0.226[3]	0.053	0.170[2]	0.055
Problems in school				
Composite of minor to serious problems	-0.052[1]	0.018	-0.046[1]	0.018
Type				
Magnet school	-0.133	0.244	-0.140	0.244
Public school of choice	-0.619[3]	0.141	-0.600[3]	0.142
Year-round school	0.996[1]	0.363	0.769[1]	0.371
Vocational-technical school	0.374	0.254	0.535[1]	0.257
Region (vs. Midwest)				
Northeast	0.796[1]	0.271	0.774[1]	0.271
South	-0.163	0.180	-0.039	0.189
West	0.210	0.221	-0.055	0.236
Urbanicity (vs. Suburban)				
Urban	-0.486[1]	0.227	-0.515[1]	0.230
Rural	-0.278	0.175	-0.223	0.177
Per-pupil expenditures				
Cost-adjusted core current PPEs (000s)	0.406[3]	0.096	0.444[3]	0.097
Special-needs students				
Percent special education	—	—	-0.025	0.017
Percent with limited English proficiency	—	—	0.118[2]	0.035
Percent below poverty level	—	—	-0.001	0.013
	n = 5,955		n = 5,955	
	R-squared = .74		R-squared = .74	

—Not applicable.

[1] Coefficient is twice its standard error.

[2] Coefficient is three times its standard error.

[3] Coefficient is four or more times its standard error.

SOURCE: Author's calculations using the National Education Longitudinal Study (NELS) and Common Core of Data (CCD) data.

justed expenditures would be stronger and positive, when controlling for the population of special-needs students. Instead, I consistently found a small positive relationship that was relatively insensitive to the cost-adjustments and special-needs controls. These results provide evidence that the lack of a strong relationship between student achievement and school expenditures cannot simply be attributed to mismeasurement of the schools' fiscal resources.

In future research I intend to test the robustness of these results. I will consider alternative model specifications and methods of accounting for differential resource costs and student needs. It may be that I find no support for my hypothesis no matter which model or adjustment factors are used, but given the dearth of work in this area, further exploration is warranted. I will examine the degree to which my results are due to assumptions linearity of the model's functional form. I will also examine the extent to which these results are dependent on my choice of cost-adjustment: Chambers' TCI. These and other avenues of exploration should shed further light on the potential effectiveness of school finance reform in affecting student equity.

Appendix A. Teacher Cost Index[13]

The theoretical basis for Chambers' teacher cost index (TCI) is the hedonic wage model. In this model, teachers care about both the quality of their work environment and the monetary rewards associated with particular employment opportunities. School districts care about the characteristics of their workers and the costs of hiring those workers. The hedonic wage model assumes that the simultaneous matching of teachers with school districts reveals the differential rates of pay associated with employee attributes and working conditions offered by employers. Thus, the model allows for decomposition of observed variations in wages into the implicit dollar values attached to each unit of the personal and workplace characteristics.

Chambers represents the reduced form of the hedonic wage model for teacher salaries as:

$$\text{(A1)}\quad \ln(SALARY_{ij}) = \alpha + \beta_D D_j + \beta_R R_j + \beta_T T_i + \beta_C C_i + \beta_S s_i + u_{ij}$$

where i indexes individual teachers and j indexes school districts. The dependent variable is the natural logarithm of the annual earnings of the teacher from the school district. The explanatory variables can be divided into two broad categories: cost factors and discretionary factors. The *cost factors* include district *(D)* and regional *(R)* attributes that affect the willingness of teachers to live and work in these localities and that are beyond the control of local decision makers, e.g., competition in the market for teachers, factors underlying cost-of-living differences, amenities of urban and rural life, climatic conditions, racial-ethnic mix of students, and district size and growth. These cost factors are directly used in calculating the TCI. The other category of explanatory variables used in the hedonic wage model includes *discretionary factors*—those within the control of local school district decision makers in the long run, such as the characteristics of the individual teachers (T), the attributes of the job or classroom to which they are assigned (C), and various school characteristics (S). These discretionary factors are included as control variables in the regression to eliminate their contribution to expenditure differences across districts. (See table 1.1 of Chambers and Fowler, 1995, for details of the specific variables included under each of these categories.)

The data used in the empirical estimation of this model are derived primarily from the Schools and Staffing Survey (SASS). They include responses from 46,750 public school teachers in 8,969 public schools and 4,884 public school districts. These data are supplemented by data from the Common Core of Data, the Census Bureau, the U.S. Geological Survey, and the National Climatic Data Center.

After estimating equation A1, a teacher cost index is calculated for each school district based on the estimated coefficients and values of the cost factors, while controlling for variations in the discretionary factors. The TCI for each school district j is calculated as: $TCI_j = \exp[\, \beta_D (D_j - D) + \beta_R(R_j - R)\,]$.

The overall mean value for the TCI is 100. The index is greater than 100 for districts facing higher non-discretionary costs (e.g., the average TCI for districts in New York City is 130) and is less than 100 for districts in low cost areas (e.g., the average TCI for districts in non-metropolitan Oklahoma is 80).

[13] This summary of the TCI draws heavily from Chambers and Fowler, 1995.

Appendix B. Definitions and Sources of the Variables

Unless otherwise noted, the variables described below are based on variables from the NELS Student Component Data Files. Other sources of data include the NELS School Component Data Files (NELS School), the Common Core of Data (CCD), and the Teacher Cost Index (TCI).

Dependent Variable

Math score, 1992: Score on the mathematics achievement test in the spring of 1992, when most of the students were in twelfth grade. Uses NELS variable F22XMTH, the IRT Theta t-score. (See Ingles et al., 1994, p. H–33 for a description of the benefits of using this metric.)

Explanatory Variables of Interest

Six variables measuring per-pupil expenditures are used in these analyses. These are based on three categories of expenditures (total, core current, and instructional salaries) and two alternative calculations of PPEs (nominal and cost-adjusted).

The three categories of expenditures are from the CCD for Fiscal Year 1992 (School Year 1991–92). Expenditures are measured for the entire school district.

- Measure 1 is **total district expenditures**, field *C_TOTEXP.*
- Measure 2 is **core current expenditures**, defined as instructional expenditures, pupil support services, and instructional staff support: *C_E13 + C_E17 + C_E07.*
- Measure 3 is **instructional salaries** only, *C_Z33.*

The two methods of calculating PPEs are described below:

- **Nominal PPEs** are calculated by simply dividing each of the expenditure measures described above by the total number of students in the school district in School Year 1991–92 *(AG_PK12).* For example, the formula for per pupil total expenditures is *C_TOTEXP/ AG_PK12.*
- **Cost-adjusted PPEs** are calculated by dividing expenditures by Chambers' teacher cost index *(TCI)* multiplying by 100, then dividing by the number of students in the district, e.g., *(C_TOTEXP/TCI*100)/AG_PK12.*

Note that the cost-adjusted measure that is used in the regressions is rescaled to be comparable to the nominal measure within each category. See "Coefficient Comparisons Across Regressions."

Control Variables

Prior Achievement

- **Math score, 1988:** *BY2XMTH,* eighth grade IRT Theta t-score.
- **Average of other scores, 1988:** Average of 1988 IRT Theta t-scores in reading, science, and social studies. *(BY2XHTH + BY2XSTH + BY2XRTH) / 3.* All these test scores are on the same metric; hence the simple average score is appropriate.

Student and Family Characteristics

- **Minority:** Student's race based on *F2RACE1,* recoded to 1=Black, Hispanic, or Native American; 0=White or Asian.
- **Female:** Student's sex based on *F2SEX,* recoded to 1=female; 0=male.
- **Single-parent family:** Adult composition of the student's household based on FAMCOMP, recoded to 1=adult female only or adult male only; 0=two parents or guardians.
- **Socioeconomic status:** *F2SES1,* SES measure based on father's education level, mother's education level, father's occupation, mother's occupation, and family income, and using Duncan's Socioeconomic Index (1961).

Student Interest and Effort

- **Interest and effort in math:** Composite variable based on the student's responses to questions *F2S21A-D*: In your current or most recent math class, how often do you:

  —Pay attention in class?

  —Complete your work on time?

  —Do more work than was required of you?

  —Participate actively in class?

  Composite ranges from 0 (little effort) to 4 (strong effort).

- **Time spent on homework:** Sum of categorical data on hours spent on homework in school *(F2S25F1)* and out of school *(F2S25F2)*. Sum ranges from 0 indicating no time to 16 indicating over 40 hours per week.

- **Class attendance:** Composite variable (uses *F2S9A–F*) measuring the student's attendance in classes, based on how often the student reports he or she:

  —Was late for school.

  —Cut or skipped class.

  —Missed a day of classes.

  —Was put on in-school suspension.

  —Was suspended or put on probation from school.

  Composite ranges from 0 to 5, where 5 indicates the student says he or she "never" did any of the above.

Student's View of the School Environment

- **Perceives disruptive environment:** Composite of the student's perception of the school's learning environment, based on how strongly the student agrees with statements *F2S7E–H*:

  —I don't feel safe at this school.

  —Disruptions by other students get in the way of my learning.

  —Fights often occur between different racial or ethnic groups.

  —There are many gangs in school.

  Composite ranges from 0 to 4, where 4 means the student agreed or strongly agreed with all four statements.

- **Experiences disruptive environment:** Composite measuring the student's personal experiences that indicate a disruptive learning environment. The composite ranges from 0 to 7 and indicates the number of affirmative responses to statements *F2S8A–G*:

  —I had something stolen from me at school.

  —Someone offered to sell me drugs at school.

  —Someone offered to sell me drugs on the way to or from school.

  —Someone threatened to hurt me at school.

  —Someone threatened to hurt me on the way to or from school.

  —I got into a physical fight at school.

  —I got into a physical fight on the way to or from school.

Peers' Characteristics

(All these variables are based on data from the NELS School File)

- **Peers from single-parent homes:** *F2C23*, estimate by school administrator of the percent of twelfth graders (in 1992) from single-parent homes. Coding: 1 indicates less than 10 percent from single-parent homes; 5 indicates more than 75 percent.

- **Percent minority peers:** Percentage of twelfth graders who are Black, Hispanic, or Native American. *F2C22B + F2C22C + F2C22E.*

- **Peers' absenteeism:** Based on *F2C21*, average daily attendance (ADA) rate for twelfth graders, recoded such that 0 indicates 95 percent ADA; 1 indicates 90 percent ADA < 95 percent; 2 indicates 85 percent ADA < 90 percent; 3 indicates ADA < 85 percent. Peers' dropout rate: Based on *F2C26*, estimate of the percent of students who enter the twelfth grade who drop out before graduation. Coded such that 0 means none drop out; 1

means 0 percent dropout rate (DR) < 3 percent; 2 means 3 percent DR < 5 percent; 3 means 5 percent DR < 7 percent; 4 means 7 percent DR < 10 percent; 5 means 10 percent DR < 20 percent; and 6 means 20 percent DR.

Special-needs students

(From the CCD Agency Database for School Year 1991–92)

- **Percent special education:** *AG_SPED/AG_PK12*100*, number of special education students in the district divided by the total number of students in the district, times 100.
- **Percent with limited English proficiency:** *P7028TP*, percentage of children in the district who speak English "not well."
- **Percent below poverty level:** *P7118TP*, percentage of children in the district living below the poverty level.

Community Characteristics

(From the CCD Agency Database for School Year 1991–92)

- **Percent adults with at least some college:** *P120403P + P120404P*, percentage of adults in the district with some college, or a bachelor's degree or higher degree.
- **Median income for household with kids:** *P3080A01*.

Size of Class; Problems in School

(From the NELS School File)

- **Twelfth grade enrollment:** Enrollment of twelfth graders as of Oct. 1991, based on *F2C2*.
- **Problems in school:** Composite of school problems as judged by the school administrator (using NELS variables *F2C57A,C–P*). Composite ranges from 0 to 15, where higher values indicate more of the following problems: tardiness, class cutting, physical conflicts, gang activity, robbery or theft, vandalism, use of alcohol, use of illegal drugs, students under the influence of alcohol or drugs while at school, sale of drugs near school, possession of weapons, physical or verbal abuse of teachers, racial/ethnic conflicts, and teen pregnancy.

School Characteristics

(From the NELS School File)

In the NELS School File, public schools are classified as the following types:

- Comprehensive school (not including magnet school or school of choice);
- Magnet school (including schools with magnet programs, schools within a school); or School of choice (open enrollment/non-specialized curriculum).

  For each of the three types of schools, I assign a 1 if the administrator indicated that the school met the characteristics of that type of school and a 0 if not. Although the definition of comprehensive schools specifically excludes magnet schools or schools of choice, the data reveal that some administrators in magnet schools and/or schools of choice marked that they were also comprehensive schools. In my regression analyses I do not include a variable for comprehensive schools; I do include dummy variables for magnet schools and schools of choice.

Zero-one dummy variables are also included for two other characteristics of schools:

- Year-round schools; and
- Vocational-technical schools.

Region of the Country

Zero-one dummy variables indicate in which of four US Census regions the student attended school in 1992, based on *G12REGON*.

- Midwest—East North Central and West North Central states;
- Northeast—New England and Middle Atlantic states;
- South—South Atlantic, East South Central, and West South Central states; and
- West—Mountain and Pacific States.

Urbanicity

Zero-one dummy variables indicate the urbanicity of the school the student attended in 1992, based on *G12URBN3*.

- Urban—central city;
- Suburban—area surrounding a central city within a county constituting an Metropolitan Statistical Area (MSA); and
- Rural—outside an MSA.

References

Becker, William E. and William B. Walstad. 1990. "Data loss from pretest to posttest as a sample selection problem," *Review of Economics and Statistics*, (1): 184–8.

Chaikind, S.; L. C. Danielson, and M. L. Brauen. 1993. "What do we know about the costs of special education: A selected review," *Journal of Special Education*, 26(4): 344–370.

Chambers, Jay G. 1981. "Cost and price level adjustments to state aid for education: A theoretical and empirical review," in *Perspectives in State School Support Programs*. Eds.: K. Forbis Jordan and Nelda H. Cambron-McCabe. Cambridge, MA: Ballinger Publishing, pp. 39–85.

Chambers, Jay G. and William J. Fowler, Jr. 1995. *Public School Teacher Cost Differences across the United States*, Washington, D.C.: Government Printing Office.

Cohn, Elchanan and Terry G. Geske. 1990. *The Economics of Education*. 3rd ed. New York: Pergamon Press.

Coleman, James S. et al. 1966. *Equality of Educational Opportunity*. Washington, D.C.: Government Printing Office.

Gamoran, Adam. 1996. "Student achievement in public magnet, public comprehensive, and private city high schools," *Educational Evaluation and Policy Analysis*, 18(1): 1–18.

Gamoran, Adam and Robert D. Mare. 1989. "Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality?" *American Journal of Sociology*, 94(5): 1146–83.

Hanushek, Eric A. 1979. "Conceptual and empirical issues in the estimation of educational production functions," *Journal of Human Resources*, 14: 351–388.

Hanushek, Eric A. 1986. "The economics of schooling: Production and efficiency in public schools," *Journal of Economic Literature*, 24 (Sept.): 1141–1177.

Hanushek, Eric A. 1989. "The impact of differential expenditures on school performance," *Educational Researcher*, 18 (May): 45–65.

Hedges, Larry V., Richard D. Laine, and Rob Greenwald. 1994. "Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes," *Educational Researcher*, 23 (April): 5–14.

Ingels, Steven J., Kathryn L. Dowd, John D. Baldridge, James L. Stripe, Virginia H. Bartot, Martin R. Frankel, and Peggy Quinn. 1994. *National Education Longitudinal Study of 1988 Second Follow-up: Student Component Data File User's Manual*. Washington, D.C.: National Center for Education Statistics.

Jencks, Christopher S. 1985. "How much do school students learn?," *Sociology of Education*, 58 (April): 128–135.

Kakalik, James; W. S. Furry, M. A. Thomas, and M. F. Carney. 1981. *The Cost of Special Education*. Santa Monica, CA: The RAND Corporation.

Lucas, Robert E. B. 1972. "Working conditions, wage-rates and human capital: A hedonic study," Ph.D. dissertation, Massachusetts Institute of Technology.

McMahon, Walter W. 1995. "Intrastate cost adjustments," paper presented at the National Center for Education Statistics Spring Meeting, Washington.

McMahon, Walter W. 1988. "Geographical cost of living differences: An update," *MacArthur/Spencer Series*, 7.

Monk, David H. 1992. "Education productivity research: An update and assessment of its role in education finance reform," *Educational Evaluation and Policy Analysis*, 14 (Winter): 307–332.

Moore, Mary T., E. William Strang, Myron Schwartz, and Mark Braddock. 1988. *Patterns in Special Education Service Delivery and Cost*. Washington: Decision Resources Corporation.

Moulton, Brent R. 1990. "An illustration of a pitfall in estimating the effects of aggregate variables on micro units," Review of Economics and Statistics, (2): 334–8.

National Center for Education Statistics. 1995. Common Core of Data (CCD): School years 1987–88 through 1992–93. Washington, D.C.: US Department of Education, Office of Educational Research and Improvement.

Odden, Allan R. and Lawrence O. Picus. 1992. *School Finance: A Policy Perspective*. New York: McGraw-Hill.

Parrish, Thomas B., Christine S. Matsumoto, and William J. Fowler, Jr. 1995. *Disparities in Public School District Spending: 1989–90*. Washington: Government Printing Office.

Sander, William. 1993. "Expenditures and student achievement in Illinois: New evidence," *Journal of Public Economics*, 52: 403–416.

School District Expenditures, School Resources and Student Achievement: Modeling the Production Function

Harold Wenglinsky

Educational Testing Service

About the Author

Harold Wenglinsky is an Associated Research Scientist in the Policy Information Center at Educational Testing Service. The Policy Information Center conducts original research as well as syntheses of existing research on important questions of educational policy for dissemination to policymaking audiences. Recent studies Dr. Wenglinsky has authored at the Center include *When Money Matters: How Educational Expenditures Improve Student Performance and How They Don't* and *Students at Historically Black Colleges and Universities: Their Aspirations and Accomplishments.* He is currently studying the application of multilevel structural equation modeling to data on school effects. Before joining the Policy Information Center, Dr. Wenglinsky was the ETS National Assessment of Educational Progress Visiting Scholar. He has been the recipient of numerous grants and awards from various organizations including the National Science Foundation and the American Educational Research Association. (The paper presented here was funded in part by a grant from NSF, REC-9628157.) Dr. Wenglinsky received his Ph.D. in Sociology from New York University in 1996.

School District Expenditures, School Resources and Student Achievement: Modeling the Production Function

Harold Wenglinsky
Educational Testing Service

Introduction

After more than 30 years of research, social scientists have made little progress in identifying the educational production function. "Production function" studies are those that use some form of multivariate analysis, such as regression analysis, to measure associations between various educational inputs, such as per-pupil expenditures, and outputs, such as academic achievement as measured by standardized tests.[1] One of the earliest studies of this type was the Equality of Educational Opportunity Report, commonly referred to as the Coleman Report (Coleman et al. 1966). This study found little association between inputs and outputs for a nationally representative sample of students and schools. Since the publication of the Coleman Report, nearly 400 additional studies of this sort have been conducted. Their results have been mixed, fueling, rather than resolving, the debate as to whether money matters to educational achievement (see Hanushek 1997 for list of studies).

Because of the mixed results of this very large number of studies, some researchers have concluded that the production function approach is flawed and should be abandoned. In their view, production function studies suffer from a multitude of problems, including their failure to analyze different types of educational expenditure (such as spending on instruction and administration) and their failure to adjust for regional variations in the cost of education (Fortune and O'Neil 1994). Some researchers suggest alternate approaches to estimating the relationship between expenditures and achievement. Monk (1992) suggests conducting small-scale studies at low levels of aggregation, such as the classroom level; Fortune and O'Neil (1994) suggest comparing the achievement levels of specific subgroups, such as high-spending and low-spending urban school districts.

This paper contends that the production function approach is salvageable; the problems researchers have identified can be addressed, producing meaningful results. The present study provides an example of how this may be done. It applies structural equation modeling and multilevel modeling to recently developed databases of fourth-graders. The study is national in scope, distinguishes between different types of spending, and adjusts for regional variations in the cost of education, thus addressing many of the issues raised by critics of the production function approach. The study finds that, at least for fourth-graders, some inputs are strongly associated with academic achievement, while others are not: Instructional expenditures, central office administration expenditures, and teacher-student ratios are all associated with achievement; principal's office expenditures, capital outlays, and teacher education levels are not. Before discussing these results and their derivation,

[1] For the purposes of this paper, "expenditures" refers to actual dollars spent by school districts, "resources" to quantifiable goods made available to schools, and "inputs" to both expenditures and resources.

however, it is necessary to touch upon the methodological issues involved in production functions.

Background

Production function studies of education have been undertaken for more than 30 years. By one estimate nearly 400 production function studies have been conducted and published since the Coleman Report of 1966 (Hanushek 1997, 1996). These studies have tended to use samples that are smaller in their geographical scope than the national Coleman Report, and have studied the same sorts of inputs that report did (aggregate per-pupil expenditures). These studies have come to different conclusions regarding the production function, some finding relationships between a given input and academic achievement, and others finding no such relationship.

More recently, studies known as "meta-analyses" have applied statistical techniques to synthesize the findings from production function studies; these too arrived at contradictory conclusions. Hanushek (1989) conducted a meta-analysis covering both expenditure and resource measures, and found no relationship between these inputs and academic achievement. Hanushek synthesized the findings of 187 production function studies using the technique of vote counting. He first divided each study into its component inputs. A study that related class size and teacher education to achievement, for example, was divided into those two inputs. Each input was then placed in one of seven categories: per-pupil expenditures, teacher experience, teacher education, teacher salary, teacher-student ratio, administrative inputs, and facilities. Within each category, the relationship of the input to the studied output was classified as positive and statistically significant, positive and statistically non-significant, negative and statistically significant, negative and statistically non-significant, and non-significant but of unknown direction. Hanushek found most relationships to be non-significant. Of 65 aggregate per-pupil expenditure relationships, for example, he found 13 to be positive and significant, 3 negative and significant, and 49 to be non-significant. He concluded that "there is no strong or systematic relationship between school expenditures and student performance (1989, 47)."

> *More recently, studies known as "meta-analyses" have applied statistical techniques to synthesize the findings from production function studies; these too arrived at contradictory conclusions.*

Hedges, Laine, and Greenwald (1994) reanalyzed most of the same studies, and drew the opposite conclusion. They first excluded from their analysis the relationships Hanushek had classified as non-significant but of unknown direction. For the remaining relationships, they reinterpreted Hanushek's vote counting in the context of rules regarding statistical significance. They argued that, if the relationships are treated as a sample, in order to draw the conclusion that there is no relationship between an input and achievement, no more than 5 percent of the relationships could be significant, and these relationships would have to be equally divided between the positive and negative directions. Yet, in fact, if relationships of unknown direction are excluded, many more than 5 percent of the relationships are significant (up to 30 percent for per-pupil expenditures); most of the significant relationships are in the positive direction. The bulk of insignificant relationships are also in the positive direction.

After reinterpreting the vote count, Hedges, Laine, and Greenwald (1994) applied a significance test, the inverse chi-square, to combine the relationships for each input into a single significance measure. They tested two hypotheses: that each input is positively related to achievement, and that each is negatively related to achievement. They found, for the full sample of relationships (as well as for various subsamples), that almost all relationships were significant in the positive direction, with a few others being significant in the negative direction. Finally, for each input, Hedges and his colleagues combined the coefficients from those studies that reported them by calculating their median. They found positive coefficients for per-pupil expenditures, teacher experience, teacher salary, administrative inputs, and facilities, and mixed results for class size, and concluded that resources affect achievement.

Hanushek (1996) continued the debate, countering the meta-analysis of Hedges, Laine, and Greenwald. He updated his sample of studies to include those published after his 1989 meta-analysis, making a total of 377 studies. Hanushek again found, when he classified relationships into the seven categories, that the bulk of studies indicated no significant relationship between resources and achievement. In a counter-study of their own, Greenwald, Hedges, and Laine (1996) created their own sample of studies, and placed the relationships they identified from the studies into seven somewhat different categories: per-pupil expenditures, teacher ability, teacher education, teacher experience, teacher salary, teacher-pupil ratio and school size. They again found, for both this new sample and for various subsamples, that the combined significance test and median effect sizes supported the hypothesis that resources affect achievement. Most recently, Hanushek (1997) has compared his sample of 377 studies to the sample of Greenwald, Hedges, and Laine (1996), and found that the latter sample systematically over-represented positive relationships.[2]

The fact that different meta-analyses can reach different conclusions from similar sets of studies indicates that the underlying studies are quite volatile in their results when subjected to different assumptions. This volatility was even revealed within the meta-analyses. For instance, Hedges, Laine, and Greenwald (1994) were able to find support both for the hypothesis of a positive relationship and that of a negative relationship between a given resource and achievement when using combined significance tests. Both Greenwald, Hedges, and Laine (1996) and Hanushek (1997) found that the results from a subsample of longitudinal studies differed markedly from those of the full sample. What the meta-analyses reveal most clearly, then, is that the original studies do not provide a clear answer to the question of whether or not money matters.

The fact that different meta-analyses can reach different conclusions from similar sets of studies indicates that the underlying studies are quite volatile in their results when subjected to different assumptions.

This lack of consensus among the meta-analyses reflects to some degree shortcomings in the methods of the original studies. Six shortcomings have been commonly noted.

First, unlike the Coleman Report, most subsequent studies were not nationally representative, but instead studied a particular state or school district. This hampers development of a consensus, because different regions of the country may have different spending patterns and different relationships between these spending patterns and student achievement.

Second, the studies did not distinguish among different types of spending. While they measured multiple inputs, such as teacher experience and teacher-student ratios, the only expenditure measure used was aggregate per-pupil expenditures. Using such a gross measure risks missing certain dynamics in the relationship between school spending and academic achievement, as increases in some types of spending may have an effect while increases in others may not. For instance, increased spending on administration may not significantly raise achievement, while increased spending on instruction may. If these types of spending are not measured separately, the apparent effects of spending on instruction will be reduced or eliminated when combined with the lack of effects from administration.

[2] Other meta-analyses have also arrived at contradictory conclusions. With regard to class size, Glass and Smith (1979) found a clear and consistent relationship while Odden (1990) did not. The effect of class size on student achievement has also been the subject of a controlled experiment in which students in kindergarten and first grade were randomly assigned to small and large classes. The study found significant achievement differences that persisted even after the students in small classes were returned to large ones (Finn and Achilles 1990; Mosteller 1995; Mosteller, Light, and Sachs 1996). This finding, like those of production function research, has been the subject of controversy (Hanushek 1997).

Third, the studies did not take into account the ways in which other influences on the process of schooling may mediate between spending and achievement. Effective schools research suggests that certain aspects of the school environment, particularly supportive relations between teachers and principals, positively influence achievement.[3] Yet none of the prior research has sought to measure the influence of school spending patterns on school environment.

Fourth, not all of the studies provided rich measures of student background.[4] While the research on measures of the socio-economic characteristics of students indicates that a single measure, socio-economic status (SES), can be generated by adding together responses to a relatively small number of questions, many studies did not include such questions. If SES is poorly measured, it is difficult to determine if relationships between spending and achievement are attributable to some degree to SES differences between students in high- and low-spending districts.

Effective schools research suggests that certain aspects of the school environment, particularly supportive relations between teachers and principals, positively influence achievement.

Fifth, most studies did not control for variations in cost between regions. The cost of living in New York City is higher than the cost of living in Montgomery, Alabama, and presumably this difference means that teachers paid the same actual dollars in the two cities are not able to maintain the same standard of living; a dollar will buy less in New York City. As a result, New York City would have to offer higher salaries to recruit successfully the same teachers as Montgomery.[5] Other factors may also influence the cost of hiring comparable teachers, including union pressure to increase wages and the overall quality of life in the region. Most studies did not take these factors into account, and they may be as important as SES, in that differences in achievement between two districts may be due to some degree to differences in how much it costs to hire teachers.[6]

Sixth, many of the measures of achievement used by earlier studies were unsophisticated. Some did not use achievement measures at all but merely relied on proxies, such as graduation rates. Some used measures as simple as whether or not a student passed a minimum competency test. Few took into account modern developments in test theory, such as Item Response Theory (IRT).[7]

Finally, the prior research has not taken into account the multilevel nature of school effects. Measuring the relationship between school characteristics and student achievement entails relating variables whose level of analysis is the school or school district to an outcome whose level of analysis is the student. Various estimation techniques have been developed that take the multilevel nature of school effects into account, and it has been found that these techniques sometimes produce results that differ from more conventional techniques. In particular, conventional techniques often underestimate standard errors and, in some cases, fail to identify important components of school effects (Raudenbush and Willms 1995; Bryk and Raudenbush 1992). Production function models have generally not made use of estimation techniques that are sensitive to multilevel data, and consequently may produce inaccurate results.

[3] Despite some early criticism of effective schools research (e.g., Cuban 1984; Purkey and Smith 1983), later large scale multivariate studies have persuaded most researchers that there is a social dimension to school life that plays some independent role in student achievement. The extent of this role is, however, still being debated (Lee, Bryk, and Smith 1993).

[4] This was pointed out by Hedges, Laine, and Greenwald (1994, 12).

[5] When cost of living is taken into account, differentials in per-pupil expenditures between high-spending and low-spending states decrease markedly, indicating that states with fewer resources often tend to be states with lower costs of living (Barton et al. 1991).

[6] This was pointed out by Fortune and O'Neil (1994, 24).

[7] For a discussion of this shortcoming in production function research, see Fortune and O'Neil (1994, 24). For a discussion of IRT, see Hambleton et al. (1991).

Some researchers doubt that these problems can be addressed, and have argued that the production function approach should be abandoned, and alternate approaches explored. Monk (1992) proposes to shift the unit of analysis for school resource studies to the classroom level. He notes that prior research has found a great deal of variation in the efficacy of teachers within the same school, as well as variation in the efficacy of a particular teacher during different classes. He classifies teachers as being of two types, those who are engaged with their classes, trying actively to address any problems in them, and those who are accommodating, seeking only to avoid dealing with problems. He views whether the teacher chooses the engagement or accommodation route as dependent upon a number of factors, including resource decisions made at the school level. Monk calls for a research program in which teachers are interviewed to provide retrospective information on the problems they face, their responses to these problems, and the degree to which resources operate as a constraint.

Another alternative to production function research is the threshold approach, proposed by Fortune and O'Neil (1994). They argue that the key problem with production function research is its use of linear models. They hypothesize that input-output relationships occur in a punctuated manner, with small increments of inputs having no effect on achievement, but large increments having a large effect. To estimate this effect, they propose comparing the mean achievement levels of school districts that are in the top 30 percent in terms of spending to school districts that are in the bottom 30 percent. To address the problem that demographic variables might be at the root of achievement differences, they propose using demographically similar school districts for the comparison. They also propose eliminating outlying cases. In applying this approach to samples of school districts in Missouri and Ohio, they find that while correlation coefficients rarely uncover an input-output relationship, the threshold approach often finds one.

Such alternatives, however, raise their own methodological issues. The most significant is that in both cases it is difficult to separate the factors contributing to student achievement.

Such alternatives, however, raise their own methodological issues. The most significant is that in both cases it is difficult to separate the factors contributing to student achievement. In the classroom-based approach, efficacious and non-efficacious teachers are identified and the resource constraints traced. Yet, the availability of one type of resource tends to be highly correlated with availability of another. Since the efficacious teacher may have many resources available at once, it will be difficult to determine which is the basis of high teacher efficacy. In addition, it will be difficult to determine whether high student achievement is primarily attributable to teacher efficacy or student characteristics.[8] In the threshold approach, only school districts with extremely high and extremely low levels of expenditures are compared. In many areas the high expenditure districts will have high levels of resources and the low expenditure districts low levels of resources. It will thus be difficult to determine which resource is responsible for achievement levels. In addition, it will be difficult to determine whether or not the resource levels in the schools or the resource levels in the communities account for achievement differences.[9]

Given that these alternate approaches raise their own difficulties, it may also be worthwhile to salvage the production function approach through addressing its problems. The present study is an attempt to do just that.

The Design of the Study

Hypotheses

This study hypothesizes that there are various potential paths through which school district expen-

[8] It is also worth noting that to draw conclusions about expenditures and resources, the classroom approach will still have to collect district- and school-level data, because expenditure and resource decisions are made at those levels, not the classroom level. Thus many of the methodological problems of production function studies will also hold for the classroom approach.

[9] NCES (1995a) found that school district expenditures and the average SES of the districts were strongly related.

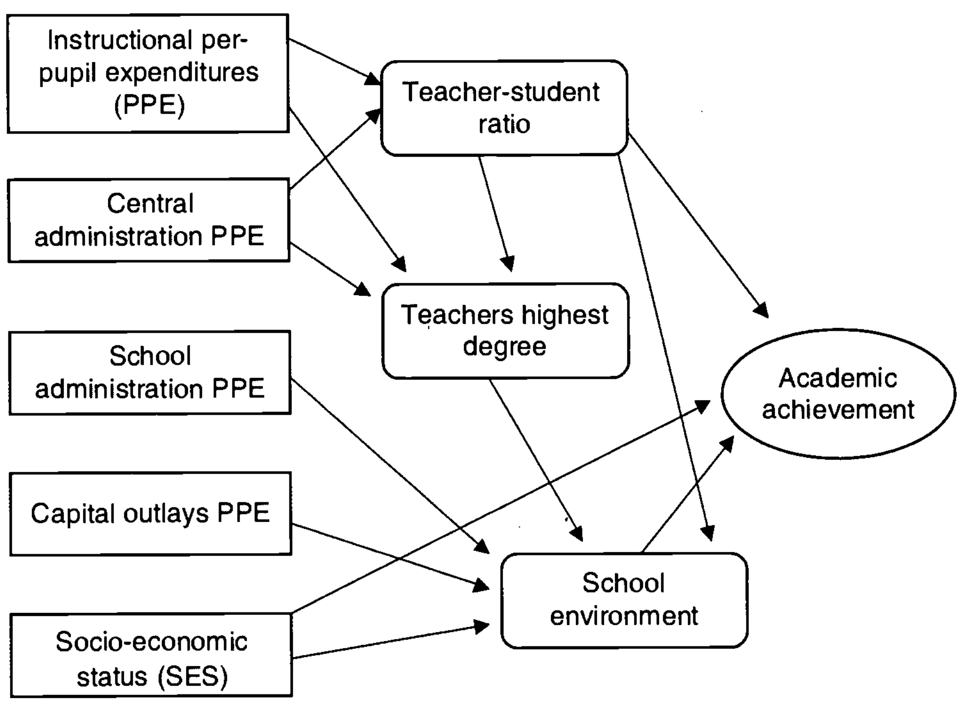
ditures and school resources can influence student achievement (figure 1). These paths occur in three basic steps. First, the allocation of money at the school district level influences the availability of resources at the school level. Most decisions about how to spend money are made by school superintendents and their staffs. These spending decisions determine how much of each school resource is purchased, and therefore what is available in the school. For instance, more spending on instruction will lead to some combination of more teachers per student, higher teacher salaries and more instructional materials.

Second, the availability of resources has consequences for the school climate. Schools vary widely in environment, some possessing low levels of student and teacher absenteeism, collegial relationships between teachers and principals, and a lack of disruptive and delinquent behaviors, and others possessing the opposite. In part, environment is influenced by the availability of resources; teachers who are paid lower salaries, for example, might be expected to be more frequently absent.

The third step involves the influence of school climate on student achievement. Effective schools research suggests that school climate strongly influences student performance (Lee, Bryk, and Smith 1993; Austin and Garber 1985; Brookover et al. 1979; Edmonds 1979). Disruptive students, high levels of student and teacher absenteeism, and frayed principal-teacher relations can be expected to interfere with the ability of teachers to instruct and students to learn.

It is hypothesized here that some educational expenditures influence achievement via these steps. Four types of expenditures are considered: instructional expenditures, central office administration expenditures, principal's office administration expenditures, and capital outlays. The first two, it is hypothesized, will directly affect school resources.

Figure 1.—Hypothesized paths to achievement



SOURCE: Wenglinsky, unpublished.

BEST COPY AVAILABLE

Research has shown that expenditures are typically invested in one of two resources, increasing the number of teachers per student or improving teacher quality (Odden and Clune 1995). It is therefore expected that instructional and central office administration expenditures will influence one or both of these resources. These resources will, in turn, affect the school climate, which will itself affect student achievement. It is also expected that capital outlays and principal's office administration will play a role in the learning process. While it is unlikely that spending in either area would directly affect the number of teachers in the classroom or the types of teachers hired, it is expected that they will influence the school climate which will itself influence student achievement.

The model hypothesized here also must take into account the role of two factors outside the school in the spending-achievement relationship. First, student SES can be expected to affect the school climate and student achievement; students from more affluent backgrounds will be more likely to meet the social demands of the school, develop a rapport with teachers, and be better prepared to achieve at high levels (Hauser, Sewell, and Alwin 1976; Jencks et al. 1972; Coleman et al. 1966). Second, the cost of education can be expected to affect the ability of expenditures to purchase school resources and influence the school climate. A given level of expenditures will not go as far in a high-cost region.

> *. . . unlikely that spending in either area . . . directly affect the number of teachers in the classroom or the types of teachers hired, . . . will influence the school climate which . . . influence student achievement.*

Data

The data employed to test this model are drawn from three sources: the National Assessment of Educational Progress (NAEP), the Common Core of Data (CCD), and a Teacher's Cost Index (TCI). NAEP is a nationally representative database of students and schools collected by the Educational Testing Service (ETS) under a contract from the National Center for Education Statistics (NCES); CCD is a database consisting of the universe of school districts in the United States, collected by NCES; and the TCI was developed by NCES to measure regional variations in the price of teachers. Three data sources had to be used because none contains all of the necessary measures.

NAEP is administered by ETS every 2 years to nationally representative samples of fourth-, eighth- and twelfth-graders, and to their teachers and principals. The subject areas tested vary, but have included at various times mathematics, reading, history, geography, and science. The information collected by NAEP is used to assess the knowledge of students throughout the country; to make comparisons in the levels of knowledge of various regional, ethnic, socio-economic, and gender subgroups; and to measure the progress of students in the nation, both over time and between grades (see Johnson 1994 for overview of NAEP; Mullis, Dossey, Owens, and Phillips 1993 for report card for 1992 mathematics assessment.) The 1992 mathematics assessment of students attending fourth grade was used in this study.[10] It contains measures of mathematics achievement, school environment, teacher education levels, teacher-student ratios, and student- and school-level SES.[11]

CCD is a database of financial information provided by the universe of U.S. school districts. All school districts send this information to the U.S. Department of Education on a yearly basis. While

[10] Eighth-graders are analyzed in another study (Wenglinsky 1997).

[11] The NAEP SES variables have been criticized for relying on student self-reports and not including a family income measure. In their comparison of various large scale databases that used both a student and parent self-report, however, Berends and Koretz (1995) found little difference between the two types of reports, suggesting that the use of student self-reports is not problematic. In terms of the lack of a family income measure, while this may be true on the student level, there is an indicator of family income at the school level—the percentage of students receiving free or reduced-price lunches, which is used as part of the SES measure in this study.

the information provided can be used to measure district-by-district per-pupil expenditures in broad spending categories, such as instruction or capital outlays, it cannot be relied upon for more detailed information because differences in the charts of accounts of school districts result in their categorizing specific expenses differently. Therefore, CCD was used to provide measures of expenditures on instruction, central office administration, school-level administration, and capital outlays only. CCD was used here, even though the district level is its lowest level of aggregation, because no nationally representative database exists that measures different types of expenditures at a lower level of aggregation.[12]

The TCI is the result of a study by NCES. NCES has conducted analyses to develop an index of the cost of hiring teachers for particular regions of the country (NCES 1995b). This cost can be expected to vary by region, even for teachers of similar levels of experience and education, because the cost of living, quality of life, and other factors all differ by region. The TCI was developed by applying regression analysis to the Schools and Staffing Survey (SASS), an NCES survey, conducted in 1990–91. The regression analysis estimates the influence of various factors on teacher salaries; these include factors that are under the control of schools and school districts, such as teacher experience and education, as well as those that are not, such as the cost of living and quality of life. The resulting estimates of the impact of these non-discretionary characteristics on teacher salaries can then be used as estimates of teacher costs in a particular region, holding constant the discretionary factors. TCI scores have been estimated for each state, and these are used in this analysis to adjust the per-pupil expenditure measures (NCES 1995b, 51).[13]

For this study the NAEP data were for fourth-graders taking the 1992 mathematics assessment. This sample consisted of 9,414 students in 270 school districts around the United States.

To analyze data from these sources, all three needed to be linked together. For this study the NAEP data were for fourth-graders taking the 1992 mathematics assessment. This sample consisted of 9,414 students in 270 school districts around the United States. Of the school districts, 48 were private schools and therefore no corresponding information was available in CCD. Of the remainder, 195 school districts were linked to CCD through common identification numbers, 8 were linked through common address information, and 19 (7 percent of the sample) could not be matched. State-level TCI scores were linked to CCD and NAEP by locating the state in which each school district was located and entering the appropriate TCI score. These linking procedures were used to produce two databases, one at the district level and one at the student level. The district-level database was produced by aggregating NAEP data to the district level and linking it to the already district-level CCD. The student-level database was produced by disaggregating CCD to the student level and linking it to the already student-level NAEP. The district-level database was used for all analyses except the multilevel one, for which the student-level database was used. Because NAEP is a sample while CCD and TCI are universes, the two databases took on the sampling characteristics of NAEP; this means that the databases are nationally representative samples of public schools and their students and that the weighting techniques and standard error adjustments required for NAEP apply.[14]

[12] Few school systems collect budget information at the school level. For a recent study addressing this issue in Texas and Ohio, see NCES (1996).

[13] The TCI is a cost-of-education index. It differs from a cost of living index in that a cost of living index only measures the cost of living, while a cost-of-education index measures other factors that affect the cost of education as well. See Barro (1994) for a discussion of the differences between cost-of-living indices and cost-of-education indices.

[14] Because NAEP is not nationally representative at the district level, all results refer to students and schools in districts, and not the districts themselves (Johnson, Rust, and Wallace 1994).

The district-level database was then used to produce measures of the variables needed to test the hypotheses (see table 1 for means and standard deviations and appendix A for full definitions). The database included the four expenditure measures, the number of pupils in the school district and the TCI score for that state. Cost-adjusted per-pupil expenditures in the four areas were calculated by dividing each by the number of pupils and the TCI. The database also included seven SES measures summed to create an SES variable; seven school environment measures summed to create a school environment variable; a measure of teacher's highest degree attained; the number of full-time teachers and students in the school, used to calculate the school teacher-student ratio; and five measures of mathematics achievement known as "plausible values," the use of which will be discussed below.

Method

The bulk of analyses were conducted on the district-level database using a structural equation modeling program, LISREL 8. LISREL requires as input rules regarding which variables are allowed to be related to one another and which are not, and a covariance matrix calculated from data. The program then estimates parameters relating the variables allowed to be related while maximizing the goodness of fit between the covariance matrix these parameters imply and the input covariance matrix. LISREL produces three principal outputs: the estimates of the direct effects between variables; estimates of the total effects between variables; and the goodness of fit as measured by adjusted goodness-of-fit and normed goodness-of-fit indices (Joreskog and Sorbom 1993). Models are considered to have a satisfactory fit when the chi-square is statistically insignificant (indicating that there is no significant difference between the input covariance matrix and the implied covariance matrix) and the adjusted and normed goodness-of-fit indices are more than 0.9 (Bentler and Bonnett 1980). LISREL also allows for the comparison of goodness of fit between the hypothesized model (referred to here as the full model) and a model in which the relationships found the be significant in the full model are fixed as being unrelated to one another (referred to here as the nested model). By running such a nested model and comparing its chi-square to that of the full model, it is possible to reject the nested model in favor of the full one (Hayduk 1987).

First, full and nested models were designed to test the hypothesized relationships. For the full model, the four cost-adjusted per-pupil expenditure measures and the SES index were treated as exogenous variables; their values were not allowed to depend on those of the other variables. Per-pupil expenditures on instruction and central office administration were allowed to affect school environment; SES was allowed to affect school environment and academic achievement; teacher-student ratio was allowed to affect teacher education, school environment and academic achievement; teacher education was allowed to affect school environment and academic achievement; and school environment was

Table 1.—Means and standard deviations

	Mean	Standard deviation
Instructional per-pupil expenditures (PPE)* (dollars)	2999.89	754.47
Central administration PPE* (dollars)	113.38	99.29
School administration PPE* (dollars)	287.72	89.65
Capital outlays PPE* (dollars)	499.54	548.66
Socio-economic status (summated scale)	13.58	2.36
Teacher-student ratio (number of teachers/students)	0.05	0.01
Teacher's highest degree (1=<BA, 2=BA, 3=MA, 4=>MA)	2.53	0.40
School environment (summated scale)	22.39	2.14
Mathematics achievement (mean for five plausible values)	210.65	17.41

*Adjusted for regional variations in the cost of education.

SOURCE: Wenglinsky, unpublished tabulations.

allowed to affect academic achievement.[15] For the nested model, the relationships that were found to be significant in the full model and were either directly or indirectly associated with achievement were fixed at zero (making them unrelated to one another).

A design effect was then calculated through running a series of preliminary LISREL models. LISREL parameter and standard error estimates assume a simple random sample, and since NAEP is a clustered, stratified sample, these estimates are inaccurate (Johnson 1989). To adjust parameters for the NAEP sample design, covariance matrices used in all analyses were weighted by a student base weight, provided by the NAEP database. Covariance matrices were also weighted by the number of students in each school district. To adjust standard errors for the NAEP sample design, a design effect that estimated the amount by which the standard error estimate was downwardly biased in assuming a simple random sample had to be calculated. This was accomplished by first running a LISREL analysis for the full model on a covariance matrix weighted by only the student base weight and the number of students per school district, thus producing baseline estimates. LISREL analyses were then conducted for the full model on 56 covariance matrices, each weighted by the jackknife replicate weight provided by the NAEP database. For three representative relationships, the variance of the 56 estimates was calculated and the variance for the baseline model was divided by this jackknife variance, producing three estimated design effects, the most conservative of which was used for subsequent analyses (1.75).

Five full models were then run on five covariance matrices. Five models needed to be run to take into account "plausible values" methodology in the measurement of academic achievement. Students who take the NAEP examination each receive only a subset of the items. In order to impute total scores, it is necessary to use models that take into account other information about the students, including their demographic characteristics. Five achievement scores are produced for each student, each based upon slightly different models. The variability of the scores needs to be taken into account in the estimation of standard errors of all coefficients in which achievement scores are involved (Johnson, Mislevy, and Thomas 1994).[16] This analysis employed a standard methodology, conducting five LISREL analyses for the full model on five covariance matrices, each using one of the plausible values as its achievement measure; calculating parameters as the mean of those for the five analyses; and then adjusting the mean of the standard errors for the five analyses by multiplying by the square root of the design effect and, for the parameters involving achievement, adding the product of 1.2 and the variance of the five parameter estimates (O'Reilly et al. 1996, 78–79). In order to assess goodness of fit, five nested models were run on the same covariance matrices as were used for the full models, and the mean of the goodness-of-fit statistics for the five full models were compared to the mean of the goodness-of-fit statistics for the five nested models.

Finally, a multilevel estimation program, Hierarchical Linear Modeling (HLM) was applied to the student-level database to test the sensitivity of the LISREL model to multilevel data structure.

Finally, a multilevel estimation program, Hierarchical Linear Modeling (HLM) was applied to the student-level database to test the sensitivity of the LISREL model to multilevel data structure. Much of the LISREL model involved a single-level data structure and was therefore not re-estimated as a multilevel model; the relationships among the first three steps of the model, expenditures, resources and social environment, all involve district- or school-level variables. The relationships between resources and student achievement, however, involve school-

[15] Teacher education was not allowed to reciprocally affect teacher-student ratio, in order to keep the model recursive. The choice of having teacher-student ratio precede teacher education was arbitrary, but, as indicated by modification indices, did not significantly affect the goodness-of-fit of the model.

[16] Researchers have recently proposed an alternate approach to plausible values, known as direct estimation (Cohen 1998).

level independent variables and a student-level dependent variable, the situation under which multilevel techniques are appropriate. The HLM thus consisted of student achievement as the dependent variable and the two resources (teacher's highest degree and teacher-student ratios) as independent variables. As in the LISREL model, SES was incorporated as a statistical control. School-level SES, the school-level aggregate of student-level SES, was included as an independent variable, and the student-level relationship between SES and achievement was included as an additional dependent variable. Plausible values methodology is handled automatically by HLM, which ran separate models for each plausible value and combined them into a single model (Bryk, Raudenbush, and Congdon 1996). The resulting model thus takes the underestimation of standard errors due to measurement variability into account, although it does not take the underestimation of standard errors due to sampling variability into account.

To confirm that a particular expenditure or resource is part of the production function, four results must occur. First, the direct effects measured that trace a path to student achievement must be statistically significant; if they are not, it brings into doubt the reliability of the model. Second, the goodness-of-fit measures for the full models must all confirm the models, while those for the nested models must be unsatisfactory; if not, the null hypothesis may hold. Third, the total effects should be substantial enough for a feasible level of investment to produce marked improvements in student performance; if not, the inputs are not of interest from a policy standpoint. Fourth, the HLM results should be consistent with the LISREL results; otherwise, the latter may be rejected for failing to take into account the multilevel nature of the data.

The expenditure and resource variables measured in the structural equation model are consistent with what is generally known.

Results

The expenditure and resource variables measured in the structural equation model are consistent with what is generally known (table 1).[17] Instructional per-pupil expenditures are, on average, \$3,000 per student, and 68 percent of the school districts in the sample spend between \$2,620 and \$3,380. This spending level constitutes 60 percent of current per-pupil expenditures. Central administration per-pupil expenditures are \$113 per student, and school administration per-pupil expenditures are \$288, constituting 3 percent and 6 percent of current per-pupil expenditures, respectively. These amounts for administrative expenditures might appear low, but are in fact consistent with estimates from other studies. Administrative expenditures refer only to superintendents, principals and their staffs, and so do not include support services, from student transportation to janitorial services, that are often perceived as being part of the administrative category. Five hundred dollars are spent per-pupil on capital outlays, and here there is wider variation than with the other expenditure variables; the standard deviation is nearly \$550. The average teacher-student ratio is 0.05 teachers per student, which means 1 teacher for every 20 students. This seems to be a low number, except that it also includes special education classes, which may have teacher-student ratios as low as 1:1. The average teacher's highest degree is somewhere between a bachelor's and a master's.

The estimates from the full structural equation model reveal that some expenditures and resources are part of the production function while others are not (table 2). Instructional and central office admin-

[17] For examples of distributions of expenditures found in other studies that conform to those found here, see Adams (1994) and Miles (1995).

Table 2.—LISREL estimates of direct effects

	Teacher-student ratio	High degree	School environment	Mathematics achievement
Instructional per-pupil expenditures (PPE) (thousands of dollars)	0.0000051** 0.0000015 0.3000870	0.0000787 0.0000524 0.1484077	—	—
Central administration PPE (thousands of dollars)	0.0000368** 0.0000108 0.2850938	0.0000938 0.0003962 0.0232586	0.0023330 0.0019294 0.1080068	—
School administration PPE (thousands of dollars)	—	—	0.0019241 0.0020389 0.0804268	—
Capital outlays PPE (thousands of dollars)	—	—	-0.0001847 0.0003314 -0.0472520	—
Socio-economic status (summated scale)	—	—	0.3679910** 0.0778429 0.4416251	5.6298038** 0.4050784 0.6395601
Teacher-student ratio (number of teachers/students)	—	-2.1375013 3.2277413 -0.0683488	-17.6074397 15.0059229 -0.1050825	152.1944001** 74.5571103 0.1119985
Teacher's highest degree	—	—	-0.8180802* 0.4473701 -0.1526883	1.5975948 2.4070308 0.0367765
School environment (summated scale)	—	—	—	0.5822595 0.4903571 0.0718114
Mathematics achievement (plausible values)	—	—	—	—

— Relationship fixed at zero.

* p<.10

** p<.05

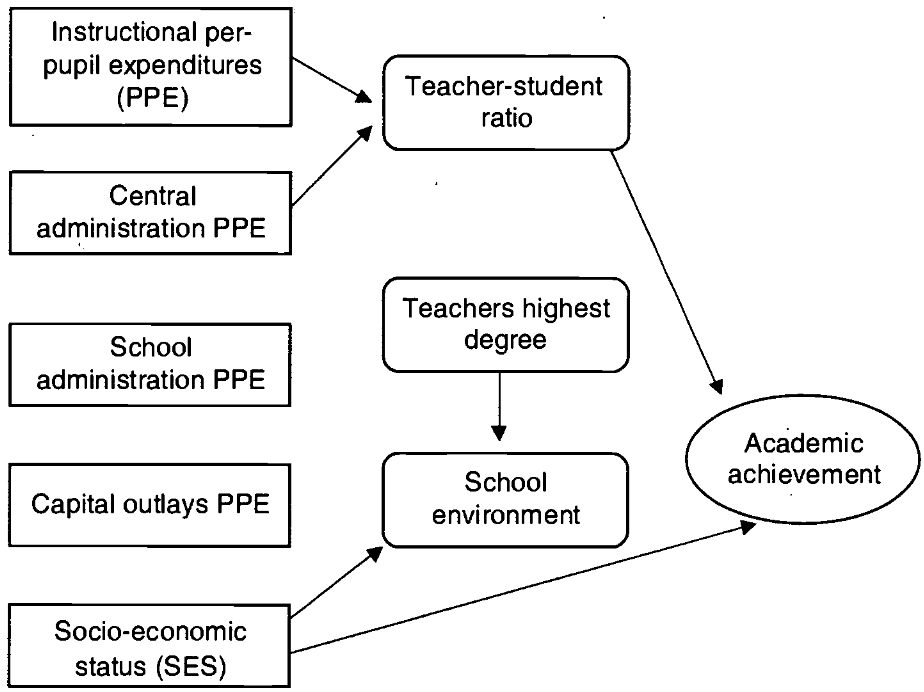
NOTE: Cells contain unstandardized parameters, standard errors, and standardized parameters.

SOURCE: Wenglinsky, unpublished tabulations.

istration expenditures do result in improved achievement. They positively affect teacher-student ratios, with standardized coefficients of 0.30 for instruction and 0.29 for central office administration. Teacher-student ratios, while not being associated with school environment as was expected, are directly associated with mathematics achievement, with a standardized coefficient of 0.11. On the other hand, school-level administration and capital outlays proved not to be related to school climate or mathematics achievement. Teacher's highest degree is weakly related to school environment (albeit in the counterintuitive direction), but school environment appears not to be related to mathematics achievement. Thus instructional expenditures, central office administration expenditures and teacher-student ratios appear to be part of the production function, while school-level administration, capital outlays and teacher's highest degree are not (see figure 2 for a schematic representation of results).[18]

[18] The analysis of eighth graders found the same three input variables to be components of the production function. It differed from the fourth grade analysis in that school climate mediated between the inputs and achievement. Instructional and central office administration expenditures were positively related to teacher-student ratios, which, rather than being directly related to achievement, were directly related to school climate. School climate, in turn, was related to achievement (Wenglinsky 1997).

Figure 2.—Paths to fourth grade achievements



SOURCE: Wenglinsky, unpublished.

To confirm this set of findings, goodness of fit was measured and compared to the goodness of fit of a model in which instructional and central office administration expenditures and teacher-student ratios were eliminated from the production function. In the full model, the chi-squares proved statistically insignificant, indicating good fit, with a mean chi-square of 25.67 across the five plausible values and a significance level of 0.06. The goodness-of-fit indices were also of sufficient size, with a mean adjusted goodness-of-fit index of 0.925 across the five plausible values and a mean normed goodness-of-fit index of 0.936. In the nested model, the chi-squares proved statistically significant, with a mean chi-square of 78.73 and a significance level better than 0.0001. The goodness-of-fit indices were of insufficient size, with a mean adjusted goodness-of-fit index of 0.817 and a mean normed goodness-of-fit index of 0.804. The goodness-of-fit measures, then, confirm that the model with the three production function components has an adequate fit and that an alternate model that excludes the components does not.

Estimates of the total effects of the production function components indicate that their effect on achievement can be substantial (table 3). The total effect of instructional per-pupil expenditures on mathematics achievement is statistically significant and amounts to 3.2 points of achievement for every \$4,000 dollars. The total effect of central office administration on mathematics achievement is 3.3 points for every \$500. Given that 12 points repre-

Table 3.—LISREL estimates of total effects

	Teacher-student ratio	High degree	School environment	Mathematics achievement
Instructional per-pupil expenditures (PPE)	0.0000051**	0.0000679	-0.0001452	0.0007987*
(thousands of dollars)	0.0000015	0.0000499	0.0000951	0.004659
	0.3000870	0.1278971	-0.0510623	0.0346460
Central administration PPE	0.0000368**	0.0000152	0.0016734	0.0065922*
(thousands of dollars)	0.0000108	0.0003787	0.0018442	0.0034844
	0.2850938	0.0037728	0.0776724	0.0376322
School administration PPE	—	—	0.0019241	0.0011203
(thousands of dollars)			0.0020389	0.0015197
			0.0804268	0.0057755
Capital outlays PPE	—	—	-0.0001847	-0.0001076
(thousands of dollars)			0.0003314	0.0002136
			-0.0472520	-0.0033932
Socio-economic status	—	—	0.3679910**	5.8440701**
(summated scale)			0.0778429	0.4059927
			0.4416251	0.6395601
Teacher-student ratio	—	-2.1375013	-15.8587921	139.5456070*
(number of teachers/students)		3.2277413	0.0000951	75.2240084
		0.1278971	-0.094645	0.1026882
Teacher's highest degree	—	—	-0.8180802*	1.1212598
			0.4473701	2.3910553
			-0.1526883	0.0258117
School environment	—	—	—	0.5822595
(summated scale)				0.4903571
				0.0718114
Mathematics achievement	—	—	—	—
(plausible values)				

— Relationship fixed at zero.

* p<.10

** p<.05

NOTE: Cells contain unstandardized parameters, standard errors, and standardized parameters.

SOURCE: Wenglinsky, unpublished tabulations.

sents a grade level, these effects are fairly substantial.[19] The effect of teacher-student ratios is still stronger. The total effect of teacher-student ratios on student achievement is 140 points for an increase of 1 teacher per student. Translated into class sizes, this means that a reduction in class size from 25 students to 15 students would result in an achievement gain of 14 points, well over a grade level.[20]

[19] It could be argued that a $4,000 increase in instructional expenditures is infeasible. Yet, a district would not need to raise all $4,000; some money could be obtained through the reallocation of existing funds. Thus, if a school district is currently spending $7,000, of which it allocates $3,000 for instruction, it could potentially increase spending on instruction by $4,000 by increasing aggregate expenditures by $2,000, to $9,000 and reallocating $2,000 of existing funds. It should also be noted that translating dollars into achievement assumes linearity, which may not be the case. It may be that only spending changes of a certain threshold translate into achievement changes. It may also be that only spending changes for school districts that begin at a certain level of expenditure result in achievement changes.

[20] It should not be surprising that the effect of teacher-student ratios is stronger than the effects of the two expenditure measures. To the extent that instructional and administrative dollars are spent on teacher-student ratios, they are conducive to academic achievement. Yet, not all instructional and administrative dollars are invested in ways that raise teacher-student ratios. Thus, while the most effective investment strategy to increase achievement would be to raise directly teacher-student ratios, where this is not feasible it is still possible to produce gains through allocating expenditures to the two areas known to raise these ratios.

Finally, the HLM analyses of the student-level database are consistent with the LISREL findings (table 4). As in the LISREL model, teacher-student ratios are significant related to achievement levels. The unstandardized coefficient is 153.8, as opposed to 152.1 for the LISREL model. Also, as in the LISREL model, socio-economic status is significantly related to achievement levels. The unstandardized coefficient is 6.01, as opposed to 5.63 in the LISREL model. Further, teacher's highest degree is not significantly related to achievement levels as in the LISREL model. It is also interesting to note that the only independent variable found to be significantly related to the SES-achievement relationship is district-level SES, suggesting that while resources can be associated with the level of achievement, they cannot affect its social distribution, at least for the population of fourth graders.

In sum, a series of structural equation models made it possible to identify some expenditures and resources that affect student achievement. Expenditures on instruction and central office administration affect teacher-student ratios, which, in turn affect student achievement. On the other hand, capital outlays, school-level administration and teacher education levels were found not to be associated with student achievement. These relationships persisted when subjected to multilevel analysis using HLM. It remains to discuss the implications of these results and the techniques employed to obtain them for the viability of the production function approach.

Conclusions

The models described here show that the key shortcomings of production functions can be addressed. First, the study was able to produce results that are national in scope. Since no single national database contains all of the variables needed for a production function, data were drawn from two universes and a sample and linked to one another. Second, the study distinguished between different types of expenditure. CCD made it possible to measure four types of expenditure, and the structural equation model made it possible to relate these to different parts of the learning process, such as school climate. This proved an important innovation because not all expenditures had an effect on achievement; those for the central office and instruction did, but those for capital and the principal's office did not. Third, the study took into account the role of school climate. NAEP provided a set of indicators of school climate that could be used to create a scale, and the structural equation model made it possible to measure both the influence of expenditures and resources on school climate and the influence of school climate on student achievement. In this study, however, the innovation proved of limited utility, since school climate was found not to play a mediating role in the production function. Fourth, the study measured student SES in a reasonably robust fashion, using a scale calculated from the measures provided by NAEP. Structural equation modeling made it possible to measure its influ-

Table 4.—Hierarchical linear modeling (HLM) unstandardized estimates of direct effects on intercept and slope of mathematics achievement

	Relationship to achievement intercept	Relationship to achievement slope
Teacher-student ratio	153.764687*	-5.686118
(number of teachers/students)	61.900057	28.600768
Teacher's highest degree	3.028100	0.826238
	2.533728	0.976926
Socio-economic status (summated scale)	6.008911*	0.444126*
	0.333620	0.139585

*p<.05

NOTE: Cells contain unstandardized parameters and standard errors.

SOURCE: Wenglinsky, unpublished tabulations.

ence on two variables, school climate and student achievement. This proved important because both relationships were significant. Fifth, the study adjusted the expenditure measures by the cost of education, using the TCI. This proved important as well, since the relationships would have been markedly different without these adjustments. Sixth, the study used a sophisticated achievement measure, drawn from NAEP, and applied it appropriately through adapting plausible values methodology to structural equation modeling. This innovation also proved important, as illustrated by the fact that many of the relationships which were found to be statistically insignificant would have appeared significant using the unadjusted mean of the plausible values. Even slight changes in the measurement of achievement can have significant effects on production function results. Finally, the study applied HLM to student-level data. This innovation actually proved unimportant; the HLM results did not differ substantially from the LISREL results.

Much more remains to be done, however. First, there were important differences in the findings from this study of fourth graders and a similar study of eighth graders. It therefore cannot be presumed that the production function for one grade level is the same for all; other grade levels should be studied. Second, many resource variables that might affect achievement were omitted from this analysis. The study used teacher education as a measure of teacher quality and found no relationship. Other measures need to be tested, however, before researchers arrive at the counterintuitive finding that teacher quality does not matter; for instance, teacher experience, teacher proficiency on standardized tests, and teachers having majored in the subject matter they are teaching, all may potentially influence student achievement. Finally, the current study uses cross-sectional data; meta-analyses (Hanushek 1997; Greenwald, Hedges, and Laine 1996) suggest that longitudinal data produces somewhat different findings. It is therefore important that a database be developed that tracks both inputs and outputs for a sample of students and schools over time.

References

Adams, Jacob E. 1994. "Spending School Reform Dollars in Kentucky: Familiar Patterns and New Programs, But Is This Reform?" *Educational Evaluation and Policy Analysis.* 16 (4): 375–390.

Austin, Gilbert R. and Herbert Garber (eds.). 1985. *Research on Exemplary Schools.* New York: Academic Press, Inc.

Barro, Stephen M. 1994. *Cost-of-Education Differentials Across the States.* Washington, D.C.: U.S. Department of Education.

Barton, Paul, Margaret Goertz, and Richard Coley. 1991. *The State of Inequality.* Princeton, NJ: Educational Testing Service.

Bentler, Peter M. and Douglas G. Bonnett. 1980. "Significance Tests and goodness of fit in the analysis of covariance structures." *Psychological Bulletin.* 88: 588–606.

Berends, Mark and Daniel M. Koretz. 1995. "Reporting Minority Students' Test Scores: How Well Can the National Assessment of Educational Progress Account for Differences in Social Context?" *Educational Assessment.* 3(3): 249–285.

Brookover, Wilbur, Charles Beady, Particia Flood, John Schweitzer, and Joe Wisenbaker. 1979. *School Social Systems and Student Achievement: Schools Can Make a Difference.* Brooklyn, NY: J.F. Bergin Publishers.

Bryk, Anthony S. and Stephen W. Raudenbush. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Newbury Park, CA: Sage Publications.

Bryk, Anthony S., Stephen W. Raudenbush, and Richard T. Congdon. 1996. *Hierarchical Linear and Nonlinear Modeling with HLM/2L and HLM/3L Programs.* Chicago, IL: Scientific Software International.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of Educational Opportunity.* Washington, D.C.: U.S. Government Printing Office.

Cohen, Jon. 1998. "Redesigning NAEP to Increase Its Usefulness." San Diego, CA: American Educational Research Association Annual Meeting.

Cuban, Larry. 1984. "Transforming the Frog into a Prince: Effective Schools Research, Policy and Practice at the District Level." *Harvard Educational Review.* 54: 129–151.

Edmonds, Ronald. 1979. "Effective Schools for the Urban Poor." *Educational Leadership.* 37 (1): 15–24.

Finn, Jeremy D. and Charles M. Achilles. 1990. "Answers and Questions about Class Size: A Statewide Experiment." *American Educational Research Journal.* 27 (3): 557–577.

Fortune, Jim C. and John S. O'Neil. 1994. "Production Function Analyses and the Study of Educational Funding Equity: A Methodological Critique." *Journal of Education Finance.* 20 (Summer): 21–46.

Glass, Gene V. and Mary Lee Smith. 1979. "Meta-analysis of Research on Class Size and Achievement." *Educational Evaluation and Policy Analysis.* 1 (1): 2–16.

Greenwald, Robert, Larry V. Hedges, and Richard D. Laine. 1996. "The Effect of School Resources on Student Achievement." *Review of Educational Research.* 66(3): 361–396.

Hambleton, Ronald K., H. Swaminathan, and H. Jane Rogers. 1991. *Fundamentals of Item Response Theory.* London: Sage Publications.

Hanushek, Eric A. 1989. "The Impact of Differential Expenditures on School Performance." *Educational Researcher.* 18 (4): 45–65.

Hanushek, Eric A. 1996. "School Resources and Student Performance." In Gary Burtless (ed.). *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success.* Pp. 43–73.

Hanushek, Eric A. 1997. "Assessing the Effects of School Resources on Student Performance: An Update." *Educational Evaluation and Policy Analysis.* 19 (2): 141–164.

Hauser, Robert M., William H. Sewell, and Duane F. Alwin. 1976. "High School Effects on Achievement." In William H. Sewell, Robert M. Hauser and David C. Featherman (eds.). *Schooling and Achievement in American Society.* Pp. 309–342. London: Academic Press.

Hayduk, Leslie A. 1987. *Structural Equation Modeling with LISREL: Essentials and Advances.* Baltimore, MD: Johns Hopkins University Press.

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

Jencks, Christopher, Marshall Smith, Henry Ackland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephan Michelson. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America.* New York: Basic Books.

Johnson, Eugene. 1989. "Considerations and Techniques for the Analysis of NAEP Data." *Journal of Educational Statistics.* 14 (4): 303–334.

Johnson, Eugene. 1994. "Overview of Part I: The Design and Implementation of the 1992 NAEP." In Eugene G. Johnson and James E. Carlson (eds.). *The NAEP 1992 Technical Report.* Pp. 9–32. Princeton, NJ: Educational Testing Service.

Johnson, Eugene, Robert J. Mislevy, and Neal Thomas. 1994. "Scaling Procedures." In Eugene G. Johnson and James E. Carlson (eds.). *The NAEP 1992 Technical Report.* Pp. 241–256. Princeton, NJ: Educational Testing Service.

Johnson, Eugene G., Keith Rust, and Carol Wallace. 1994. "Weighting Procedures and the Estimation of Sampling Variance." In Eugene G. Johnson and James E. Carlson (eds.). *The NAEP 1992 Technical Report.* Pp. 193–239. Princeton, NJ: Educational Testing Service.

Joreskog, Karl G. and Dag Sorbom. 1993. *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language.* Chicago: Scientific Software International.

Lee, Valerie E., Anthony S. Bryk, and Julia B. Smith. 1993. "The Organization of Effective Secondary Schools." *Review of Research in Education.* 19: 171–267.

Miles, Karen H. 1995. "Freeing Resources for Improving Schools: A Case Study of Teacher Allocation in Boston Public Schools." *Educational Evaluation and Policy Analysis.* 17 (4): 476–493.

Monk, David H. 1992. "Educational Productivity Research: An Update and Assessment of Its Role in Education Finance Reform." *Educational Evaluation and Policy Analysis.* 14: 307–332.

Mosteller, Frederick. 1995. "The Tennessee Study of Class Size in the Early School Grades." *The Future of Children: Critical Issues for Children and* Youths. 5(2): 113–127.

Mosteller, Frederick, Richard J. Light, and Jason A. Sachs. 1996. "Sustained Inquiry in Education: Lessons from Skill Grouping and Class Size." *Harvard Educational Review.* 66(4): 797–842.

Mullis, Ina V.S., John A. Dossey, Eugene H. Owen, and Gary W. Phillips. 1993. *NAEP 1992 Mathematics Report Card for the Nation and the States.* Princeton, NJ: Educational Testing Service.

National Center for Education Statistics. 1995a. *Disparities in Public School District Spending: 1989–90.* Washington, D.C.: Government Printing Office.

National Center for Education Statistics. 1995b. *Public School Teacher Cost Differences Across the United States.* Washington, D.C.: Government Printing Office.

National Center for Education Statistics. 1996. *Assessment and Analysis of School-Level Expenditures: Working Paper 96–19.* Government Printing Office: Washington D.C.

O'Reilly, Patricia E., Christine A. Zelenak, Alfred M. Rogers, and Debra L. Kline. 1996. *1994 Trial State Assessment Program in Reading Secondary-Use Data Files User Guide* Washington, D.C.: U.S. Department of Education.

Odden, Allan, and William H. Clune. 1995. "Improving Educational Productivity and School Finance." *Educational Researcher.* 24 (9): 6–10,22.

Odden, Allan. 1990. "Class Size and Student Achievement: Research-Based Policy Alternatives." *Educational Evaluation and Policy Analysis.* 12 (2): 213–227.

Purkey, Stewart C. and Marshall S. Smith. 1983. "Effective Schools: A Review." *Elementary School Journal.* 83: 427–452.

Raudenbush, Stephen W. and J. Douglas Willms. 1995. "The Estimation of School Effects." *Journal of Educational and Behavioral Statistics.* 20 (4): 307–355.

Wenglinsky, Harold H. 1997. "How Money Matters: Models of the Effect of School District Spending on Academic Achievement." *Sociology of Education.* 70 (3).

Appendix A: Variable Definitions

Capital Outlays Per-pupil Expenditures: Derived from data in CCD for Fiscal Year 1992. Calculated by dividing total capital outlays, as defined in CCD, for each school district by the number of students in the school district and the Teacher Cost Index. Measured in thousands of dollars.

Central Administration Per-pupil Expenditures: Derived from data in CCD for Fiscal Year 1992. Calculated by dividing total expenditures on central administration, as defined in CCD, for each school district by the number of students in the school district and the Teacher Cost Index (TCI). Measured in thousands of dollars.

Highest Degree: Taken from NAEP data for mathematics for 1992. Consists of the highest level of education attained by teacher responding to NAEP on behalf of a student. Responses were coded "1" for less than a Bachelor's degree, "2" for a Bachelor's degree, "3" for a Master's degree and "4" for more than a Master's degree.

Instructional Per-pupil Expenditures: Derived from data in CCD for Fiscal Year 1992. Calculated by dividing total expenditures on instruction, as defined in CCD, for each school district by the number of students in the school district and the Teacher Cost Index. Measured in thousands of dollars.

Mathematics Achievement: Taken from NAEP data for mathematics for 1992. Consists of the five plausible values for students responding to NAEP. Means and standard deviations presented in this paper are means of these statistics for the five plausible values. For all maximum likelihood estimates, plausible values were analyzed in accordance with plausible values methodology. Measured on common proficiency scale for all grades (fourth, eighth and twelfth).

School Administration Per-pupil Expenditures: Derived from data in CCD for Fiscal Year 1992. Calculated by dividing total expenditures on school-level administration, as defined in CCD, for each school district by the number of students in the school district and the Teacher Cost Index (TCI). Measured in thousands of dollars.

School Environment: Derived from NAEP data for mathematics for 1992. Calculated as summated scale of the following items: for each school in NAEP the degree to which teacher absenteeism is not a problem; the degree to which student tardiness is not a problem; the degree to which student absenteeism is not a problem; the degree to which class cutting is not a problem; and the degree to which there is a regard for school property; for each teacher in NAEP the degree to which teachers have control over instruction; and the degree to which teachers have control over course content. Measured as total of that scale.

Socio-economic Status (SES): Derived from NAEP data for Mathematics for 1992. Calculated as summated scale of the following items: for each student whether or not family receives newspaper; whether or not there is an encyclopedia in the home; whether or not there are more than 25 books in the home; whether or not the family subscribes to magazines; the highest level of education attained by the mother; the highest level of education attained by the father; and for each school in NAEP the percentage of students who receive reduced price or free lunches. Measured as total of that scale.

Teacher-Student Ratio: Derived from NAEP data for mathematics for 1992. Calculated by dividing total number of teachers in school by total number of students in school.

The Development of School Finance Formulas to Guarantee the Provision of Adequate Education to Low-Income Students

Andrew Reschovsky

Jennifer Imazeki

University of Wisconsin-Madison

About the Authors

Andrew Reschovsky is a professor in the Robert M. La Follette Institute of Public Affairs and the Department of Agricultural and Applied Economics at the University of Wisconsin-Madison, where he teaches public finance, microeconomic analysis, and urban economics. He has a Ph.D. in Economics form the University of Pennsylvania, and has held previous academic positions at Rutgers University and Tufts University. He has also served as a consultant to several federal agencies and to a number of state and local governments throughout the country. During 1994 and 1995 he served as an economist in the Office of Tax Analysis of the U.S. Department of the Treasury. His research focuses on tax policy and on state and local government finance. He has written numerous articles in professional journals and has contributed articles to a number of books. In addition to his work on education fiance, he is currently conducting research on state government fiscal responses to block grants for welfare and on the possibilities of converting the mortgage interest credit into a tax deduction.

Jennifer Imazeki is a research assistant with the Consortium for Policy Research in Education, and a Ph.D. candidate in Economics at the University of Wisconsin-Madison. Her dissertation research focuses on education production and the teacher labor market.

The Development of School Finance Formulas to Guarantee the Provision of Adequate Education to Low-Income Students

Andrew Reschovsky

Jennifer Imazeki

University of Wisconsin-Madison

Introduction

In a series of legal actions starting with the landmark ruling by the California Supreme Court in *Serrano v. Priest* (1971), state courts have grappled with the problem of inequities in the financing of schools. Spurred by these court decisions, a majority of state legislatures increased the level of state funding for education, and adopted formulas for the distribution of school aid which were designed to increase the equity in school finance.[1] In particular, some states attempted to equalize per pupil spending across school districts. Other states attempted to guarantee that property-poor school districts would be able to achieve a given level of spending per pupil as long as they levied a standard property tax rate. Still other states attempted to guarantee that all school districts that chose the same property tax rate would be able to spend the same amount of money per pupil regardless of district property wealth.

The focus of most of these attempts to reduce inequities in school finance has been on the distribution of dollars. The implicit (and sometimes explicit) assumption behind these efforts is that a more equal distribution of fiscal resources will lead to increased equity in educational opportunities and in educational outcomes. There continues to be a debate, however, about the strength of the relationship between spending on education and educational outcomes; some scholars, notably Eric Hanushek (1989, 1997), argue that no consistent relationship exists between spending and educational outcomes, while others (e.g., Hedges, Laine, and Greenwald, 1994) challenge Hanushek's conclusions.

Even if it can be shown conclusively that spending money on public education results in substantial improvements in student performance, it is important to recognize that there is not a one-to-one relationship between spending and educational outcomes. A comparison of two districts with equal spending per pupil reveals that educational performance may be lower in one of the districts if the costs of providing any given level of education are higher in that district, or if that district is more inefficient in its use of resources.

[1] For a discussion of alternative definitions of equity in school finance see Berne and Stiefel (1984) and Reschovsky (1994).

The *cost* of education can be defined as the minimum amount of money that a school district must spend in order to achieve a given educational outcome, such as reading at a third-grade level at the end of the third grade. Costs differ across school districts for reasons that are outside the control of local school boards, such as the number of children with "special needs" or factors that increase the amount of money needed to attract good teachers, such as the area cost-of-living. Although actual expenditures are influenced by the costs districts face, they also reflect choices made by local school boards concerning the type and amount of education they provide and the ways they choose to allocate and organize resources used in achieving their educational objectives. Thus, a school district with below-average costs could have above-average expenditures because it chooses to provide its students with the opportunity to take a particularly wide range of advanced courses, or because it is relatively inefficient in its use of resources.

> . . . equality of education, . . . cannot be achieved unless explicit account is taken of the higher costs . . . associated with educating children . . . from poor or otherwise disadvantaged backgrounds.

The importance of costs in any discussion of equity in school finance is that as long as equity is defined in terms of equal educational outcomes, the achievement of equity will require higher spending in districts facing higher costs. Conversely, equal per pupil spending across districts will not result in equal educational outcomes as long as some districts face higher costs than other districts.

Over the past decade, a number of state courts have begun to recognize the important role cost differences play in the design of policies for achieving equity goals. The courts have realized that equal per pupil spending or equal tax effort do not guarantee equal educational outcomes. This has led them to address issues of student performance more directly, by recognizing that equality of education, however defined, cannot be achieved unless explicit account is taken of the higher costs that are generally associated with educating children who come from poor or otherwise disadvantaged backgrounds. As William Clune (1994) has argued, the courts are moving from a focus on equity in spending to one of educational adequacy, with adequacy defined in terms of minimum standards of student performance. The courts appear to be arguing that states are responsible for assuring that all school districts provide an adequate level of education. A prerequisite for designing a school finance system that is capable of achieving this goal is knowledge of how much it will cost each school district to provide an adequate education for its students.

In Kentucky, the court ruled that the state constitution required the state do more to raise the level of student achievement in poor school districts (*Rose v. Council for Better Education, Inc., 1989*). In both New Jersey and Texas, state courts concluded that the state legislatures' responses to previous court cases had been inadequate, and further efforts must be made to allocate more resources to poor districts plagued by low student achievement. In a Massachusetts decision (*McDuffy v. Secretary of Education, 1993*), the state's Supreme Court specified seven specific "capabilities" that an educated child must possess. In effect, the court ruled that the state must develop a system of school finance which guarantees that all children be provided with an adequate education, which is defined in terms of a specified set of skills.

The establishment of a school financing system that guarantees all students an adequate education requires that we be able to measure the costs of providing an adequate education in each school district. The purpose of this paper is to estimate a cost function for K–12 education and, using data from the state of Wisconsin, demonstrate how these cost estimates can be integrated into state aid formulas in a way that is consistent with the achievement of educational adequacy.[2]

[2] For a detailed discussion of how costs can be integrated into aid formulas designed to achieve various educational equity goals, see Ladd and Yinger (1994).

We start with a brief discussion of the limited ways in which cost considerations have been included in school aid distribution formulas. We also review several previous efforts to develop cost measures. We then discuss the methodological approach used in estimating public education costs and describe the data used in the analysis. Our estimated cost functions are presented next, followed by a discussion of developing a cost index for school districts in Wisconsin based on our estimated cost function. We demonstrate that costs vary substantially among Wisconsin's 368 school districts.[3] Next, we develop a school aid formula designed to achieve education adequacy and then simulate the distribution of aid for the academic year 1997–98 using both this formula and a conventional foundation formula. In a number of states, cost factors are introduced into school aid formulas by "weighting" poor or disabled students more heavily than "regular" students. We then use our analysis of costs to define an appropriate weight for poor children. Finally, we summarize our results and draw some conclusions.

Accounting for Costs in the Distribution of Education Aid

State government grants financed about 45 percent of total spending on elementary and secondary education in 1993–94, the latest year for which we have data (National Center for Education Statistics, 1997). Most of these grants were distributed using foundation or guaranteed tax base formulas. Foundation formulas are designed to equalize per-pupil expenditures. Guaranteed tax base or district power equalizing formulas are intended to equalize the tax rates necessary to provide any given level of per pupil spending. In most states, neither of these formulas explicitly take into account inter-district differences in costs.

> *. . . equalization aid formulas generally do not include adjustments for cost differences . . .*

Although equalization aid formulas generally do not include adjustments for cost differences, state governments do provide categorical aid to local districts for the education of certain disabled or "special-education" students. In fact, federal legislation, the *Individuals with Disabilities Education Act,* requires that school districts provide all children with physical, mental, or emotional disabilities with a public education "...in the least restrictive environment appropriate for their educational progress... (p. 346)" (Chaikind, Danielson, and Brauen, 1993). Although the federal government financed nearly $4 billion in special education grants in fiscal year 1997, these funds accounted for only a small portion of total expenditures by local school districts on special education programs.

Some states help to finance the education of these special-needs students by giving them a heavier weight in equalization aid formulas. For example, by assigning a weight of 2.3 to each "disabled" student attending public schools, the state signals that it believes that the per pupil cost of educating these students is 2.3 times the cost of providing education to "regular" students. A number of states also assign extra weight to students from economically disadvantaged families.

As described in detail by Chaikind, Danielson, and Brauen (1993), estimates of the cost of educating disabled and other "special education" students come primarily from a limited number of detailed surveys of special education programs in small samples of school districts.[4] These surveys provide a detailed accounting of the resources expended to educate special education students. It should be noted, however, that tabulating spending on special education is inherently difficult, particularly when some special education students

[3] We have excluded Norris from our analysis. Due to an historical anomaly, Norris is officially a K–12 school district, but it is in fact a private "school for wayward boys" with a 1996–97 enrollment of about 75 students and a per pupil property tax base that is less than three percent of the state average.

[4] The national expenditure survey discussed by Chaikind, Danielson, and Brauen (1993) samples 60 school districts around the country.

divide their time in school between regular classrooms and separate special education classes. Furthermore, as both federal and state special education grants are allocated on the basis of the number of students classified as eligible for special education and the spending on special education programs, school districts may have strong incentives both to declare as many students as possible eligible for special education and to attribute to special education as much general spending as possible.

Even if special education spending data are accurate, they do not necessarily provide full information on the costs of special education. Spending data tell us how much money districts allocate to special education, but provide no information on the services actually provided to special education students. Because some states excuse students in special education programs from taking standardized tests, it is particularly difficult to assess how effective schools are at educating special-education students. Unless an effort is made to account for differences across school districts in the level and quality of special education actually provided, the use of "weights" for special education pupils in school finance formulas may either over or under count the true costs of educating special education students.[5]

Even if special education spending data are accurate, they do not necessarily provide full information on the costs of special education.

Although most state aid formulas do not account in any systematic way for differences in costs, several cost indices have been developed that could be used in school finance formulas to adjust for differences in costs. One approach, followed by Walter McMahon (1991, 1994), has been to estimate cost-of-living indices for school districts. A second approach, primarily associated with Jay Chambers (1981, 1995), has been to estimate hedonic wage equations for teachers and use the results to compute a teacher salary index or, more broadly, a cost of education index for individual school districts.

Although both of these approaches provide valuable information about differences in the costs of providing education, they go only part of the way towards the goal of providing a comprehensive cost index that can be used in school aid formulas. By definition, the concept of costs links school district spending to school performance. For reasons outside the control of local school officials, districts with higher costs must spend more to provide any given level of educational services than districts with lower costs. The higher salaries that school districts in high cost-of-living areas have to pay to attract teachers is only one reason why costs may be high. For example, depending on the composition of their student bodies, some districts may have to provide special programs and hire additional employees in order to achieve the same educational outcomes that other districts can provide without special programs or extra employees. For this reason, cost-of-living indices provide an inadequate basis for making cost adjustments to school aid formulas. Furthermore, as pointed out by Chambers (1995), the extent to which an area's high cost of living reflects attributes of a given location that teachers find attractive, cost-of-living indices may overstate the true costs of hiring teachers in attractive locations.[6]

The teacher salary indices developed by Chambers provide a more direct measure of school district costs than cost-of-living indices. Chambers estimates hedonic wage equations in an attempt to isolate those factors outside the control of local school districts that require some districts to pay higher salaries than others in order to employ teachers with similar qualifications to carry out similar teaching assignments. In his re-

[5] Duncombe, Ruggiero, and Yinger (1995) provide evidence that school aid formulas in New York state that include "weighted pupils" are likely to under-adjust for cost differences, and in some cases may actually magnify, rather than reduce, the underlying cost differences among school districts.

[6] McMahon (1994) recognizes that cost-of-living indices will reflect locational amenities. He suggests, however, that *ad hoc* adjustments can be made to cost-of-living indices to adjust for the presence of amenities (and presumably, locational disamenities).

cent report to the National Center for Education Statistics, Chambers (1995) identifies as cost factors the racial and ethnic composition of the student body, land costs, pupil-teacher ratios, and a range of variables that influence the attractiveness of any given geographical area, such as weather conditions and crime rates. Using the coefficients of the cost factors, Chambers constructs two teacher salary indices, one that varies by county and the other by school district.

Although these indices provide useful information about costs, they almost certainly understate the contribution to costs of various school districts' characteristics. To develop a comprehensive measure of costs, one must not only account for differences across districts in the cost of hiring teachers of a given quality to carry out a given assignment, but one must also account for the fact that some districts will have to hire more teachers and incur more non-teacher expenditures (e.g., on textbooks, social workers) in order to achieve any specific educational goal.

. . . a comprehensive measure of costs . . . must not only account for differences . . . in the cost of hiring teachers . . . but . . . must also account for the fact that some districts will have to hire more teachers.

Economists generally define costs as the value of resources needed to produce any given level of output. The typical cost-of-living or cost-of-education index is designed to measure the dollar cost of purchasing a given set of inputs to be used in the production of education. While indices of this type provide useful information, they fail to provide a comprehensive measure of costs. In recent years there have been several attempts by economists to develop more comprehensive measures of costs. These studies have been motivated by a desire to develop a straightforward way to account for cost differences in state government grant formulas to local school districts. They include studies of school districts in Nebraska (Ratcliff, Riddle, and Yinger, 1990), in Arizona (Downes and Pogue, 1994), in New York (Duncombe, Ruggiero, and Yinger, 1996), and in Michigan (Courant, Gramlich, and Loeb, 1994).[7]

In each of these studies, the authors find that costs varied substantially among school districts. The studies identify a number of local school districts' characteristics that influence the cost of public education. For example, Ratcliff, Riddle, and Yinger found five factors that both influence costs and lie largely outside the control of local public officials: the number of handicapped students, the number of students that the school district is required to transport, secondary school students as a proportion of a district's total enrollment, and the size and type of school districts. Downes and Pogue show that school "maintenance and operations" costs per student in Arizona are related to the ethnic composition of the student body, the incidence of poverty, the proportion of students with limited English proficiency, and school size. Cost factors identified by Duncombe, Ruggiero, and Yinger include district size, the percentage of children living in poor families, the percentage living in female-headed households, the percentage of students with disabilities and the percentage with limited English proficiency.

The above-mentioned studies have followed one of two approaches in estimating the costs of public education. One approach is to estimate cost functions directly for public education. By definition a cost function provides a measure of the value of total resources necessary to produce any given level of output or performance. Thus, the use of this approach requires that one be able to develop measures of public school output. Difficulties in measuring public sector outputs have led some researchers to attempt to identify costs through the estimation of reduced-form public education expenditure functions. Although the estimation of an expenditure function does not require the use of educational output measures, it is likely to lead to underestimates of the impact of various cost factors on education spending. These underestimates are likely to occur because in a reduced-form regression it is impossible to separate

[7] Paralleling these studies of the costs of education, several recent studies have attempted to measure the costs associated with the provision of municipal government services. These include studies of local government in Massachusetts (Bradbury, et al., 1984), in Minnesota (Ladd, Reschovsky, and Yinger, 1992), and in Wisconsin (Green and Reschovsky, 1994).

the impact of cost factors on the demand for public education from their direct impact on costs. Downes and Pogue (1994) provide a detailed discussion of the strengths and weaknesses of each approach, and then proceed to estimate the costs of public education in Arizona using both approaches.

In this paper, we attempt to estimate cost functions directly for the provision of K–12 public education in Wisconsin. As we will explain in detail in the next section, we pursue a methodological approach that is very similar to that used by Duncombe, Ruggiero, and Yinger.

Methodology

Following Bradford, Malt, and Oates (1969) and Inman (1979), it is useful to think of public school output as a function of school resources, such as teachers and textbooks, the characteristics of the students, and the family and neighborhood environment in which the students live. This relationship is represented by equation (1), in which S_{it} represents an index of school output, X_{it} is a vector of direct school inputs, Z_{it} is a vector of student characteristics, and F_{it} is a vector of family and neighborhood characteristics. The subscript i refers to the school district and subscript t refers to the year.

Of critical importance in estimating an education cost function is the accurate measurement of school district output . . .

(1) $S_{it} = g(X_{it}, Z_{it}, F_{it})$

To move from this education production function to a cost function, we must specify the relationship between school inputs and educational spending. Equation (2) indicates that per-pupil expenditures, E_{it}, are a function of school inputs, X_{it}, a vector of input prices, P_{it}, and ϵ_{it}, a vector of unobserved characteristics of the school district.

(2) $E_{it} = f(X_{it}, P_{it}, \epsilon_{it})$

Since cost functions are defined as the spending necessary to provide any given level of output, we can formulate a cost function for public education by solving equation (1) for X_{it}, and then substituting X_{it} into equation (2). The results is represented by equation (3), in which u_{it} is a random error term.

(3) $E_{it} = h(S_{it}, P_{it}, Z_{it}, F_{it}, \epsilon_{it}, u_{it})$

In the "Results" section, we present estimates of equation (3) using 1994–95 data for K–12 school districts in Wisconsin. In the remainder of this section, we discuss a number of methodological and data issues that must be addressed in order to carry out these estimates. Table 1 displays descriptive statistics of the variables we use in our analysis.

Of critical importance in estimating an education cost function is the accurate measurement of school district output, S_{it}. The vast literature on educational production function tends to focus on student cognitive achievement as measured by standardized test scores. A commonly used measure of school output or performance is average test scores from achievement tests administered to all students. It seems reasonable, however, to assume that most voters, including parents, judge the effectiveness of schools by their ability to generate annual improvements in test scores. Robert Meyer (1996) demonstrates that average test scores alone provide a highly flawed measure of school output. He points out that average achievement on a test administered to tenth grade students, for example, measures the average level of achievement prior to entering first grade, plus the average effects of school performance, and of family, neighborhood, and student characteristics on the growth of student achievement from the first through the tenth grade. It is thus likely that rather than providing a measure of the contribution of schools to the growth in student achievement, the tenth grade score primarily reflects the impact of family and neighborhood environment on student achievement.

Table 1.—Descriptive statistics

Variables	Mean	Standard deviation	Minimum value	Maximum value
Per-pupil expenditures, 1994–95	$6,327	$759	$4,690	$9,053
Tenth grade exam score, 1995–96	75.9	7.2	29	94
Eighth grade exam score, 1993–94	73.6	7.5	34	94
Number of advanced courses	17.9	19.2	3	227
Teacher salary index	1	0.1	0.8	1.3
Percent of students eligible for free and reduced-price lunch	23.1	13.4	0	79.8
Percent of students with disabilities	12.2	3	0.8	43.9
Percent of students with severe disabilities	0.2	0.2	0	1.1
Percent of students enrolled in high school	32.4	3.4	22.6	45.9
Student enrollment	2,192	5,676	109	97,555
Median income	$27,821	$7,184	$14,122	$56,859
Residential tax base/Total tax base	58.7	15.1	14.8	98.6
Tax price	0.6	0.3	0.2	1.5
Categorical state aid	$289.2	$126.2	$48.9	$772.3
Percent of households with children	35.3	4.9	18.6	54.5
Percent homeowners	75.6	7.1	44.7	90.8
Percent elderly	14.8	3.9	5.6	28.3
Percent with 4-year college degree	12.5	6.9	3.4	60.1

SOURCE: U.S. Bureau of the Census (with data from the 1990 Census of Population and Housing) and the Wisconsin Department of Public Instruction.

For a more accurate measure of school output (at least the portion of output measured by increases in cognitive skills), it is important to use a "value-added" measure of pupil achievement. By focusing on the changes in test scores over time, this type of school output or performance measure isolates the contribution of school resources to increases in student achievement as measured by scores on standardized achievement tests. Meyer points out that because student mobility among school districts tends to be quite high, the construction of value-added measures of school output should be based on tests of the same students at regular intervals, preferably annually.

Although we do not have annual data, we are able to construct a value-added measure of student achievement in Wisconsin schools using biannual test scores. In the 1993–94 academic year, Wisconsin began to require that all students take standardized exams during the eighth and tenth grades and that the test results be reported to the Wisconsin Department of Public Instruction. Thus, we can construct a value-added measure for students who were eighth graders in 1993–94 and tenth graders two years later, in 1995–96.

While standardized tests are targeted to specific knowledge about core subject areas (in Wisconsin, these are reading, mathematics, language, general science, and social studies), another measure of the quality of schools is the breadth of the course offerings. The education a child receives will be enriched if the child is exposed to a wide range of subjects above and beyond the core subject areas. One measure of richness of the course offerings is the number of advanced courses offered. Data on the number of advanced courses offered provide a measure of the opportunities available to students. Although no direct information on the actual number of students enrolled in these courses is available, the fact that few school districts can afford to continue to offer specialized courses unless the courses have reasonable enrollments, suggests that the use of data on the number of courses provides a good measure of the actual richness and diversity of the school curriculum benefiting students.[8]

In the vector of school input prices, salaries are a crucial component. Teacher salaries account for the largest share of school expenditures. In our estimate of an education cost function, we include only teacher salaries, excluding explicit treatment of other public school employees. It is important to recognize that teacher payrolls are determined both by factors under the control of local school boards, and factors that are largely outside of their control. In setting hiring policies, districts make decisions about the quality of teachers they will recruit and these decisions have obvious fiscal implications. For example, a district can limit its search for new teachers to those with advanced degrees, to those with high grade point averages, or to those with a certain number of courses in their teaching specialty. Teacher salary levels are generally determined through a process of negotiation with teacher unions, and school boards have a substantial impact on the outcome of these negotiations.

At the same time, the composition of the student body, working conditions within schools, and area cost of living play a potentially large role in determining the salary a school district must offer in order to attract teachers of any given quality. These factors will be reflected in student and district cost variables, to be described below.

In estimating the cost index, we would like a measure of teacher salaries that only reflects differences in salaries that are outside the control of local school districts. One possibility is to use the Chambers teacher cost index, discussed in the section on "Accounting for Costs in the Distribution of Education Aid." However, because we have access to detailed information on individual teacher characteristics, we chose to construct our own index of teacher salaries. To construct this index, we use data collected by Wisconsin's Department of Public Instruction on the salary and fringes, education, and experience of every public school teacher in the state. We regress the log of the sum of salary and fringes for all full-time teachers on each teacher's background characteristics (including years of teaching experience and highest degree earned) plus a dummy variable for each school district. The coefficients on the district dummies are then used as the values of the teacher salary index. That is, the teacher salary index represents differences in salaries across districts, holding teacher background constant.[9] As explained below, we treat the teacher salary index as endogenous when estimating the cost function.[10]

The vectors of student, family, and neighborhood characteristics, Z_{it} and F_{it}, are made up of several variables that we believe influence a district's level of spending per pupil. First, enrollment and enrollment squared are included. The literature suggests that a U-shaped relationship exists between

[8] Another measure of school district output that could be included is the ability of the school system to prevent dropouts. Thus a school district will be more effective to the extent that it can minimize its dropout rate. Unfortunately, because enrollment numbers are collected only once each academic year, accurate estimates of dropout rates are difficult to calculate. A comprehensive list of school performance measures should include a measure of each district's success in educating students with mental, physical, and learning disabilities. These performance measures are particularly important as our test score data exclude the performance of most special education students. Unfortunately, Wisconsin does not compile comprehensive data on the performance of special education students.

[9] Although our construction of the teacher cost index is similar, in spirit, to the methods used by Chambers (1995), our index differs from Chambers index in several important ways. First, Chambers' index numbers for Wisconsin are based on parameter estimates from a national sample while our index is based solely on Wisconsin data. As our objective is to analyze school costs within one state, it seems appropriate to use parameter estimates that are specific to the state. Moreover, we have data for the full population of Wisconsin public school teachers, rather than just a sample, increasing confidence in the estimates. Because we have data for the state population, our index is based on deviations of actual salary values from the average, rather than hedonic, predicted salaries. Because we use district dummies, our index also differs from Chambers in that it captures salary differentials that may be based on immeasurable factors. Chambers calculates cost differentials based on differences in measurable factors alone. If there are district-specific factors that affect salary differentials and that are left out of Chambers' list of exogenous variables, then, relative to our index, his index will understate inter-district differentials.

[10] The endogeneity of the teacher salary index reflects the fact that while higher teacher salaries lead to higher per-pupil expenditures, decisions by school districts to raise spending are likely to result in higher teacher salaries.

per-pupil expenditure and district size (both measured in logs), reflecting diseconomies of scale associated with both very small and very large school districts. Next, we include a measure of economic disadvantage. The evidence from previous studies (cited earlier) indicates that there are higher costs associated with the education of children from low-income families. In our estimation procedure, we use the percentage of students who qualify for the federal government-financed free and reduced-price lunch program as a measure of the share of students coming from economically disadvantaged families. There is also substantial literature that documents the extra costs associated with educating students with various kinds of disabilities.[11] Therefore, we include two measures of disabled students. One is the percentage of students who are classified with any type of disability. The other is the percentage of students who are classified as autistic, deaf, or blind. Studies have shown that the education costs for students with these disabilities is far greater than the extra costs associated with educating students with other disabilities (Chaikind, Danielson, and Brauen, 1993).

The evidence from previous studies . . . indicates that there are higher costs associated with the education of children from low-income families.

To reflect the possibility that educating high school students requires more resources than educating elementary school students, we also include as a cost factor the proportion of a district's student body that is enrolled in high school.

The variable ϵ_{it} in equation (3) represents the unobserved factors in each school district that influence school district spending. One such factor that has received much attention is the "inefficiency" of a district; that is, the extent to which spending in a district is in excess of the amount necessary to obtain its chosen level of educational output. A number of recent papers have applied various methods of frontier analysis in an attempt to systematically measure this inefficiency for each district (Bessent and Bessent, 1990; Deller and Rudnicki, 1993; Duncombe, Ruggiero, and Yinger, 1996; McCarty and Yaisawarng, 1993; and Ruggiero, 1996). These authors have used these techniques to gauge the amount of inefficiency involved in the provision of public education.

Great care, however, must be taken before one interprets the results of such analysis as providing evidence about the inefficiency of public schools. This is because the standard measure of "inefficiency" that arises from applying frontier analysis captures the effect of all factors that lead spending to be higher than the minimum cost of providing any given mix of public school output. Thus higher spending in one school district that is attributable to the higher costs of educating an above-average share of economically disadvantaged students will, at least in part, be characterized as "inefficiency." As pointed out by Duncombe, Ruggiero, and Yinger (1996), the fact that these higher costs will be attributed in part to the efficiency measure and in part to the cost factors included in equation (3) means that the cost function estimates provide an underestimate of the full effects of the cost factors on education spending.

The correct interpretation of these efficiency measures also requires that we have adequately measured public school output. If our cost function fails to include a school output measure that is important to local residents, any expenditures attributable to achieving that object will be classified as due to inefficiency rather than to higher costs. For example, in many states students eligible for special education classes are not required (or allowed) to take standardized achievement tests. For this reason, if one school district devotes extra resources to bringing children enrolled in special education classes up to their grade level in reading, while another district

[11] A number of other studies have found that educating students who enter school with a limited knowledge of English results in higher costs. However, when we included the percentage of students enrolled in English as a Second Language programs in our cost functions, it was always statistically insignificant and had a negative sign. Thus, we do not include this variable in the estimates presented in this paper.

provides only limited resources to special education students, a standard frontier analysis is likely to characterize the first district as "inefficient" relative to the second district.

Because of these complexities, our analysis here does not include a measure of efficiency. Plans for future research include using both parametric and nonparametric frontier analysis techniques to estimate "efficiency" measures for inclusion in the estimated cost functions. Results from previous studies (particularly Duncombe, Ruggiero, and Yinger, 1996) suggest that the exclusion of an efficiency measure will not strongly affect the relative ranking of the estimated cost index, but will decrease variation and range.

Results

The cost function represented by equation (3) is estimated using two-stage least squares, with the school output variables and the teacher salary index treated as endogenous. Following much of the literature on education costs, the cost function is estimated in natural logs, the dependent variable being the natural log of total operating expenses per pupil during the 1994–95 school year.

. . . we model the demand for public education as a function of . . . residents' preferences for education, their incomes, the tax prices they face . . . , and the intergovernmental aid their school district receives.

We use three measures of education outcome: the district average battery score from the Tenth Grade Knowledge and Concepts Exam, administered in the fall of the 1995–96 academic year; the district average battery score from the Eighth Grade Knowledge and Concepts Exam, administered in 1993–94; and a count of the number of advanced courses offered by each school district during the 1994–95 school year (all in logs). The battery scores are themselves an average of the national percentile rank on multiple-choice exams in reading, math, language, science and social science. Because the exams are administered in October, the tenth-grade scores are actually reflections of knowledge acquired prior to the tenth grade; therefore, we use 1995–96 scores with 1994–95 spending data. The eighth-grade scores are included as a control for past achievement, thus isolating the relationship between spending and *growth* in achievement between the eighth and tenth grades.

As noted in equation (1), school output is, in part, a choice that reflects the "tastes" of the community. The decision about the mix and level of output is made in conjunction with the decision about how much to spend. We therefore treat the school output variables as endogenous. As instruments for these output measures, we draw upon a set of variables that are related to the demand for public education. Following a large literature on the determinants of local government spending, we model the demand for public education as a function of school district residents' preferences for education, their incomes, the tax prices they face for education spending, and the intergovernmental aid their school district receives. To the extent that the median voter model provides a reasonable explanation for school district spending decisions, it is appropriate to use median income and the tax price faced by the median voter as instruments. Since most state school aid in Wisconsin is distributed through a matching grant formula, we use the tax price implied by the aid formula. It should be noted that because Wisconsin distributes aid through a complex three-tier district power equalizing (DPE) formula, some districts, particularly those with modest property wealth and above average spending, may face a tax price that is greater than one because for every dollar of additional spending, the size of their grant is reduced. We also include categorical aid received by the district as another instrument. The ratio of the residential property tax base as a proportion of the total tax base serves as a rough measure of the district's ability to export the tax burden to commercial and industrial properties. Finally, we include as instruments several socioeconomic variables that may be related to the preferences for public education. These include the percentage of households with children, the percentage of household heads who are homeowners, the percentage of the population age 65 or older, and the percentage of adults who have earned a four-year college degree.

As instruments for the salary index variable, we include a set of variables that reflect differences in the cost of living in various parts of the state. Deller, Green, and Voss (1996) have divided Wisconsin's 72 counties into five cost-of-living groups based on median household incomes, median housing values and rents. As instruments, we use dummy variables reflecting the assignment of each school district to its appropriate cost-of-living group.

Over 50 of Wisconsin's K–12 school districts are very small, with fewer than 500 students. At the other extreme of the size distribution is Milwaukee with nearly 100,000 students; an enrollment that is four times greater than the next largest school district in the state (Madison). Fitting a regression in which every district is treated equally may mask the true relationship between per pupil spending and the covariates. To account for this, we weight the regressions by district membership.

Table 2 presents our estimates of a cost function for public education in Wisconsin for the 1994–95 academic year. The test scores have the expected signs; since eighth grade scores are a proxy for past levels of students' achievement, high scores mean that districts can spend less to achieve a given level of progress. The negative coefficient for the number of advanced courses is counter-intuitive; one would expect spending to be higher in districts offering a wide range of advanced courses. However, the negative coefficient may reflect, in part, economies of scale as the number of advanced courses offered is highly correlated with enrollment.

Determining a level of educational output that is considered adequate for each state is obviously a public policy decision.

The cost variables generally have the expected signs and most are statistically significant. Our constructed salary index and proportion of students from poor families (as measured by the percentage eligible for participation in the free and reduced-price lunch program) are related to higher spending and are statistically significant; the percentage of students with disabilities (severe and otherwise) is also associated with higher costs though the coefficients of those variables are not statistically significant. Consistent with previous studies, we find a U-shaped relationship between spending per pupil and school district size. The estimated coefficients imply that average costs are lowest in districts with 5,694 students. In contrast to the results of other studies, we find a significantly negative relationship between per pupil education spending and the percentage of students who are in high school.

The Construction of a Cost Index

Estimating a cost function provides us with information about the contributions of various characteristics of school districts to the costs of education. To use this information in school aid formulas, we develop a cost index, which allows us to isolates the variation in school spending attributable to the exogenous cost factors, while holding constant variables that are under the control of the district. In the section "The Design of School Finance Formulas to Achieve Adequacy," this index is integrated into a foundation formula designed to ensure that each school district receives sufficient resources to provide an adequate education for its students.

With a properly constructed cost index we can determine how much each school district must spend in order achieve any given level of educational outcome. Determining a level of educational output that is considered adequate for each state is obviously a public policy decision. One possibility is to define the standard of adequacy as the average level of current student performance within a state (Clune, 1995).

A cost index is constructed by using the results of our cost function estimation to predict hypothetical spending for each district. These predictions are then compared to actual spending in a district with average costs that provides average levels of educational output. Specifically, to determine each school district's hypothetical spending, we multiply the regression coefficients from our estimated cost function with the actual values of the cost fac-

Table 2.—Education cost function, 1994–95 Wisconsin's 368 K–12 school districts		
Variable	Coefficient	t-statistic
Intercept	4.808*	4.508
Log of tenth grade exam score, 1995–96	2.796*	2.282
Log of eighth grade exam score, 1993–94	-1.650	1.573
Number of advanced courses	-0.002*	-2.065
Teacher salary index	1.583*	6.158
Percent of students eligible for free and reduced-price lunch	0.004*	3.078
Percent of students with disabilities	0.004	1.038
Percent of students with severe disabilities	0.131**	1.807
Percent of students enrolled in high school	-0.012*	-2.349
Log of student enrollment	-0.593*	-4.631
Square of log of student enrollment	0.034*	4.106
Sum of squared errors (SSE)	4.594	

* Indicates statistically significant at the 5 percent level.

** Indicates statistically significant at the 10 percent level.

SOURCE: Author's calculations based on the U.S. Bureau of the Census (with data from the 1990 Census of Population and Housing) and the Wisconsin Department of Public Instruction.

tors in each district and the state average values of the educational outcome values.[12] Thus, we set the number of advanced courses and tenth-grade score at the average for all Wisconsin districts. We should emphasize, however, that alternative standards of adequacy could be used in calculating cost indices. The use of different standards will not affect the relative ranking of districts, but will alter the absolute cost index numbers, and hence, will influence any distribution of state aid that is dependent on the cost index.

As discussed above, we use a value-added measure of student achievement in our cost function; that is, the coefficient on tenth-grade scores reflects the increase in spending associated with an increase in achievement, given an initial level of achievement in the eighth grade. Therefore, the expenditures necessary to reach an average level of tenth-grade achievement will depend on the level of student achievement in the eighth grade for that district. Lower eighth-grade achievement implies that it will be more costly to achieve average tenth-grade achievement. In the estimation of the cost function, eighth-grade achievement is treated as an endogenous variable because, like tenth-grade achievement, it is, in part, a choice of the district. In creating the cost index, we want to hold constant any variation in spending that is under the control of the district. Thus, to account for the endogeneity of the eighth-grade scores, we calculate the cost index using predicted eighth-grade scores, with the predictions based on the coefficient estimates from the first-stage regression, actual values of the cost factors, and state average values for the demand instruments. That is, a district's predicted eighth-grade score reflects the score expected from a district with average preferences and observed cost factors. Combined with the average tenth-grade score, the level of spending predicted by the cost function is the spending required to reach average tenth-grade achievement, given average preferences for education and observed cost factors.

Descriptive statistics for Wisconsin's cost index are presented in the first column of table 3. The data clearly show that costs vary tremendously across

[12] Since the salary index is treated as endogenous in the cost function estimation, a predicted salary, based on the first-stage regression, is used in constructing the cost index.

Table 3.—Distribution of education cost indices

	Cost index	Poverty-weighted index	Chambers teacher salary index
Mean	100.0	100.0	99.9
Median	88.8	98.6	99.0
Standard deviation	39.6	9.9	8.2
Range	411.3	58.4	48.4
Minimum	48.9	83.2	82.6
Maximum	460.2	141.6	131.0
Restricted range	73.4	24.9	20.9
Minimum at 10 percent	68.3	88.7	90.1
Maximum at 90 percent	141.7	113.6	110.9
Correlations:			
Cost index	1.000		
Poverty-weighted	0.810	1.000	
Chambers	-0.308	-0.362	1.000

SOURCE: Author's calculations based on the U.S. Bureau of the Census (with data from the 1990 Census of Population and Housing) and the Wisconsin Department of Public Instruction.

school district in Wisconsin. The district with the lowest costs could attain an average level of achievement by spending about half as much per pupil as the district with average costs. At the other extreme, the district with the highest costs must spend more than four and one-half times more than the average cost district to provide an average educational outcome for its students. The large range of the index reflects, in part, the values of the index in a few districts. Ignoring the 10 percent of districts with the lowest index values and the 10 percent of districts with the highest values substantially reduces the range of the cost index. The restricted range in table 3 shows that the district at the 10th percentile has costs that are 32 percent below average and the district at the 90th percentile has costs that are 42 percent above average.

Two school districts have cost indexes that are much higher than the indexes of any other district. Milwaukee's index is 460 and White Lake's is 352, while the district with the next highest index has a cost index of 238. The major reasons for Milwaukee's high cost index are its large size and its high concentration of economically disadvantaged students. With nearly 100,000 students, the district is 45 times the size of the average Wisconsin school district. Seventy-two percent of its students are eligible for free or reduced-price lunches, a proportion that is higher than all but one other Wisconsin school district. White Lake's cost index is high primarily because of its extremely small size (the entire school district has only 250 students) and its very high concentration of children from poor families.

Because our estimated cost functions include no measure of efficiency, it is possible that we are interpreting extra spending that is caused by inefficiencies on the part of school districts as higher costs. In their estimate of cost indexes for school districts in New York State, Duncombe, Ruggiero, and Yinger (1996) report that the maximum cost index declines from 356 to 240 when they replace a cost index calculated without a measure of efficiency with a cost index based on cost function estimates that include an endogenous measurement of efficiency. These New York State results suggests that the high cost indexes for Milwaukee and White Lake may reflect in part some degree of inefficiency on the part of these two local school districts. Duncombe, Ruggiero, and Yinger also report that the cost indexes measured with and without a control for efficiency are highly correlated, with a correlation coefficient equal to 0.94. This suggests that including a measure of efficiency may have relatively little impact on the rank ordering of districts in terms of costs.

The data in the second column of table 3 allow us to compare our cost index to an index that measures the concentration in each district of children from economically-disadvantaged families. We constructed this index by comparing the percentage of low-income students in each district to the average percentage of low-income students in the state. The third column of table 3 displays statistics describing the distribution of the teacher salary index developed by Jay Chambers (1995).

Recall that the Chambers teacher salary index only reflects factors outside school district control that require some districts to pay more or less for teachers with similar qualifications. Thus, while the Chambers index will reflect the higher salaries that school districts may have to pay to induce teachers to work in districts with high concentrations of poor children, these higher salaries are only one reason for the possibly high costs of educating poor children. For example, in order to overcome the educational disadvantages faced by many children from poor families, extra teachers may be needed to staff smaller class sizes and special remedial programs. By contrast, our cost index provides a comprehensive measure that reflects all the factors that lead to costs of achieving any educational outcome to be higher in some districts than in others. It is not surprising that our cost index shows larger variation (as measured by both the standard deviation and the range) than either of the other two indices.

Most members of the educational community use the term adequacy to refer to the achievement of minimum standards of educational performance or outcome.

Table 4 displays the distribution of the three indices across school districts characterized by size (in terms of enrollment), by property wealth per pupil, by urban-rural status, and by the concentration of students from poor families. The data clearly show the U-shaped relationship between district size and costs. They also indicate that costs tend to be high in rural districts, reflecting both small district size and relatively high concentrations of low-income students. Costs also tend to be higher in both the property-poor and the property-rich districts. Property-poor districts tend to be characterized by higher concentrations of students with disabilities and students from economically disadvantaged families. Higher costs in property-rich districts tend to reflect higher than average costs of living in those districts.

Table 4 illustrates quite clearly the differences between our cost index and the Chambers teacher salary index. The highest values of the Chambers index are found in Milwaukee and other urban school districts reflecting primarily the relatively high cost of living in these areas as compared to Wisconsin's rural areas. Average costs in the Milwaukee suburbs (listed as *Urban Fringe, Large City* in table 4) are 24 percent below the state average when measured using our index. At 17 percent above average, the Chambers index indicates that the Milwaukee suburbs have higher costs than any other area in Wisconsin. Also, in contrast to our cost index, the Chambers index tends to be highest in school districts with relatively few pupils from poor families.

Finally, because we are interested in the relationship between costs and achievement, table 5 shows the average index scores by performance on the tenth-grade exam. Using our cost index, or the poverty-weighted index, costs are higher in low-performing districts. This implies that even more resources will be needed to get students in these districts up to "adequate" levels of achievement.

The Design of School Finance Formulas to Achieve Adequacy

Most members of the educational community use the term adequacy to refer to the achievement of minimum standards of educational performance or outcome. Not surprisingly, disagreements arise concerning the level and composition of performance standards that should be considered as adequate. William Clune (1994), for example, defines true adequacy as the achievement of "...high minimum standards in low-income schools..." (p. 378). Although achieving agreement at a national level about the precise definition of high minimum standards may be impossible, individual state governments may be able to decide on a set of performance

Table 4.—Distribution of education indices, by school district characteristics

Social characteristics	Number of K–12 districts	Cost index	Poverty-weighted index	Chambers teacher salary index
District size (number of pupils)				
Less than 500	53	141.6	104.6	91.8
500–999	110	108.6	103.2	96.4
1,000–2,499	126	85.2	97.8	102.3
2,500–9,999	70	75.9	95.1	106.2
10,000–24,499	7	100.2	99.3	108.2
Milwaukee (97,555)	1	460.2	135.6	115.0
Equalized property values (EQV)/pupil				
Less than $125,000	54	117.9	107.0	95.6
125,000–174,999	108	103.2	102.0	97.3
175,000–249,999	127	90.2	97.6	100.3
250,000–399,999	55	96.8	96.9	106.2
400,000 or more	24	104.4	95.8	105.1
Urban/rural status				
Large city	1	460.2	135.6	115.0
Mid-size	18	91.9	99.4	104.7
Urban fringe, large city	20	76.3	91.4	117.2
Urban fringe, mid-size	14	75.8	92.2	104.4
Large town	3	68.4	91.9	99.3
Small town	96	82.7	97.3	100.8
Rural	216	110.9	102.6	97.2
Poverty concentration				
Less than 10 percent	55	72.3	87.8	106.7
10–19.9 percent	114	80.7	93.8	101.3
20–29.9 percent	101	95.4	101.1	98.0
30–39.9 percent	53	118.8	108.4	96.9
40 percent or more	45	170.9	118.7	96.0

SOURCE: Author's calculations based on the U.S. Bureau of the Census (with data from the 1990 Census of Population and Housing) and the Wisconsin Department of Public Instruction.

standards that they believe must be met to provide students with an adequate education.[13]

Foundation formulas used by the majority of states distribute grants so as to guarantee that each school district will be able to achieve a "foundation" level of per pupil spending as long as each district uses a state-determined "minimum" property tax rate. If costs were identical in all school districts, by defining the foundation level as the spending necessary to achieve the state-specified minimum performance level, the state could guarantee that each school district had sufficient resources necessary to provide an adequate level of education.

[13] The establishment of performance standards requires that decisions be made about precisely how a standard is defined. Is a standard achieved if all students meet it, or is it defined in terms of mean performance, or in terms of the percentage of students who perform above a given level?

Table 5.—Distribution of education indices, by student performance

	Number of K–12 districts	Cost index	Poverty-weighted index	Chambers teacher salary index
Tenth grade exam score decile				
1 (lowest)	37	132.8	107.6	97.5
2,3,4,5	147	102.6	101.8	98.7
6,7,8,9	147	92.4	98.1	99.9
10 (highest)	37	87.3	93.0	107.2

SOURCE: Author's calculations based on the U.S. Bureau of the Census (with data from the 1990 Census of Population and Housing) and the Wisconsin Department of Public Instruction.

The results presented in the previous section indicate that costs differ substantially among school districts. Thus, to guarantee the provision of adequate education, we need to develop a foundation formula in which each school district's foundation level of spending varies according to differences in costs and in which the average foundation level equals the dollar amount necessary to meet the performance standards associated with educational adequacy in districts with average costs.

A conventional foundation aid formula is presented in equation (4), where A_i equals foundation aid per pupil in district *i*, E^* is the foundation level of per pupil spending, t^* the mandated local property tax rate, and V_i the property value per pupil in school district *i*:

(4) $A_i = E^* - t^*V_i$

Equation (4) will generate negative aid in districts with high per pupil property values. In practice, these districts are allocated zero aid, or, in some cases, a minimum per pupil grant. The first step in adapting the foundation formula so that it will guarantee that every district has sufficient resources to provide an adequate level of education is to determine a standard of educational performance that is considered adequate. Referring to this standard as S^*, we can define $\tilde{E}$ as the amount a school district with average costs must spend to obtain an adequate educational outcome, S^*. A foundation formula designed to guarantee that every school district has sufficient resources to provide S^* can be written as:

(5) $A_i = \tilde{E}c_i - t^*V_i$

where c_i is the value of the cost index in school district *i*.[14]

To simulate the distribution of aid using this formula we have defined a standard of adequacy as the statewide average score on the tenth-grade Knowledge and Concepts Exam. $\tilde{E}$ is thus defined as the expenditure needed to achieve the average tenth-grade test performance in a district with average costs. The amount of aid allocated to district *i* using this cost-adjusted aid formula will be a function of the per pupil property wealth in *i* and the relative costs in district *i*. Lower average student performance on the eighth-grade tests (holding preferences constant) will lead to higher costs in district *i*, and hence to additional aid.

To provide a baseline upon which to judge the impact of using a cost-adjusted foundation formula, we first simulate the distribution of aid to Wisconsin's 368 K–12 districts using a conventional foundation formula. We have chosen $6,372 as the foundation level of per pupil spending, which is the amount needed to achieve the average tenth-grade test performance in a district with average costs.[15] To add some realism to the simulation, we adjust the required property tax rate (t^*) so that the total

[14] See Ladd and Yinger (1994) for a detailed derivation of a cost-adjusted foundation formula.

[15] The state average expenditure per pupil was $6,084 in 1996–97.

amount of foundation aid distributed is equal to $3.03 billion, the actual amount of equalization aid allocated to K–12 districts in Wisconsin for the 1996–97 school year.[16]

In our second simulation, we allocate foundation aid using the $6,372 foundation level, the 11.8 mill tax rate, and the cost-adjusted foundation aid formula (equation 5). Before simulating the distribution of cost-adjusted foundation aid we adjusted downward the reported cost index for the nine school districts with the highest index values. In particular, we truncated the index at 200; that is, we make the (admittedly arbitrary) decision that no district could have costs that were more than twice the average. This adjustment reflects our view that very high-cost adjustments are not politically feasible and the fact that our index may overstate costs in some districts because it fails to account explicitly for school district inefficiencies.

Even if the state government is willing to reform its school finance system to account for cost differences among school districts and to provide funds to achieve educational adequacy, it may not be willing to devote additional state funds to this effort. To account for this possibility, we simulate a revenue-neutral cost-adjusted foundation formula. The foundation level is adjusted downward so that the total budgetary cost of foundation grants does not exceed the $3.03 billion budgetary cost of the non-cost-adjusted foundation formula. Revenue neutrality requires a lowering of the foundation level from $6,372 to $6,158, with a corresponding reduction in the standard of adequacy that can be financed.

Milwaukee receives the largest per-pupil grant; at $11,532 it is more than twice the largest grant distributed through the non-cost-adjusted formula.

The first column of data in table 6 summarizes the distribution of per pupil foundation aid using a foundation formula without cost adjustments. The average grant equals $3,900 per pupil and the largest grant is $5,404. In nine school districts, per pupil property tax bases are large enough to yield more revenue per pupil at the mandated tax rate than the $6,372 foundation level. These nine districts receive no foundation aid. Milwaukee's grant is equal to $4,635. Although this grant is greater than average, 97 other school districts in the state receive larger per pupil grants.

The second column of table 6 summarizes the distribution of cost-adjusted foundation aid. Because most of the state's largest school districts have above average costs, total cost-adjusted foundation aid totals $171 million more than non-cost-adjusted aid. As expected, the standard deviation of per pupil grants is higher ($2,388 compared to $1,133). Milwaukee receives the largest per pupil grant; at $11,532 it is more than twice the largest grant distributed through the non-cost-adjusted formula. As some relatively high-wealth districts have below-average costs, the number of school districts now getting zero aid increases from 9 to 18.

The data in the third column of table 6 shows that achieving revenue-neutrality results in a distribution of per pupil foundation grants with both a smaller mean and standard deviation. As expected, grants to school districts with relatively high costs are reduced. Milwaukee's grant allocation, for example, is reduced by over $400 per pupil. Since the foundation level is reduced in the revenue-neutral formula, 19 school districts receive zero aid under the revenue-neutral, cost-adjusted formula.

Table 7 provides additional information to allow us to compare a cost-adjusted and a non-cost-adjusted foundation formula. Both formulas use a foundation level that has been defined as the spending per pupil necessary to achieve an adequate educational outcome in districts with average costs. Thus, adjusting the foundation formula for cost differences will increase aid for districts with above-average costs and decrease aid for districts with below-average costs. For 130 of the 368 K–12 districts, using the cost-adjusted formula results in an

[16] The resulting property tax rate is 11.8 mills (1.18 percent).

Table 6.—Distribution of aid per pupil under alternative foundation formulas

	Conventional (no cost adjustment)	Cost-adjusted	Revenue-neutral cost-adjusted
Mean	$3,900	$3,824	$3,622
Median	4,170	3,517	3,328
Standard deviation	1,133	2,388	2,313
Range	5,404	11,532	11,103
Minimum	0	0	0
Maximum	5,404	11,532	11,103
Restricted range	2,606	6,118	5,984
Minimum at 10 percent	2,371	1,014	862
Maximum at 90 percent	4,977	7,132	6,846

SOURCE: Author's calculations based on the U.S. Bureau of the Census (with data from the 1990 Census of Population and Housing) and the Wisconsin Department of Public Instruction.

increase in per pupil aid. The top panel of table 7 illustrates that while per pupil aid remains substantially higher in low-property wealth as compared to high-property wealth districts, the largest percentage increases in aid go to high-wealth school districts. At the same time the largest percentage reductions in aid go the wealthiest districts. This pattern only serves to emphasize that the occurrence of high costs is not closely correlated with school district property wealth.

The data in the bottom panel of table 7 illustrate that the largest increases in aid resulting from using a cost-adjusted foundation formula benefit both small and large districts. While aid increases in over three-fourths of the smallest districts, in those small districts where aid does decline, the declines are generally quite small. The eight school districts with between 10,000 and 25,000 students are evenly split between those that gain and those that lose aid as a result of using a cost-adjusted formula.

The data in table 8 allows us to assess the impact of moving from a non-cost-adjusted to a revenue-neutral, cost-adjusted foundation formula. Because the cost-adjusted formula also has a lower foundation level (E*), 116 of the 368 school districts would receive an increase in aid. The lowering of the foundation level means that some school districts with above-average costs would face a reduction in foundation aid as the aid increases are concentrated among districts with the highest costs. The general pattern of changes in aid across districts characterized by both per pupil property wealth and district size is similar to that displayed in table 7, however, the average increases in aid are smaller and the average reductions in aid are larger.

Poverty Weights in School Aid Formulas

The use of a cost index as part of a state aid formula allows states to simultaneously account for all the factors that lead to cost differences among school districts. Although there are advantages to a comprehensive treatment of cost differences, a number of states have taken a partial approach by replacing actual student enrollment with a weighted student count. In this approach, the weights are designed to reflect the higher costs associated with educating particular groups of students. While these weights are most commonly used for pupils with mental or physical disabilities, roughly one-fourth of all states use some kind of weight to allocate extra funding for either or both low-income and low-achieving students. These weights, which reflect the extra costs associated with low-income students, range in value from 0.15 (Vermont) to 0.625 (Illinois), with most states falling somewhere around 0.25 (Odden and Picus, 1992).

Table 7.—Distribution of state aid under a conventional foundation formula and a cost-adjusted foundation formula

		By district equalized property values (EQV)/pupil									
		Districts with increased aid					Districts with decreased aid				
EQV/pupil	Number of K–12 districts	Number with increased aid	Aid/pupil no cost adjustment	Aid/pupil with cost adjustment	Dollar change	Percent change	Number with decreased aid	Aid/pupil no cost adjustment	Aid/pupil with cost adjustment	Dollar change	Percent change
Less than											
$125,000	54	36	$5,088.41	$7,084.30	$1,995.89	39%	18	$5,021.72	$4,288.66	($733.05)	-15%
125,000–174,999	108	45	4,617.37	6,154.46	1,537.09	33%	63	4,574.06	3,545.89	(1,028.17)	-22%
175,000–249,999	127	32	3,884.50	5,783.59	1,899.09	49%	95	3,899.47	2,425.85	(1,473.61)	-38%
250,000–399,999	55	11	2,673.37	6,481.00	3,807.63	142%	44	2,832.26	1,346.06	(1,486.20)	-52%
400,000 or more	24	6	598.11	4,558.05	3,959.93	662%	18	744.81	80.21	(664.60)	-89%
Total	**368**	**130**	**$4,217.42**	**$6,274.61**	**$2,057.20**	**49%**	**238**	**$3,727.03**	**$2,486.19**	**($1,240.83)**	**-33%**

		By district size									
		Districts with increased aid					Districts with decreased aid				
EQV/pupil	Number of K–12 districts	Number with increased aid	Aid/pupil no cost adjustment	Aid/pupil with cost adjustment	Dollar change	Percent change	Number with decreased aid	Aid/pupil no cost adjustment	Aid/pupil with cost adjustment	Dollar change	Percent change
Less than 500	53	43	$4,218.18	$7,200.02	$2,981.84	71%	10	$2,456.12	$2,235.60	($220.52)	-9%
500–999	110	55	4,173.97	6,036.93	1,862.96	45%	55	4,251.18	3,444.07	(807.11)	-19%
1,000–2,499	126	24	4,425.63	5,414.75	989.13	22%	102	3,808.35	2,427.82	(1,380.53)	-36%
2,500–9,999	70	3	4,421.06	5,087.97	666.91	15%	67	3,380.70	1,837.55	(1,543.15)	-46%
10,000–24,999	8	4	3,300.27	4,460.79	1,160.51	35%	4	3,424.42	2,295.18	(1,129.24)	-33%
Milwaukee	1	1	4,634.97	11,007.12	6,372.15	137%	0	—	—	—	—
Total	**368**	**130**	**$4,217.42**	**$6,274.61**	**$2,057.20**	**49%**	**238**	**$3,727.03**	**$2,486.19**	**($1,240.83)**	**-33%**

— Not applicable.

SOURCE: Author's calculations based on the U.S. Bureau of the Census (with data from the 1990 Census of Population and Housing) and the Wisconsin Department of Public Instruction.

Table 8.—Distribution of state aid under a conventional foundation formula and a revenue-neutral cost-adjusted foundation formula

		By district equalized property values (EQV)/pupil										
		Districts with increased aid					Districts with decreased aid					
EQV/pupil	Number of K–12 districts	Number with increased aid	Aid/pupil no cost adjustment	Aid/pupil with cost adjustment	Dollar change	Percent change	Number with decreased aid	Aid/pupil no cost adjustment	Aid/pupil with cost adjustment	Dollar change	Percent change
Less than $125,000	54	33	$5,098.13	$6,984.01	$1,885.88	37%	21	$5,015.98	$4,201.29	($814.70)	-16%
125,000–174,999	108	38	4,632.11	6,158.14	1,526.03	33%	70	4,570.39	3,472.33	(1,098.06)	-24%
175,000–249,999	127	29	3,899.08	5,701.51	1,802.43	46%	98	3,894.69	2,302.65	(1,592.04)	-41%
250,000–399,999	55	11	2,673.37	6,138.99	3,465.61	130%	44	2,832.26	1,192.03	(1,640.24)	-58%
400,000 or more	24	5	717.74	5,052.91	4,335.18	604%	19	705.61	47.79	(657.82)	-93%
Total	**368**	**116**	**$4,226.96**	**$6,229.47**	**$2,002.51**	**47%**	**252**	**$3,749.88**	**$2,421.85**	**($1,328.02)**	**-35%**

		By district size									
		Districts with increased aid					Districts with decreased aid				
EQV/pupil	Number of K–12 districts	Number with increased aid	Aid/pupil no cost adjustment	Aid/pupil with cost adjustment	Dollar change	Percent change	Number with decreased aid	Aid/pupil no cost adjustment	Aid/pupil with cost adjustment	Dollar change	Percent change
Less than 500	53	42	$4,214.58	$6,948.03	$2,733.46	65%	11	$2,630.05	$2,291.00	($339.05)	-13%
500–999	110	48	4,212.27	6,046.99	1,834.72	44%	62	4,212.81	3,324.61	(888.20)	-21%
1,000–2,499	126	18	4,445.95	5,476.77	1,030.82	23%	108	3,839.26	2,376.79	(1,462.46)	-38%
2,500–9,999	70	3	4,421.06	4,851.47	430.41	10%	67	3,380.70	1,698.66	(1,682.04)	-50%
10,000–24,999	8	4	3,300.27	4,207.70	907.43	27%	4	3,424.42	2,119.03	(1,305.39)	-38%
Milwaukee	1	1	4,634.97	10,578.94	5,943.97	128%	0	—	—	—	—
Total	**368**	**116**	**$4,226.96**	**$6,229.47**	**$2,002.51**	**47%**	**252**	**$3,749.88**	**$2,421.85**	**($1,328.02)**	**-35%**

— Not applicable.

SOURCE: Author's calculations based on the U.S. Bureau of the Census (with data from the 1990 Census of Population and Housing) and the Wisconsin Department of Public Instruction.

As far as we can determine, the process of determining the weights assigned to low-income children often reflect political considerations rather than estimates of the true costs of educating children from economically disadvantaged families.

Although we believe that it is preferable to use a comprehensive index of costs, using weights for specific populations of students is still an improvement over state aid formulas that do not make any attempt to account for cost differences. In this section, we use our estimated cost functions to calculate a weight based on the extra costs associated with educating children from poor families. If poverty weights are going to be used in state aid formulas, in our view it is preferable that the magnitude of the weights reflect as accurately as possible the extra costs associated with educating poor children.

The use of a poverty weight implies that the first poor pupil in a school district contributes the same amount to extra costs as the 500th poor pupil. It appears more reasonable to assume that the first few poor students contribute little to extra costs, while after a threshold proportion of poor students, costs begin to rise as the number of poor students increases. By estimating equation (3) for different subgroups of districts defined by their percentage of students from economically disadvantaged families, we find support for this hypothesis. The data suggest that the threshold above which additional poor children lead to higher costs is at the eighth percentile on the distribution of the percentage of students eligible for the free and reduced-price lunch program.

. . . determining the weights assigned to low-income children often reflect political considerations rather than estimates of the true costs of educating children from economically disadvantaged families.

To create a poverty weight, we used the results of our estimation of equation (3) to predict total expenditures in each district. For each district, we then set the percentage of students from poor families equal to the threshold level and recalculated predicted expenditures. This latter prediction tells us the cost of educating a regular mix of students. The difference between the two predictions, divided by the number of poor students in each district in excess of the threshold level provides a measure of the additional costs in each district associated with educating students from poor families. A district-specific poverty weight is determined by dividing this measure by the cost for each regular (non-poor) student.[17] The results of these calculations indicate that both the mean and median weight equals 1.59, with the individual district weights distributed very tightly around the mean value.

A poverty weight of 1.59 indicates that to achieve any given level of educational outcome costs two and a half times as much money as required to educate a regular student. The fact that our poverty weight is considerably larger than the largest poverty weight used by those states that include such weights in their equalization aid formulas, suggests that these other states underestimate the true costs of educating poor children.

Although our poverty weight for Wisconsin is high relative to weights used in state aid formulas, it is much closer to William Clune's (1994) estimate of the additional "cost" of educating students in high-poverty schools. Clune argues that these extra costs are about $5,000 per poor pupil. As the national average spending in these schools was also about $5,000, Clune's estimate implies a poverty weight of about 1.0. Using current spending data, Clune's estimate of the per pupil cost of educating poor children would be closer to $6,000. Although Clune admits that his cost estimate is more of an educated guess than a precise calculation, it is more or less consistent with our results that are based on a complex statistical estimate of the costs of education.[18]

[17] School districts with few poor children are assigned a poverty weight of zero.

[18] In Wisconsin, the average spending per pupil for regular students is $5,082. Thus, using a poverty weight of 1.59, the average district would require an additional $8,080 for each poor student.

Conclusions

There appears to be a growing public awareness that the receipt of a high-quality education is the key to economic success, and at the same time a realization that the education received by a substantial number of students in the United States, especially in large cities, is not of high enough quality to prepare them for well-paying jobs. Although these failures of the U.S. system of public education have been well documented, a heated public debate is raging over how to improve public education.

A number of scholars have argued that improving the performance of public education requires, as a necessary though not sufficient step, reform of the financing of public schools. While most efforts over the past several decades to reform school financing have focused on equalizing the resources available for education, in recent years reformers have attempted to link financing to the actual educational performance of students. A relatively new goal of school finance reformers is the achievement of educational adequacy which is defined in terms of a minimum acceptable level of educational performance for all students, including those who come from economically disadvantaged families.

The key to linking educational outcomes to school financing is the integration of cost considerations into school financing formulas. Costs are defined as the minimum amount of money that a school district must spend in order to achieve a given educational outcome. In this paper, we estimated a cost function for elementary and secondary public education using data from Wisconsin school districts. We used the results of our estimate to construct a cost index. We then integrated the cost index into a foundation formula designed to guarantee that each school district would have sufficient resources available to achieve educational adequacy, which we defined in this paper as state average performance on a comprehensive achievement examination taken by tenth-grade students. We concluded that the State of Wisconsin could finance adequacy by increasing state aid to local school districts by approximately 6 percent, with aid distributed using a cost-adjusted foundation formula.

It is important to emphasize that providing school districts with enough resources to achieve educational adequacy does not in itself guarantee that students will be provided with an adequate education. Additional financial resources must be accompanied with strict accountability standards. States will need to develop financial incentives or penalties, plus other administrative mechanisms, to assure that local school districts actually improve educational outcomes and meet their goals of educational adequacy. If local school districts fail to meet these standards of performance, state governments may have to assume direct administrative control over local districts.

References

Bessent, Authella M. and E. Wailand Bessent, "Determining the Comparative Efficiency of Schools through Data Envelopment Analysis." *Educational Administration Quarterly* 16 (Spring 1980): 57–75.

Berne, Robert and Leanna Stiefel. *The Measurement of Equity in School Finance.* Baltimore: Johns Hopkins Press 1984.

Bradbury, Katharine L., Helen Ladd, Mark Perrault, Andrew Reschovsky, and John Yinger. "State Aid to Offset Fiscal Disparities Across Communities." *National Tax Journal* 37 (2) (June 1984): 151–170.

Bradford, David F., Robert A. Malt, and Wallace E. Oates. "The Rising Cost of Local Public Services: Some Evidence and Reflections," *National Tax Journal* 22 (June 1969): 185–202.

Chaikind, Stephen, Louis C. Danielson, and Marsha L. Brauen. "What Do We Know About the Costs of Special Education? A Selected Review." *The Journal of Special Education* 26 (4) (1993): 344–370.

Chambers, Jay G. "Cost and Price Level Adjustments to State Aid for Education; A Theoretical and Empirical Review." In *Perspectives in State School Support Programs,* Second Annual Yearbook of the American Education Finance Association, edited by K. Forbis Jordan and Nelda H. Cambron-McCabe. Cambridge, MA: Ballinger Publishing Company 1981.

Chambers, Jay. *Public School Teacher Cost Differences Across the United States,* Analysis/Methodology Report, National Center for Education Statistics (NCES 95–758). Washington, D.C.: U.S. Department of Education, October 1995.

Clune, William H. "The Shift from Equity to Adequacy in School Finance." *Education Policy* 8 (December 1994): 376–394.

Clune, William H. "Building a Systemic Remedy for Educational Adequacy: Starting With What We Know." Paper prepared for the Consortium for Policy Research in Education (CPRE), Wisconsin Center for Educational Research, University of Wisconsin-Madison, Madison, WI, November 1995.

Courant, Paul N., Edward M. Gramlich, and Susanna Loeb. "Educational Reform in Michigan." Paper presented at a Federal Reserve Bank of Chicago Symposium, *Midwest Approaches to School Finance,* October 26–27, 1994.

Deller, Steven, Gary Green, and Paul Voss. "Wisconsin Cost of Living Estimates." Memorandum to Carl O'Conner and Pat Walsh, Madison, WI: Department of Agricultural and Applied Economics, University of Wisconsin-Madison, November 13, 1996.

Deller, Steven C. and Edward Rudnicki. "Production Efficiency in Elementary Education: The Case of Maine Public Schools." *Economics of Education Review* 12 (1) (January 1993): 45–57.

Downes, Thomas A. and Thomas F. Pogue. "Adjusting School Aid Formulas for the Higher Cost of Educating Disadvantaged Students." *National Tax Journal* 47 (1) (March 1994): 89–110.

Duncombe, William, John Ruggiero, and John Yinger. 1996. "Alternative Approaches to Measuring the Cost of Education." In *Holding School Accountable; Performance-Based Reform in Education*, edited by Helen F. Ladd, 327–356. Washington, D.C.: The Brookings Institution, 1996.

Green, Richard K. and Andrew Reschovsky. "Fiscal Assistance to Municipal Governments." In *Dollars & Sense; Policy Choices and the Wisconsin Budget*, Volume III, edited by Donald A. Nichols, 91–117. Madison, WI: The Robert M. La Follette Institute of Pubic Affairs, 1994.

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

Hanushek, Eric A. "Assessing the Effects of School Resources on Student Performance: An Update." *Educational Evaluation and Policy Analysis* 19 (2) (Summer 1997): 141–164.

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

Inman, Robert P. "The Fiscal Performance of Local Governments: An Interpretive Review." In *Current Issues in Urban Economics*, edited by Peter Mieszkowski and Mahlon Straszheim, 270–321. Baltimore: The Johns Hopkins Press: 1979.

Ladd, Helen F., Andrew Reschovsky, and John Yinger. "City Fiscal Condition and State Equalizing Aid: The Case of Minnesota." *Proceedings of the 84th Annual Conference on Taxation*, 42–49. Columbus, Ohio: National Tax Association: 1992.

Ladd, Helen F. and John Yinger. "The Case for Equalizing Aid." *National Tax Journal* 47 (March 1994): 211–224.

McCarty, Therese A. and Suthathip Yaisawarng. "Technical Efficiency in New Jersey School Districts." In *The Measurement of Productive Efficiency; Techniques and Applications*, edited by Harold O. Fried, C.A. Knox Lovell, and Shelton S. Schmidt. New York: Oxford University Press 1993.

McDuffy v. Secretary of Education. 1993. 615 N.E.2d 516 (Mass).

McMahon, Walter W. "Geographical Cost of Living Differences: An Update." *American Real Estate and Urban Economics Journal* 19 (3) (1991): 426–450.

McMahon, Walter W. *Interstate Cost Adjustments*, Unpublished paper 1994.

Meyer, Robert H. "Value-Added Indicators of School Performance." In *Improving the Performance of America's School*, edited by Eric A. Hanushek and Dale W. Jorgeson, 197–223. Washington, D.C.: National Academy Press: 1996.

Odden, Allan R. and Lawrence O. Picus. *School Finance: A Policy Perspective*. New York: McGraw-Hill, Inc.: 1992.

Ratcliff, Kerri, Bruce Riddle, and John Yinger. "The Fiscal Condition of School Districts in Nebraska: Is Small Beautiful?" *Economics of Education Review* 9 (1) (1990): 81–99.

Reschovsky, Andrew. "Fiscal Equalization and School Finance." *National Tax Journal* 47 (March 1994): 185–197.

Rose v. Council for Better Education, Inc. 1989. 790 S.W.2d 186, 212–213 (Kentucky).

Ruggiero, John. "Efficiency of Educational Production: An Analysis of New York School Districts." *Review of Economics and Statistics* 78 (3) (August 1996): 499–509.

Serrano v. Priest. 1971. 96 Cal. Rptr. 601, 487 P.2d 1241.

U.S. Department of Education, National Center for Education Statistics, 1997. *The Digest of Education Statistics*, 1996, table 155. Washington, D.C., available on the world wide web at www.nces.ed.gov/pubs/d96/D96T155.html.

Acknowledgments

We are grateful to William Fowler, Steven Deller, Allan Odden, and Kim Rueben for helpful comments on an earlier draft. Financial support was provided by the Consortium for Policy Research in Education (CPRE), Grant No. OERI-R3086A60003 and by the Wisconsin Center for Education Research, School of Education, University of Wisconsin-Madison. The opinions expressed in this paper are those of the authors and do not reflect the views of the National Institute on Educational Governance, Finance, Policy-Making and Management, U.S. Department of Education, Office of Educational Research and Improvement, the institutional partners of CPRE, or the Wisconsin Center for Education Research.

Using Cost and Need Adjustments to Improve the Measurement of School Finance Equity

Lauri Peternick and Becky A. Smerdon

American Institutes for Research

William Fowler, Jr.

National Center for Education Statistics

David H. Monk

Cornell University

About the Authors

Lauri Peternick is a Research Scientist at the American Institutes for Research. Her areas of specialization include education finance and organizational behavior. She is presently finishing her doctoral thesis on the variation of site-based decision making within a single school district, from the University of Chicago. She has a Bachelor's Degree in Political Science from Columbia University and a Master's Degree in Education Policy Analysis from Stanford University. In addition to her work on equity litigation, and cost-of-education indexes, she has presented, written and co-authored papers on student-level resource cost models, site-based decision making and school-level data collection. Her most recent article, "Site-Based Budgeting in Fort Worth, Texas" appeared in the spring 1998 issue of the *Journal of Education Finance.*

Becky Smerdon is a Research Scientist at the American Institutes for Research. Her areas of specialization include educational equity and school organization. She has a Bachelor's Degree from the University of Michigan in Political Science and History, and a Master's Degree in Educational Foundations and Policy. She is presently finishing her doctoral thesis titled "The Effect of High School Membership on Student's Academic Development: Implications for Students At Risk of Educational Failure," from the University of Michigan. She has presented, co-authored, and written papers on classroom instruction, educational equity and school organization.

William J. Fowler, Jr. is the director of the Education Finance Statistical Center (EFSC) at the National Center for Education Statistics (NCES). Dr. Fowler received the Outstanding Service Award of the American Education Finance Association in 1997, and served on its Board of Directors from 1992 to 1995. He also serves on the editorial board of the *Journal of Education Finance*. Dr. Fowler is a graduate of Columbia University with a doctorate in education (1977).

David H. Monk is professor of educational administration and chair of the Department of Education at Cornell University. He earned his Ph.D. at the University of Chicago and has taught in a visiting capacity at the University of Rochester and the University of Burgundy in Dijon, France.

Monk is the author of Educational Finance: An Economic Approach and Raising Money for Education: A Guide to the Property Tax (with Brian O. Brent) as well as numerous articles in scholarly journals. He serves on the editorial boards of The Economics of Education Review, The Journal of Education Finance, Educational Policy, and the Journal of Research in Rural Education. He consults widely on matters related to educational productivity and the organizational structuring of schools and school districts and is a Past President of the American Education Finance Association.

Using Cost and Need Adjustments to Improve the Measurement of School Finance Equity

Lauri Peternick and Becky A. Smerdon

American Institutes for Research

William Fowler, Jr.

National Center for Education Statistics

David H. Monk

Cornell University

Over the last 30 years, there has been a movement throughout the country to alter the sources of funding for school systems, because they are thought to be inequitable. Beginning with the McInnis v. Shapiro[1] case in 1968, the courts have been asked to review the constitutionality of educational funding systems which rely on local property tax. The focus of this debate lies in the fact that a dependence on local property tax leads to enormous disparities in education funding between school districts. It has been argued that these disparities violate the equal protection clause of state and federal constitutions as well as states' constitutional obligations and commitments with regard to education.

The legal debate over these issues has been well publicized and recognized in our society. Accompanying this public phenomenon has been a less well-known, but equally aggressive movement to analyze educational funding inequality, and to develop methods of measuring funding equity. Just as litigation in this area has grown, so to has analysis of equity within educational funding. Verstegen (1998) identifies four developments within the field of equity analysis. These include:

- Redefining the constitutionally required level of education a state must provide;
- Focusing on adequacy in addition to equity;
- Relying on the plain meaning of education clauses in state constitutions; and
- Using new criteria for measuring constitutional compliance.

The following paper will focus on the fourth of these four points. More specifically, the paper reviews two methods for adjusting per-pupil expenditure figures with the aim of more accurately measuring equity. These two methods are weighted pupil adjustments and the application of geographic cost-of-education indices. The purpose of these adjustments is to take into consideration extenuating circumstances and the additional burdens school officials face when trying to provide a quality education to students.

[1] McInnis v. Shapiro 293 F. Supp. 327 (1968).

Weighted Pupil Adjustments

The enormous differences between various communities' schools become evident immediately upon walking in the door. Not only are there obvious differences regarding the ages of the children served, the physical condition of the buildings, and the amenities provided, but also the resources and advantages or impediments which accompany children to school. The federal government has recognized the varying needs of children and in so doing has provided food and additional funds for students of varying populations, (i.e. Chapter I students, students who face language barriers, and students with special physical needs). These children require greater resources to share in comparable educational experiences with children who are not confronted with these issues. To address the varying needs of students, the federal and many state governments use a weighted student model count for the distribution of grants to school districts. Under such a system, a student with special needs might be accounted for as 1.2 or 2.3 students. The rationale for this weighted count is the needed recognition for additional resources for that particular student and the additional burden placed on the school system to provide an adequate educational experience for all the children they serve.

Much of this recognition and additional effort is mandated in P.L. 94-142, the Individuals with Disabilities Education Act (IDEA), which requires a free, appropriate public education for all children with disabilities. Passed in October 1990, this act is a reauthorization of the Education of the Handicapped Act. To meet this mandate, local school districts need to ensure that students with disabilities are placed in the least restrictive environment appropriate for their educational progress. Each student must have an Individualized Educational Plan (IEP) as well as the necessary related services.

However, these mandates, as well as services provided to students with special needs that are not covered under IDEA, (e.g., children at risk or children with limited English proficiency) can be met in a variety of ways. Services to aid these students can be provided in a self-contained classroom, resource room, residential school or through mainstreaming. Providers of such services include the district, co-operative programs, and private organizations. Moreover, in addition to special teachers, particular instructional materials, and other core educational services, additional services such as transportation and counseling may be needed to support these specialized educational experiences. On average, costs for meeting the needs of a special education child are approximately 2.3 times that of a child in regular education, which often translates into a weighted pupil system where a special education student is weighted as 2.3 students.

Studies reviewing the cost of students with . . . additional needs show that students with more prevalent disabilities tend to have lower average costs, whereas students dealing with less prevalent issues have higher costs.

Studies reviewing the cost of students with special vulnerabilities or additional needs show that students with more prevalent disabilities tend to have lower average costs, whereas students dealing with less prevalent issues have higher costs. Average special education costs can range from approximately $1,000 per pupil for students with speech or language impairments to over $30,000 per pupil for those who are deaf and blind.[2] Programs that utilize resource rooms have lower average costs, while self-contained classrooms and residential schools yield higher costs. Costs in this area may also be distinguished as either "Supplementary Costs" or "Replacement Costs." "Supplementary Costs" refer to services that are provided in addition to regular education costs while "Replacement Costs" refer to programs and services provided instead of regular education. Two important criteria for determining the cost of these programs for students are the eligibility and placement criteria used with regard to the student and

[2] Stephen Chaikind, Louis C. Danielson, and Marsha L. Braun. "What Do We Know About the Costs of Special Education? A Selected Review." *The Journal of Special Education 26,* (4)(1993):344–370.

the budgetary environments under which jurisdictions operate.

The use of weighted pupil counts represents one important way analysts may denote the additional financial burdens school officials face when working with special-needs students. It is one ingredient in the construction of a framework within which school systems must learn to function. A second tool for recognizing special circumstances faced by school systems is a geographic cost-of-education index.

Cost-of-Education Indices

For the past three decades, researchers have conducted studies to develop methodologies and empirical estimations of cost-of-education indices (CEIs). The purpose of these indices is to put into context the value of educational dollars by adjusting for differences in the purchasing power of different school systems. CEIs may be used for resource analysis in two ways. First, CEIs may influence analyses regarding estimating funds needed for educational services. Second, cost-of-education adjustments may be necessary when comparing the financial resources available to students with similar educational needs in geographically disparate locations.

The purpose of these indices [cost-of education indices] is to put into context the value of educational dollars by adjusting for differences in the purchasing power of different school systems.

When exploring the role of indices, one must note that conceptually similar geographic cost-adjustment indices rely on different approaches to account for contextual differences in the hopes of providing an accurate assessment of resources and costs. Examples of these indices include "Average-teacher-salary index" (Barro, 1992, "Cost of Living Index" (McMahon and Chang, 1991), and "Teacher Cost Index" (Chambers and Fowler, 1995). These works focus on developing an adequate methodology for determining differences in personnel costs across locations. The rationale for focusing on personnel costs is that they account for 80 percent of local school budgets (Chambers, 1996). The Barro, McMahon, and Chang, and Chambers and Fowler cost indices were calculated using the Schools and Staffing Survey (SASS) developed by the National Center for Education Statistics (NCES), Census data, the U.S. Geological Survey, and the National Climatic Data Center, data from the Bureau of Labor Statistics, and data from the American Chamber of Commerce Research Association (see Chambers and Fowler, 1995 for details). In each case, cost estimates are presented for all fifty states and provide state and regional comparisons of the alternative teacher cost indices.

Average-teacher-salary index

Barro (1992) developed a model that adjusts for variations in teacher salaries based on their level of education and experience. Referred to as the Average Teacher Salary (ATS) index, the measurement is calculated by statistically controlling for such factors as the highest degree earned by the teacher, the number of years the teacher has taught, and whether or not the teacher has professional certification. This cost index implicitly attributes all remaining variation in teachers' salaries, both above and beyond the differences in education and experience, to differences in geographic costs. Thus, all remaining differences in teacher salaries are attributed to such features as disparities in living conditions, teacher quality, local amenities, and random error. Although this model represents an improvement over using average teacher salary, Chambers and Fowler (1995) argue that it does not systematically account for other teacher characteristics (e.g., personal attributes) or attributes of the work environment that might affect the level of teacher compensation. They maintain that such variations must be addressed when assessing variations in teacher costs.

Cost-of-Living Index

Unlike Barro's salary index, McMahon and Chang (1991) developed a method for estimating a cost-of-living (COL) index to account for differences in the purchasing power of educational dollars. McMahon and Chang assert that in order to com-

BEST COPY AVAILABLE

pare salaries across geographic locations, it is necessary to adjust those salaries by the cost of living in different locations. The COL index adjusts for per capita personal income, the median sale prices of existing single family homes, and the percent change in population in the preceding years by state and region using 1981 data from the Bureau of Labor Statistics and 1990 data from the American Chamber of Commerce Research Association (see Chambers and Fowler, 1995 for details). Like Barro's (1992) salary index, the COL index does not take into consideration other important variations in the cost of school personnel. For example, the COL index does not consider that teacher salaries are higher in districts serving more challenging students or those located in high crime areas in order to compensate for the more difficult working conditions (Chambers and Fowler, 1995).

Teacher Cost Index

In addition to including the geographic cost of living, Chambers and Fowler (1995) extend the analysis of teacher costs to include amenities of the labor markets in which public school districts are located. Their teacher cost index (TCI) is based on a hedonic wage model which takes into consideration conditions that attract workers to a geographic area or a certain teaching position. This model captures variations in teacher costs through a comprehensive analysis of the patterns of teacher compensation. The TCI portrays the complexities of employment transactions between individual teachers and their school districts. In addition, it accounts for school district preferences for teacher qualifications and individual teacher preferences for working and living conditions in local communities (Chambers and Fowler, 1995). Specifically, the TCI simulates the effects of factors that reflect differences in cost of living and geographic attractiveness of local communities (e.g., climatic conditions, amenities of urban and rural life, the incidence of crime). The attractiveness of a job assignment is estimated by controlling for personal background characteristics of teachers (e.g., college major, age) and job assignment characteristics (e.g., class size, students' behavior problems).

When exploring issues of equity, the education research field . . . have relied on a variety of measures, each of which pursues different, . . . ways of gauging the magnitude of unequal distribution of resources.

In reviewing cost-of-education indices, it becomes apparent that these instruments are designed to contextualize the value of education dollars by adjusting for differences in the purchasing power of different locations. CEIs are important for estimating both need within a location or locations and equity among locations. Three CEIs were reviewed and presented in this paper, the ATS index, the COL index, and the TCI, and all indices employ different methods to adjust for local variations. The ATS index adjusts for teacher preparation and experience. The COL index adjusts for the cost of living in local communities. The TCI adjusts for personal characteristics of teachers, variations in local amenities, and the job environment. These indices are related and generally provide similar cost estimates across states. However, some interesting variations emerge. These variations suggest that in some localities, teacher costs are more strongly influenced by certain features than others (e.g., cost of living versus teacher preparation and experience). Thus, not only are there variations in the value of currency, but also differences in the cause of such variations. To further highlight the power these indices hold, we now turn to employing such indices when conducting analysis of equity in funding and the correlations between spending and measures of wealth.

Equity Analysis

When exploring issues of equity, the education research field, as well as other disciplines has relied on a variety of measures, each of which pursues different, and not always consistent, ways of gauging the magnitude of unequal distribution of resources. In so doing, the measures represent different aspects of the inequality that can exist in a distribution. Below is a description of four such measures. These include the variance, the Gini coefficient, the McLoone Index, and the slope coefficient.

Variance

Variance is the average difference between the resources received by each unit and the average amount of resources supplied. In an educational example, variance is the difference between dollars received by each school district and the average dollars administered within a state. A large variance statistic indicates a wide diversity of funding and unequal distribution of financial resources. A second way this concept is articulated is as the "coefficient of variation." The coefficient of variation is 100 times the standard deviation divided by the mean. It roughly indicates the percentage above or below the mean within which two-thirds of the observations lie. The coefficient of variation can take on any positive value, with zero indicating perfect equity. It assists standardizing and comparing variances in different locations with different mean spending values.

When used alone, the variance statistic can be somewhat misleading. Because each school district is treated equally, the variance measure is sensitive to extreme cases. Within a given state, one extreme school district, either receiving relatively large or small amounts of money, may result in a large variance statistic and lead to a conclusion of inequality despite the fact that all of the remaining districts received relatively the same amount of resources. Thus, the distribution of resources may be equitable in that state, except for one unusual district. This is a significant problem in educational applications because the distribution of educational resources is often characterized by extreme cases. Generally, to avoid this problem, educational researchers employ a weighting system that weights school district spending by the size of the school district (in enrollment). Large school districts with many students influence the equity measure more than a single, small outlier.

. . . the distribution of educational resources is often characterized by extreme cases.

Gini coefficient

The Gini coefficient is based on the Lorenz curve, which shows the cumulative proportion of the aggregated value of a variable plotted against the cumulative portion of units, when units are ranked in ascending order by the variable. Stated more simply, the Lorenz curve is calculated by first ranking units based on the magnitude that they possess of the variable being measured. In the example below, school districts serve as the units ranked and per-pupil expenditure serves as the variable. The second step is to calculate the cumulative percent distribution. One method of doing this would be to calculate the total share of the variable (e.g., per-pupil expenditures) being received by the lowest 10 percent of the recipients in the distribution, then calculate the percentage of the total received by the lowest 20 percent and so on, until the percentage of the total received by the lowest 90 percent is reached. These figures, one for each 10 percent interval, are then plotted. The axes on the graph are measured in terms of the percentages. The Lorenz curve is created by connecting these points. If the variable has the same value in every unit, the Lorenz curve is a straight line elevating at a positive 45-degree angle. The Lorenz curve would bow downward if the lowest 10 percent received less than 10 percent. The greater the departure from the diagonal, the more pronounced the inequality.

The Gini coefficient is a summary statistic that represents the departure of the Lorenz curve from the diagonal. This coefficient is estimated by calculating the ratio of the area between the diagonal and the Lorenz curve and the total area beneath the diagonal. The larger the Gini coefficient, the greater the inequality. The coefficient ranges from 0 to 1, with 0 indicating perfect equity.

McLoone coefficient

Unlike the variance and Gini coefficients, the McLoone index is sensitive to where along the distribution the inequality exists.[3] The index is used to assess equity in the distribution of variables among units in the lower half of the distribution. It compares what recipients below the median in the

[3] David H. Monk. *Educational Finance: An Economic Approach.* New York: McGraw-Hill (1990).

distribution actually received with the amount they would have received had they been given the same amount as the median recipient. As recipients in the lower half receive similar amounts to those at the middle of the distribution, the McLoone index becomes larger in absolute value. In contrast to the variance and Gini coefficients, this index may be viewed as a measure of *equality* because the measure becomes larger as the distribution becomes more equal. The other statistics may be considered measures of inequality since they become larger as inequality increases.

Slope coefficient

Unlike previously discussed indicators, the slope coefficient provides insight into who is receiving more or less. It does this by identifying the strength of the relationship that exists between two attributes of the units measured. Specifically, the slope coefficient measures the change in one attribute associated with a change in another attribute. In the example below, average household income within a school district represents one attribute and per-pupil expenditure of that district, a second. It is then possible to plot these attributes for each district, draw a line that best represents the degree to which the two attributes correspond, and calculate a slope for that line. This slope would then be the slope coefficient. In this example, a positive slope would indicate that with every unit increase in household income, there is an increase in per-pupil expenditure (i.e., as household income increases, per-pupil expenditure increases). A negative slope would indicate that an increase in household income coincides with a decrease in per-pupil expenditure. The magnitude of the coefficient indicates how much change in per-pupil expenditure is associated with every unit change in household income. Generally, educational researchers wish the relationship between school district wealth and per-pupil spending to be weak, since much of the cause of school district spending differences is the result of local property wealth.

A Case Study Employing Adjustments and Indices

To understand the power student demographic and cost-of-education adjustments hold, we now turn to a case study of financial equity for school districts within the state of New York. The purpose of this case study is to gauge the impact these adjustments may have on educational analysis, as well as the influence they hold in swaying conclusions drawn. Data for this case study comes from the Common Core Data (CCD). The CCD is the National Center for Education Statistics' (NCES) primary database on elementary and secondary public education in the United States and provides an annual, comprehensive, national statistical database of all public elementary and secondary schools and school districts. The CCD comprises a set of five surveys sent to state education departments. Most of the data are obtained from administrative records maintained by the state education agencies (SEAs). Statistical information is collected annually from public elementary and secondary schools, public school districts and the 50 states, the District of Columbia and outlying areas. The SEAs compile CCD requested data into prescribed formats and transmit the information to NCES. The five data sets within CCD can be used separately or in conjunction with one another to provide information on many topics of interest.

The purpose of this case study is to gauge the impact these adjustments may have on educational analysis, as well as the influence they hold in swaying conclusions drawn.

For issues of clarity, we chose to study financial equity within the state by conducting two sets of analyses, one including New York City and one excluding New York City. This is because large metropolitan areas often face very different issues than the surrounding districts and which can mislead equity analyses and their conclusions. By conducting two separate analyses, we hope to minimize this potential problem.

In both sets of analyses, we employed a weighted pupil model and a cost-of-education index, the TCI, to develop four different data sets. One data set includes "Unadjusted" data. A second data set presents "Needs Adjustment" data, based on a weighted pupil model. A third provides regional "Cost Adjusted" data, based on the TCI index, and a fourth set presents data that has been both "Needs and Cost Adjusted." Data used for the analyses were total expenditures per district, total students per district, the number of students with an "Individual Educational Plan" (IEP), the percent of all at-risk children enrolled in school, and the percent of children who speak English "Not Well."

To construct a weighted per-pupil average, we adopted the same method as the state government of New York. Student needs adjustments were calculated by weighting student categories as follows: students with IEPs were multiplied by 2.3, while limited English proficiency and at-risk students were weighted by a factor of 1.2. These multipliers were used on the aggregate for each district, so that an individual student may belong to more than one category and would be multiplied under each classification. Once a weighted student population was determined, "Needs Adjusted" district per-pupil expenditures were calculated by taking the "Total Expenditures" and dividing by the weighted student population.

To determine a "Regional Cost Adjusted" per-pupil expenditure, we divided "Total Current Expenditures" by "Total Students" and then multiplied this figure by the corresponding TCI adjustment. Lastly, to construct a data set that took into consideration a student need and regional cost adjustment, we took the "Needs Adjusted" data and multiplied it by the appropriate TCI adjustment used in the "Regional Cost Adjusted" data. See figures 1, 2, and 3 for graphic displays of the data set distributions. Figure 4 shows the impact adjustments may have when comparing a limited number of districts.

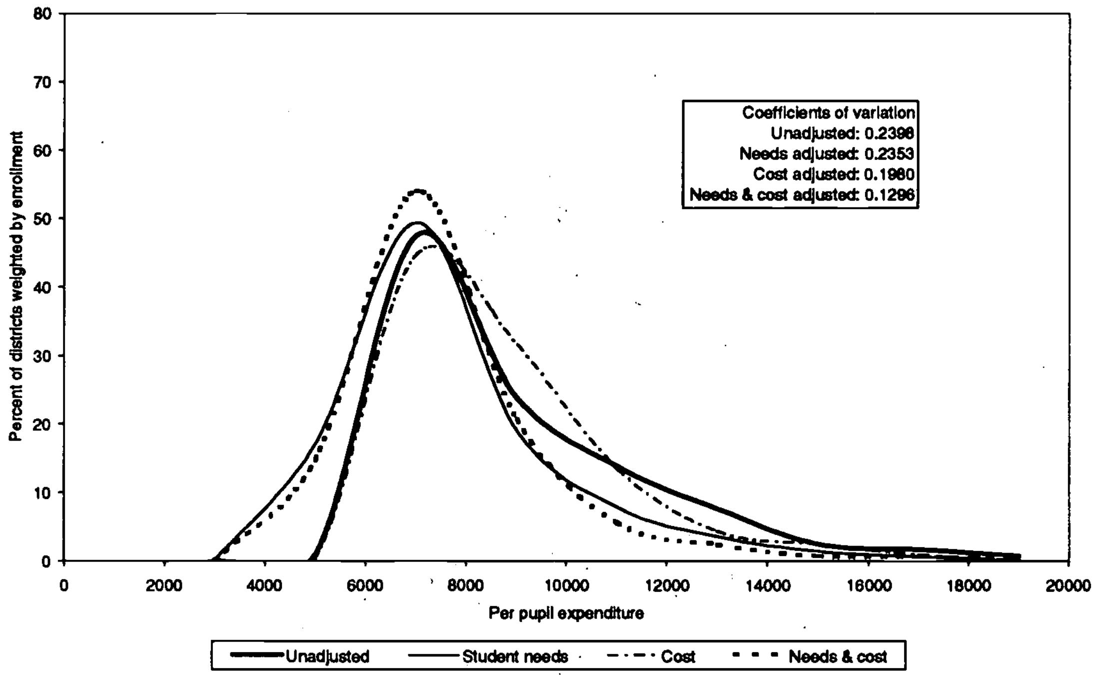
Once the four data sets were created, we applied the four different equity measures noted earlier to determine how these adjustments may affect equity analysis. Table 1 presents comparisons of the three of these equity measures, variance coefficient, Gini coefficient, McLoone coefficient, when applied to the various data sets. The first column is the variance coefficient. The Gini and McLoone coefficients are presented in the second and third columns, respectively. For purposes of comparison, it is important to remember what these coefficients measure. The variance and Gini coefficients measure *inequity* (higher coefficients reflect greater inequity). In contrast, the McLoone coefficient represents *equity* (higher coefficients reflect greater equity). Differences in per-pupil expenditures are estimated for each of the three equity measures based on each of the four data sets. Part A of the table estimates the observed inequity or equity for each of the data sets excluding New York City. Part B estimates the observed inequity or equity for each data set including New York City.

Table 1.—Comparisons of Type I equity measures: Analysis of New York State

	Variance coefficient	Gini coefficient	McLoone index
A. Excluding New York City			
Unadjusted	0.2398	0.1265	0.8878
Needs adjusted	0.2353	0.1227	0.8991
Cost adjusted	0.1980	0.1017	0.8859
Needs and cost adjusted	0.1296	0.0974	0.8947
B. Including New York City			
Unadjusted	0.2096	0.0983	0.9292
Needs adjusted	0.2093	0.0966	0.9252
Cost adjusted	0.2404	0.1240	0.7978
Needs and cost adjusted	0.2421	0.1256	0.7966

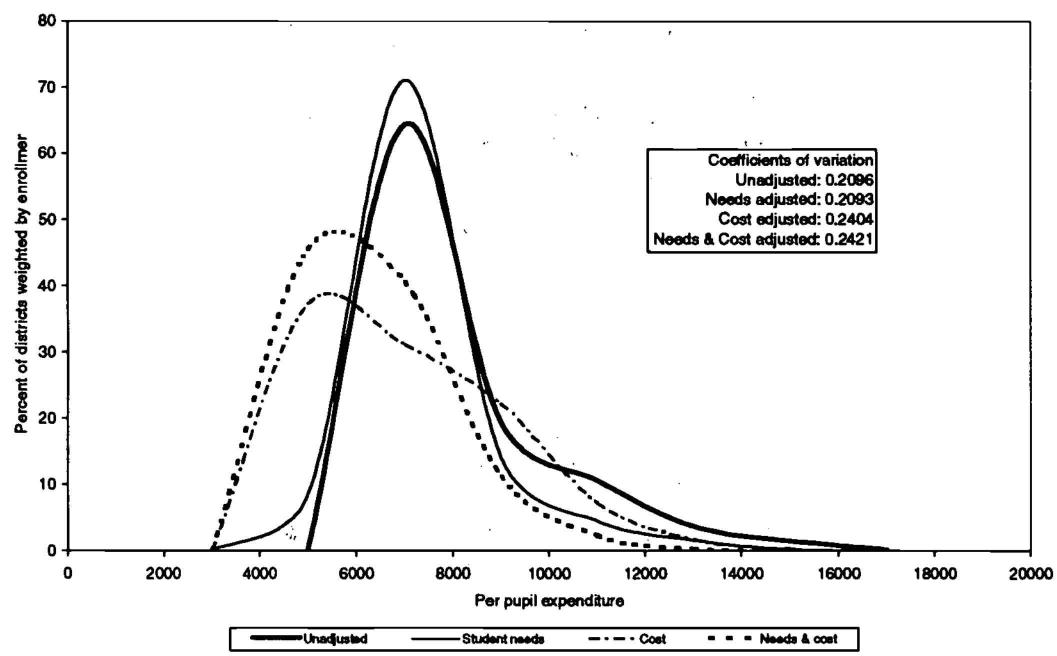
SOURCE: Special tabulation by authors from the Common Core of Data (CCD) using only New York State.

Figure 1.—Data set distribution: Analysis of New York State excluding New York City



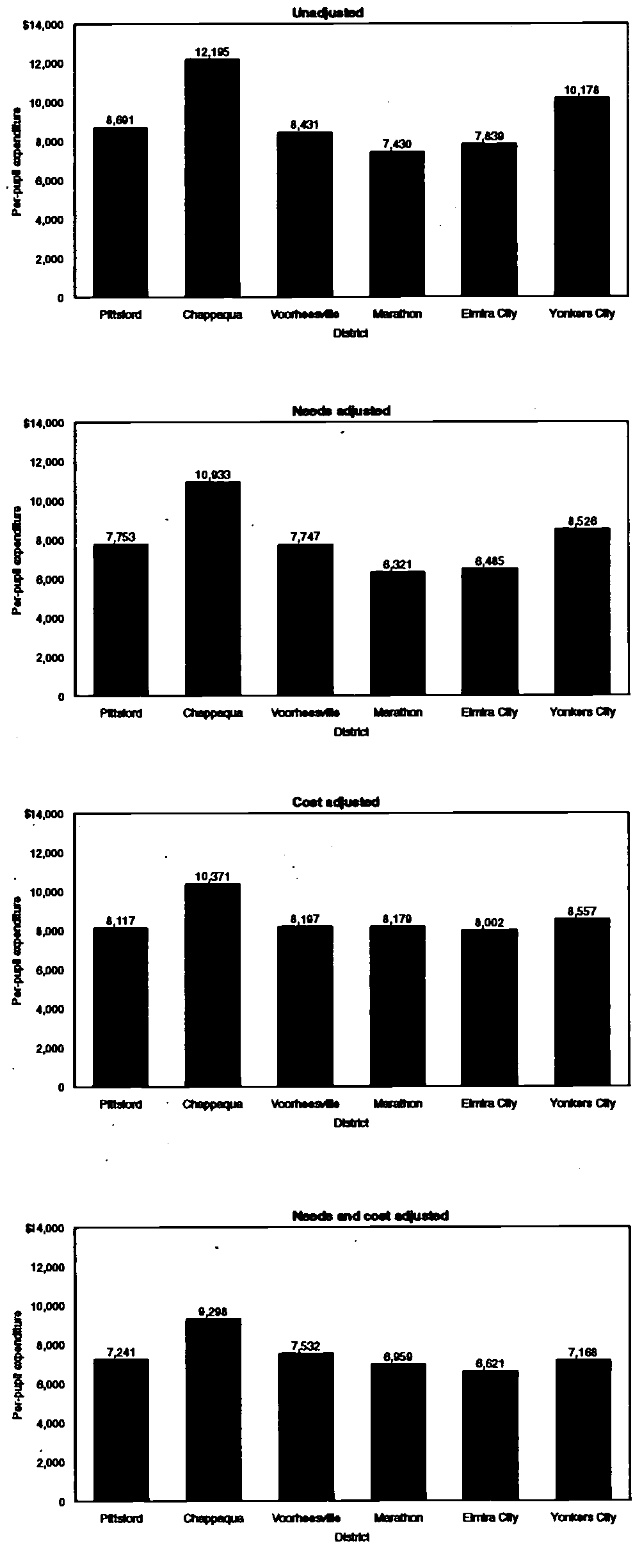
SOURCE: Special tabulation by authors from the Common Core of Data (CCD) using only New York State.

Figure 2.—Data set distribution: Analysis of New York State including New York City



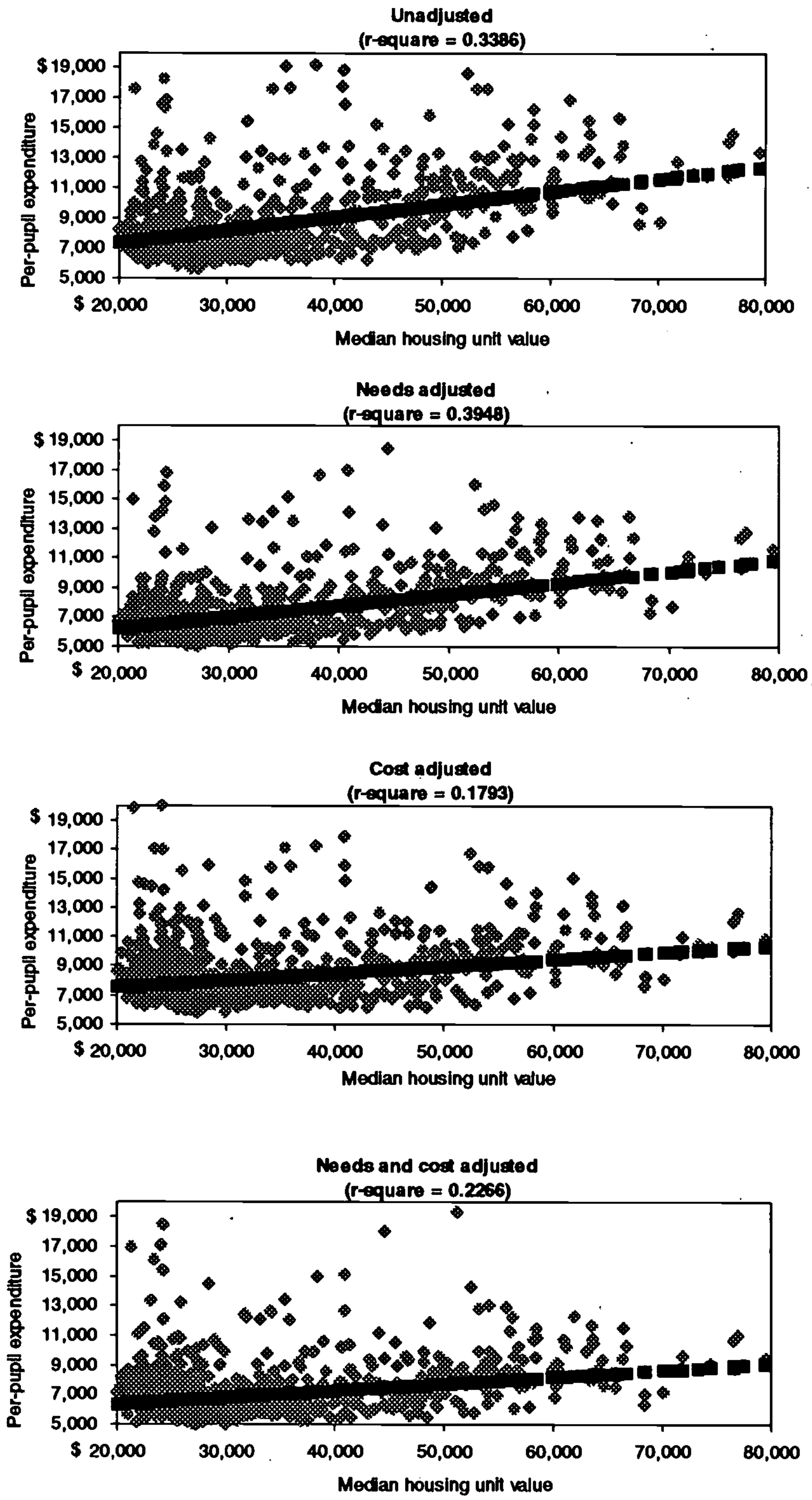
SOURCE: Special tabulation by authors from the Common Core of Data (CCD) using only New York State.

Figure 3.—Per-pupil expenditures for selected school districts



SOURCE: Special tabulation by authors from the Common Core of Data (CCD) using only New York State.

Figure 4.—Regression analysis of per-pupil expenditure by median household income: Analysis of New York State excluding New York City



SOURCE: Special tabulation by authors from the Common Core of Data (CCD) using only New York State.

With regard to the coefficient of variation presented in Part A, the greatest measured inequity appears present when no adjustments are made to the data. The value of this is 0.2398. Once a needs adjustment is made, the situation appears to improve slightly to 0.2353. This improvement may be viewed by some as so small to be deemed insignificant. Remember that a coefficient of variation of 0.24 indicates approximately the percentage above or below the mean within which two-thirds of the observations lie.

However, regional cost adjustments appear to have a stronger impact in this case than do the needs adjustment, decreasing the coefficient to 0.1980. Interestingly, the impact of the needs adjustment increases significantly when coupled with the cost adjustment. In this instance, the variance coefficient decreases to 0.1296. Thus, employing both a cost and need adjustment almost halves the coefficient of variation.

Similar findings hold true when employing the Gini coefficient in measuring equity. "Unadjusted" data provides the most significant measures of inequity (0.1265), followed by "Needs Adjusted" (0.1227), "Cost Adjusted" (0.1017), and lastly "Needs and Cost Adjusted" (0.0974).

The McLoone Coefficient, in contrast, shows a different picture. In this case, the greatest inequity in measurement appears using "Cost Adjusted" data (0.8859), followed by "Unadjusted" data (0.8878), and "Needs and Cost Adjusted" data (0.8947). The greatest measured equity occurs when the data set is only "Needs Adjusted" (0.8891). The difference in this outcome from the previous two equity measures may be traced back to the focus of the McLoone coefficient, that being the lower half of the data set distribution. The McLoone coefficient compares what recipients below the median distribution received with the amount they would have received assuming an equal distribution. In comparing the McLoone index with the variance and Gini coefficient, one may ascertain where the greatest amount of equity or inequity lies within a distribution between the various data sets.

Also of interest are the changes in equity measurement when comparing the data sets that include New York City with those that do not. When New York City is included, the "Unadjusted" and "Needs Adjusted" data set show increased equity, with regard to the variance and Gini coefficient. However, the "Cost Adjusted" and "Needs and Cost Adjusted" indicate greater inequity when New York City is included in the analyses. This information indicates that although at first glance funding for education in New York City seems strongly in line with funding levels in the rest of the state, the issues addressed under the "Cost Adjusted" and "Need and Cost Adjusted" data have a very different impact on funding in New York City than they do in other school districts within the state of New York.

Moreover, focusing solely on data sets that include New York City, applying a regional cost adjustment to the data appears to increase the variance and Gini coefficients indicating greater levels of inequity. This appears to be the case whether one applies it to "Unadjusted" or "Needs Adjusted" data. In both instances, the "Cost Adjusted" and "Needs and Cost Adjusted" indices produce measures indicating greater inequity. A nominal increase in measured equity does occur when using a "Needs Adjusted" data set. The greatest inequity is calculated when "Needs and Cost Adjusted" data are included. Once again, the impact of these cost adjustments indicates that issues considered within the cost adjustment provide a much different burden for the city of New York than they do for the rest of the districts within the state. A final point of interest that should be made is the size of the New York City school district compared with the rest of the state. Approximately one-half of the children attending public school in the state of New York attend New York City schools. This means that if New York City is included in the analysis, then approximately one-half of the data points in the analysis reflect the policies and resources of New York City.

With regard to the McLoone index, in comparing analyses excluding New York City with analyses including New York City, if New York City is included, the measures of equity increase for "Unadjusted" and "Needs Adjusted" data, but decrease for "Cost Adjusted" and "Needs and Cost Adjusted" data. In reviewing analyses that include New York City, one again sees that applying a cost adjustment to the data, whether it is "Unadjusted" or "Needs

Adjusted," produces measures of greater inequity, 0.7978 and 0.7966 with "Needs and Cost Adjusted" data indicating the greatest inequity (see table 1). "Unadjusted" data provide measurements of greatest equity, 0.9292, followed by "Needs Adjusted" data, 0.9252.

Lastly, we applied these adjustments to data when employing the slope coefficient. Eight regression equations were modeled, analyzing the effect of either median household income or median housing unit value on one of the four measures of per-pupil expenditure as weighted by enrollment. These regressions were performed both including and excluding New York City, resulting in sixteen regression results. An analysis of these results again reveals the impact of cost and needs adjustments when examining the disparity in per-pupil expenditures.

The parameter coefficients presented in tables 2 and 3 and figure 4–7 indicate the change in per-pupil expenditure predicted by a one dollar increase in the independent variable. A positive, statistically significant relationship was found in 14 of the 16 regressions. The "Cost Adjusted" and "Need and Cost Adjusted" regressions on housing value failed to reveal a statistically significant relationship when New York City was included in the sample.

When excluding New York City, an 8.48-cent increase in per-pupil expenditure is expected when there is a one dollar increase in median household income, based upon the "Unadjusted" data set. The r-square explains that 0.3386 of the variation in per-pupil expenditure is explained by median household income using an "Unadjusted" data set. The increase in per-pupil expenditure was 7.7 cents when a "Need Adjusted" data set was used. However, the ability for median household income to explain a change in per-pupil expenditure increases slightly when the data are adjusted for student need, 0.3948. When a "Cost Adjusted" data set is employed, the relationship between median household income and per-pupil expenditure decreases to 0.1793. When using "Cost Adjusted" data, a dollar increase in median household income predicts a 4.8-cent increase in per-pupil expenditure. Lastly, the slope coefficient drops to its lowest value with regard to median household income when a "Need and Cost Adjusted" data set is used. In this case a dollar increase in household income leads to a predicted value of only a 4.5-cent increase in per-pupil expenditure, with household income explaining 0.2266 of the change in per-pupil expenditure.

If one includes New York City, the relationship between household income and per-pupil expenditure increases as does the explanatory power of median household income. However, the same pattern of explanatory power resonates with a dollar increase in household income utilizing "Unadjusted" data reflecting the greatest increase in per-pupil expenditure, 8.99 cents. A dollar increase based upon "Needs Adjusted" data indicates a 8.23-cent increase, "Cost Adjusted" 7.59 cents and "Needs and Cost Adjusted" 6.96 cents. "Needs Adjusted" data shows median household income to have the strongest explanatory power, 0.4509, "Unadjusted" data providing the second strongest explanatory power, 0.3908, and cost adjusted data the weakest relationship, 0.2748.

With regard to housing unit value, excluding New York City, once again using "Unadjusted" data provides the largest slope coefficient, 0.0188, indicating a dollar increase in housing unit value predicts a 1.88-cent increase in per-pupil expenditure. "Needs Adjusted" data indicate a 1.64-cent increase in per-pupil expenditure. "Cost Adjusted" indicates a 1.11-cent increase and "Needs and Cost Adjusted" indicates a 0.98 cent increase. "Needs Adjusted" data provide evidence for the greatest explanatory power, 0.5278, with "Unadjusted" data second, 0.4895, "Needs and Cost Adjusted" third, 0.3575, and "Cost Adjusted" data showing the weakest relationship, 0.2853. Including New York City does not appear to significantly change these relationships. "Unadjusted" data still provides for the greatest increase in per-pupil expenditure, 1.88 cents, followed by "Needs Adjusted," 1.64 cents, "Cost Adjusted," 0.14 cents, and "Needs and Cost Adjusted," 0.12 cents. This time "Needs Adjusted" and "Unadjusted" data both provide the strongest evidence for the explanatory power of housing unit value, 0.2161, with "Cost Adjusted" and "Need and Cost Adjusted" data providing negligible explanatory power, 0.0031 and 0.0015, respectively.

Table 2.—Type II equity measures: Median household income

	Slope coefficient	Standard error	R-square
A. Excluding New York City			
Unadjusted	0.0848	0.0045*	0.3386
Needs adjusted	0.0770*	0.0036	0.3948
Cost adjusted	0.0480	0.0039*	0.1793
Needs and cost adjusted	0.0450	0.0032*	0.2266
B. Including New York City			
Unadjusted	0.0899	0.0043*	0.3908
Needs adjusted	0.0823	0.0034*	0.4509
Cost adjusted	0.0759	0.0047*	0.2748
Needs and cost adjusted	0.0696	0.0039*	0.3121

*Significance at p = 0.001.

SOURCE: Special tabulation by authors from the Common Core of Data (CCD) using only New York State.

Table 3.—Type II equity measures: Median housing unit value

	Slope coefficient	Standard error	R-square
A. Excluding New York City			
Unadjusted	0.0188	0.0007*	0.4895
Needs adjusted	0.0164	0.0006*	0.5278
Cost adjusted	0.0111	0.0006*	0.2853
Needs and cost adjusted	0.0098	0.0005	0.3175
B. Including New York City			
Unadjusted	0.0118	0.0009*	0.2161
Needs adjusted	0.0100	0.0007*	0.2161
Cost adjusted	0.0014	0.0010	0.0031
Needs and cost adjusted	0.0012	0.0008	0.0015

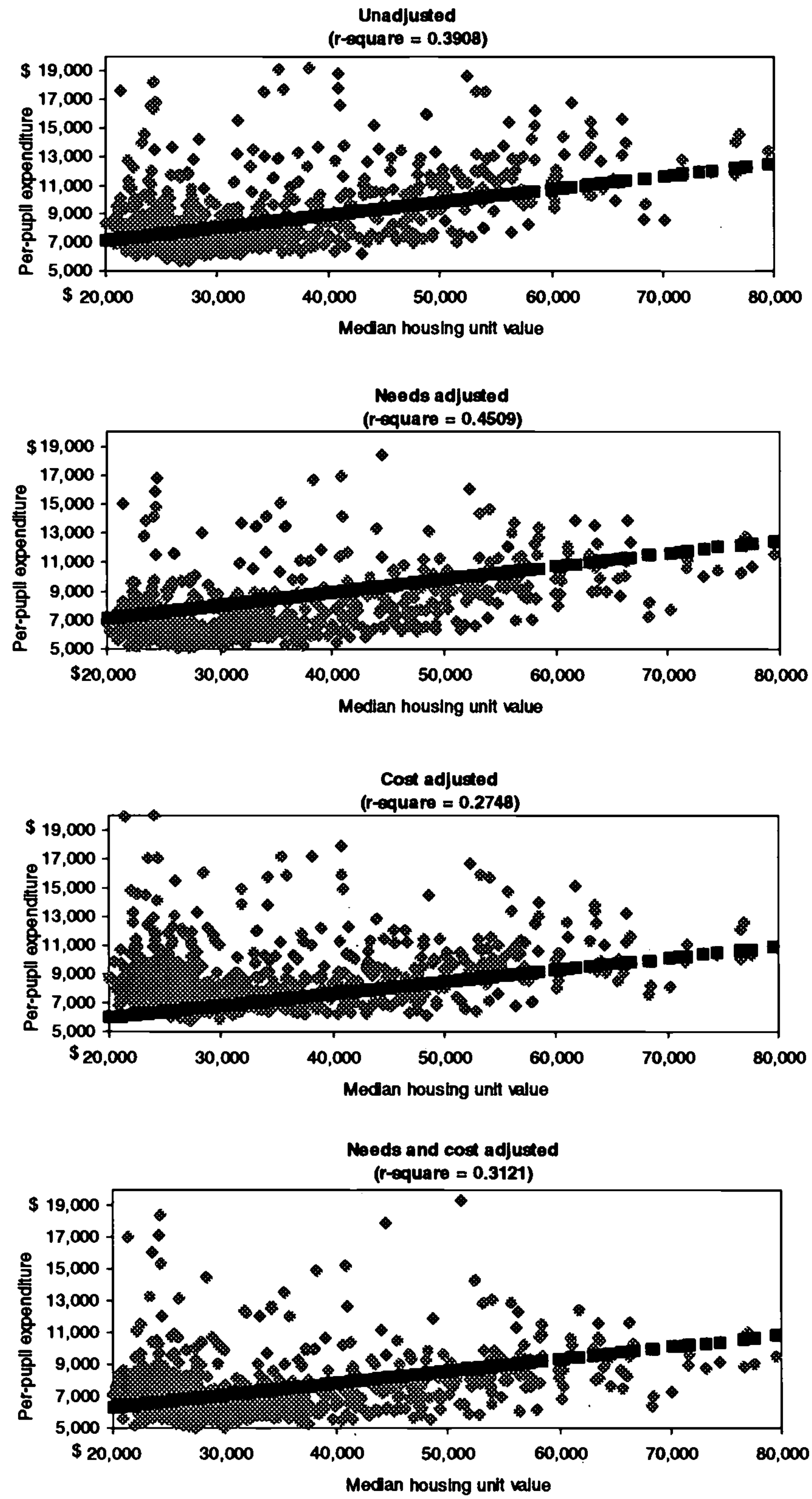
*Significance at p = 0.001.

SOURCE: Special tabulation by authors from the Common Core of Data (CCD) using only New York State.

To interpret these findings, one might conclude that cost adjustments hinder the explanatory power of median household income and housing unit value because many of the issues these adjustments address are already taken into consideration and serve as components contributing to the housing unit value and household income. In contrast, student needs adjustments serve to increase the explanatory power of the items because they provide no overlapping of issues and instead present a more accurate portrayal of the burden faced by each school district. Thus, the importance of the relationship between PPE and the independent variables depends upon the relative size of each adjustment. In addition, conclusions regarding the relationship of school district wealth and school district spending are affected by the type and nature of the measurement of school district wealth and the adjustment employed.

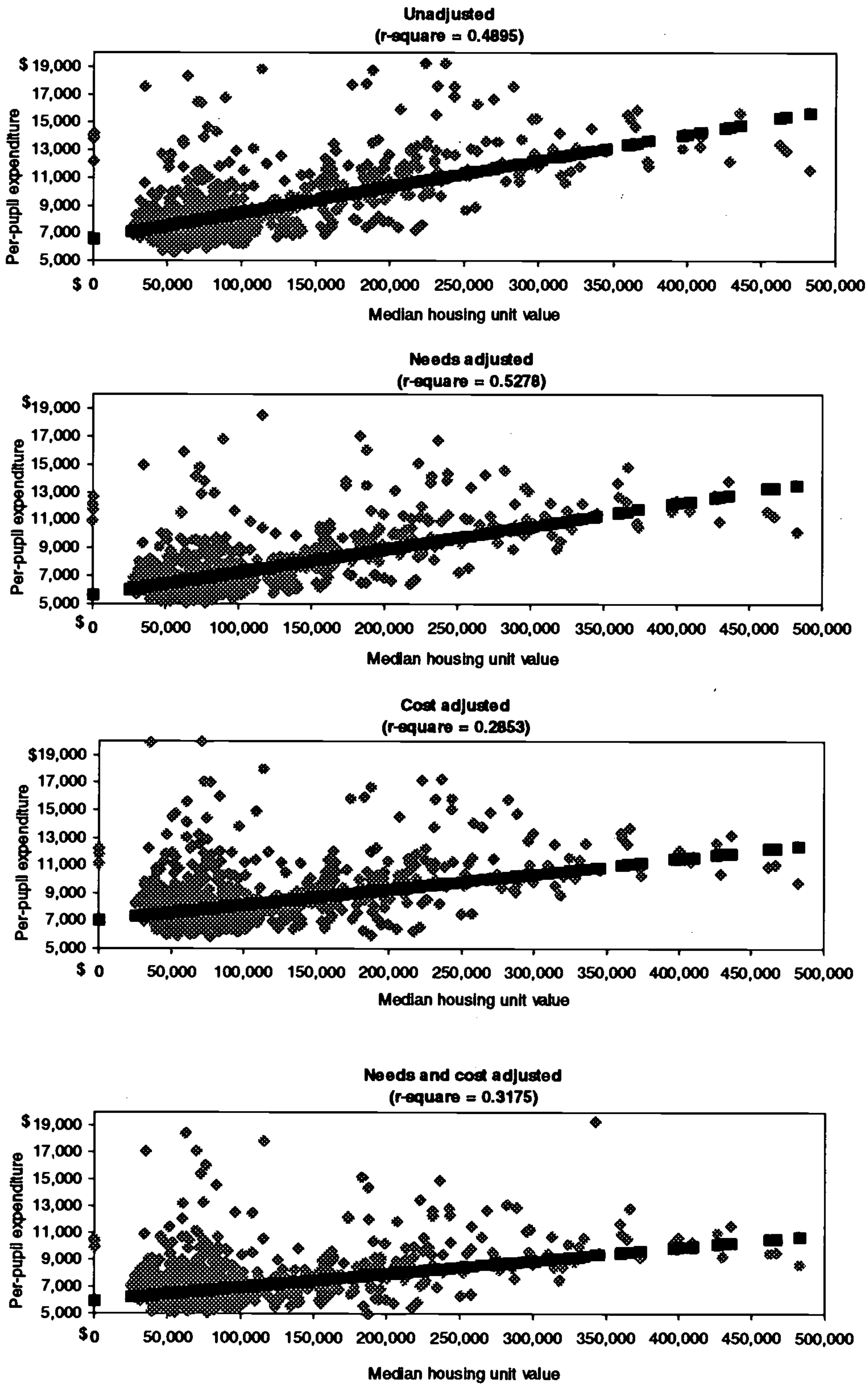
A second dynamic that is interesting to note, is that median housing unit value has a larger effect

Figure 5.—Regression analysis of per-pupil expenditure by median household income: Analysis of New York State including New York City



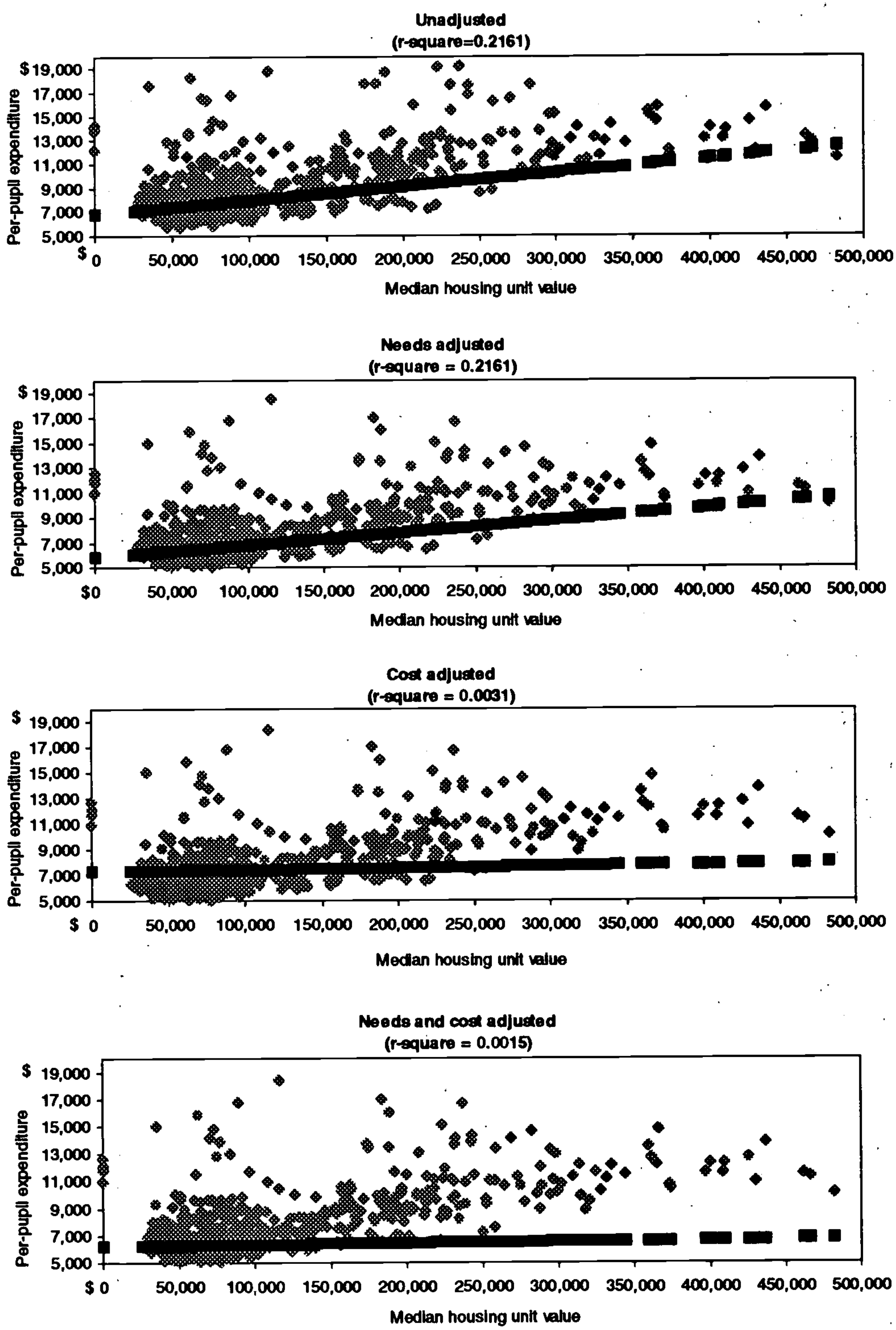
SOURCE: Special tabulation by authors from the Common Core of Data (CCD) using only New York State.

Figure 6.—Regression analysis of per-pupil expenditure by median housing unit value: Analysis of New York State excluding New York City



SOURCE: Special tabulation by authors from the Common Core of Data (CCD) using only New York State.

Figure 7.—Regression analysis of per-pupil expenditure by median housing unit value: Analysis of New York State excluding New York City



SOURCE: Special tabulation by authors from the Common Core of Data (CCD) using only New York State.

with per-pupil expenditure when New York City is excluded, but median household income has a larger effect when New York City is included. This difference may be attributed to the fact that New York City real estate is exceedingly expensive and not necessarily reflective of the city population's ability to financially support education. Household income, in contrast, is not as inflated when compared to the rest of the state.

In summary, results presented in this case study demonstrate the varying impact different adjustments may have depending upon what geographic locations are included within the data set and what measures and types of analyses are employed within one's work. Clearly, no uniform conclusions can be reached. Measures of equity do not always increase or decrease depending on the adjustment employed. Instead, these results indicate one's need to be aware of the basis for the adjustments and the power they hold when considering whether or not to employ them in one's work.

> *. . . student needs adjustments serve to increase the explanatory power . . . because they provide no overlapping of issues and instead present a more accurate portrayal of the burden faced by each school district.*

Conclusion

This study investigated a number of methods to measure and adjust for contextual variations in the cost of education based on the student population served and the costs experienced with different geographic regions. The paper began by identifying and defining various adjustments. These adjustments included a weighted pupil model and three cost-of-education indices that have been developed in prior research [i.e., Average-teacher-salary index (Barro,1992); Cost of Living Index (McMahon and Chang, 1991); Teacher Cost Index (Chambers and Fowler, 1995)]. These indices are related and generally provide similar cost estimates across states. However, some interesting variations emerge when comparing indices. These variations suggest that in some localities, teacher costs are more strongly influenced by particular features than others (e.g., cost of living versus teacher preparation and experience, or hedonic considerations). Thus, not only is there great diversity in funding, but there is diversity across local communities in the types of characteristics that influence this diversity.

The second section of the paper then defined and compared four types of equity measures previously established in the literature (i.e., coefficient of variance, the Gini coefficient, the McLoone Index, and a slope coefficient). Lastly, the final section of this paper presented a case study which applied the weighted pupil model and the TCI index to equity analyses to determine what impact these adjustments may have upon financial analyses.

Results from the study indicate that adjustments may impact results in a variety of fashions depending on the information included in the data set and the type of analyses conducted. For the state of New York, the TCI adjustment appeared to have a far more significant impact on the analysis than the needs adjustment. This may not be true in other states. In New York, the McLoone index also appeared to provide some provocative insight that the variance and Gini coefficient did not present. Once again, this is in part a function of the data sets and adjustments used and may not always appear.

These findings illustrate is the sensitivity of equity analyses and the varying and significant impact of cost and student need adjustments on the conclusions. One must be mindful of the power of these cost and student need adjustments and thoughtful in their utilization.

References

Alexander, N.A. "The Growth of Education Revenues Between 1982-83 to 1991-92: What Accounts for Differences Among States?" In W.J. Fowler Jr., Ed., *Developments in School Finance, 1995.* Washington, D.C.: National Center for Education Statistics, 1996.

Barro, S.M. *Cost-of-Education Differentials Across States.* Revised draft report. Washington, D.C.: SMB Economic Research, Inc., 1992.

Chaikind, S., L. C. Danielson, and M. L. Braun. "What Do We Know About the Costs of Special Education? A Selected Review," *The Journal of Special Education* 26 (4) (1993): 344–370.

Chambers, J. "Public School Teacher Cost Differences Across the United States: Introduction to a Teacher Cost Index (TCI)." In W.J. Fowler Jr., Ed., *Developments in School Finance, 1995.* Washington, D.C.: National Center for Education Statistics, 1996.

Chambers and Fowler. *Public School Teacher Cost Differences Across the United States: An Analysis to Develop a Teacher Cost Index (TCI).* Washington, D.C.: National Center for Education Statistics, 1995.

McInnis v. Shapiro 293 F. Supp. 327 (1968)

McMahon, W. and Chang, S. *Geographical Cost-of-Living Differences: Interstate and Intrastate, Update.* Normal, IL: Center for the Study of Educational Finance, MacArthur/Spencer Special Series of November 20, 1991.

Monk, D.H. *Educational Finance: An Economic Approach.* New York: McGraw-Hill, 1990.

Orland, M.E. and Cohen, C.E. "Meeting the Challenge of Devolution: How Changing Demographic and Fiscal Contexts Affect State Investments in Education." In W.J. Fowler Jr., Ed., *Developments in School Finance, 1995.* Washington, D.C.: National Center for Education Statistics, 1996.

Rothstein, R. "Summary of Where's the Money Gone: Changes in the Level and Composition of Education Spending (1967–1991)." In W.J. Fowler Jr., Ed., *Developments in School Finance, 1995.* Washington, D.C.: National Center for Education Statistics, 1996.

Verstagen, Deborah. "School Finance Litigation Across the United States: What are State's Core Finance Data Needs?" Unpublished report. January 1998.

A Proposal for Collecting School-Level Resource Data on the Schools and Staffing Survey

Julia B. Isaacs and Michael S. Garet,

American Institutes for Research

Stephen P. Broughman,

National Center for Education Statistics

About the Authors

Julia B. Isaacs is a senior research analyst at the Pelavin Research Center of the American Institutes for Research (AIR). Since joining AIR in February 1995, she has conducted research on public and private school finance, co-authoring four NCES working papers on this subject. She also has worked closely with the Head Start Bureau, serving as task leader on projects to revise the Head Start grant application system and to write program guidance on implementing Head Start Program Performance Standards. For 10 years she was a Principal Analyst at the Congressional Budget Office, responsible for analyzing the budgetary effects of legislative proposals in the areas of food stamps, child welfare, and child care. She holds a master's degree in Public Policy from the University of California at Berkeley.

Michael S. Garet is a senior research analyst at the American Institutes for Research (AIR). He earned his Ph.D. at the Massachusetts Institute of Technology, and he has served on the faculties of Stanford and Northwestern Universities. He is currently the deputy director of a large-scale national evaluation of the Eisenhower Professional Development Program, being conducted for the U.S. Department of Education. He has written on research methods, school organization, educational finance, and curriculum reform in mathematics and science.

Stephen P. Broughman has been an employee of the National Center of Education Statistics (NCES) since 1992 where his duties include being the project officer for the Private School Survey (PSS) and having the lead responsibility for the private school component of the Schools and Staffing Survey (SASS). He obtained a Bachelor's of Science and a master's degree in economics from Virginia Commonwealth University, and a master's degree in applied mathematics and statistics and a Ph.D. in economics from the State University of New York at Stony Brook. His current research interests include private school issues and teacher supply.

A Proposal for Collecting School-Level Resource Data on the Schools and Staffing Survey

Julia B. Isaacs and Michael S. Garet,

American Institutes for Research

Stephen P. Broughman,

National Center for Education Statistics

Introduction

Since 1987, the National Center for Education Statistics (NCES) has collected national data on the characteristics of public and private schools through periodic administrations of the Schools and Staffing Survey (SASS), which is scheduled to be administered for the fourth time in 1999–2000. The overall objective of SASS is to provide a detailed and comprehensive picture of American elementary and secondary education, through an interrelated set of questionnaires sent to local education agencies (districts), schools, principals, and teachers. Analyses of the existing SASS data have benefited from the linkages across these different components of the SASS. But analyses have been constrained by the limited information collected on certain critical issues—one of them being school resources or finances.

This paper reports on an exciting possibility being explored by NCES—a proposal to expand the resource and finance data collected as part of the 1999–2000 SASS. The proposal, which currently is being tested for feasibility, has two major components. The first part of the proposal is to collect more detailed information about staffing resources in the schools in the SASS sample in order to improve understanding of how schools allocate personnel resources, which account for more than 85 percent of expenditures in most school sites (Levine, Chambers, Duenas, and Hikido, 1998). The second component involves gathering expenditure data for individual schools in the SASS sample. This represents a departure from existing educational finance data collections, such as the National Public Education Financial Survey (NEPFS) or the Annual Survey of Local School Governments—Schools (Form F-33), which collect data at the district level, but not for individual schools. Moreover, the SASS finance survey would represent the first collection of traditional finance data from a nationally representative sample of private schools in 20 years.[1]

[1] The last national survey of private schools was conducted in 1978–79. See McLaughlin, D. H. and Wise, L.L. 1980. *Nonpublic Education of the Nation's Children.* Technical Report 9. Palo Alto, CA: American Institutes for Research.

In this paper, we discuss the rationale for the collection of school-level resource and expenditure data, and we outline the proposals that have been developed to collect such data as part of SASS. In the first section of the paper, we review the kinds of policy issues that could be addressed with improved resource and expenditure data. In the next section, we present an overview of two approaches to collecting improved school-level resource data—a Resource Cost Model (RCM) approach and a traditional finance approach. In the third section, we describe the elements of these two approaches that may be incorporated in the 1999–2000 SASS. Finally, the paper concludes with a discussion of how the proposed additions to SASS fit in with other NCES efforts to expand knowledge about the allocation of resources at the school level.

Underlying Policy Issues

The effort to collect expanded school-level resource and expenditure data has been undertaken by NCES in response to the demand of education finance researchers for improved data to address a number of important education policy issues. A review of the literature, combined with a discussion among a half-dozen prominent education finance experts,[2] suggests that the collection of improved resource and expenditure data would support analysis of the types of policy issues outlined in table 1 and discussed briefly below.

The effort to collect expanded school-level resource and expenditure data has been undertaken by NCES in response to the demand of education finance researchers for improved data . . .

Resource Allocation and Productivity Issues

One of the most hotly debated questions of educational policy concerns the effects of school resources on student outcomes.[3] Much of the research in this area has relied on district-level data on per-pupil expenditures to measure school resources, but it is clear that this measure provides only a very crude index of the educational resources allocated to particular students and programs. In order to make progress in understanding the effects of resources on student outcomes, we need a much better understanding of the ways resources are used to produce education services. In particular, we need to understand how schools differ in the resources available and the ways these resources are allocated to different services and programs (i.e., special education or bilingual education). Furthermore, we need to understand how district-level resources (i.e., resources in curriculum coordination and professional development) support school-level activities.

Costs and Effects of Policy Initiatives

Closely related to issues of resource allocation and productivity are questions concerning the costs and effects of policy initiatives. Better data are needed to evaluate such questions of interest as the effects of finance reform on district allocations to schools, the costs of modifying school programs to implement new standards in mathematics and science, the cost of new school designs (for example, the New American School designs), and the costs of new forms of professional development (i.e., mentoring, networks, and study groups).

Equity and Adequacy

Educational equity has been a major focus of both policy and research interest. Most studies of

[2] These education finance experts included Matthew Cohen (Ohio Department of Education), Margaret Goertz (University of Pennsylvania), Richard Laine (Illinois State Board of Education), David Monk (Cornell University), Allen Odden (University of Wisconsin), and Leanna Steifel (New York University). Also present for the discussion on January 9, 1998, were NCES Associate Commissioners Paul Planchon and Martin Orland, as well as Steve Broughman, William Fowler, Frank Johnson, Daniel Kasprzyk, and Mary Rollefson of NCES, and Jay Chambers, Michael Garet, Julia Isaacs, Lauri Peternick, and Joel Sherman of the American Institutes for Research.

[3] See Hedges, L.V., Laine, R.D., and Greenwald, R. April, 1994. "Does Money Matter? A Meta-Analysis of Studies of the effects of Differential School Inputs on Student Outcomes." *Educational Researcher.* 23 (3):5–14; and Hanushek, E.A. Summer, 1997. "Assessing the Effects of School Resources on Student Performance: An Update." *Educational Evaluation and Policy Analysis.* 19(2).

Table 1.—Policy issues driving demand for school-level resource data
Resource allocation and productivity How do schools allocate resources? How much is spent on instruction and how much on administration? What is the relationship between school expenditures and student outcomes?
Costs and effects of policy initiatives How does Initiative X affect school staffing patterns and expenditures?
Equity and adequacy How much variation is there in per-pupil expenditures among schools?
School-based management What data are needed to inform school management decisions?
Accountability Are resources under grant Y being spent as intended? How do resource allocations in school Z compare with allocations in similar schools?
Congressional interests and public inquiries How much is spent on administrative expenditures at the school site and the central office?
SOURCE: American Institutes for Research.

educational equity have used district-level data, and these studies have documented wide disparities in per-pupil spending across districts within a state as well as across states. Of significant interest, but much less studied, is whether resources are distributed in an equitable manner across schools within a district.

In addition to examining equity issues, researchers have also focused on the *adequacy* of resource provision—that is, the minimum resources required to insure that all students have an appropriate opportunity to learn. Differences in student populations affect the level of resources that are required to provide an adequate level of educational services. For example, students with limited English proficiency (LEP) or in need of special education may require more services and, thus, more resources than other students.

School-Based Management

Recent reforms in school organization have sought to increase the degree to which staff at the school-site level are involved in making key educational decisions. But most districts lack the capacity to provide detailed school-level financial and resource data to support decision making. To the extent resource allocation decisions are made at the school level, school staff require detailed information on school budgets and expenditures. Such information is critical, for example, to support principals and teachers in understanding the budgetary tradeoffs involved in allocating resources to types of staff—for example, teachers, teacher aides, and clerical staff. In making decisions about such allocations, schools may also require "benchmark" information about the staffing allocations in high-performing schools serving similar student populations.

Accountability

One key function of information on school expenditures is to determine whether resources are being spent as intended. Such information is required to inform parents and community members on what is happening at the school-level (in charter schools, choice programs, etc.), as well as to inform state and federal agencies and private foundations on the ways in which resources for special programs are deployed.

Congressional Interests and Public Inquiries

NCES often is asked to address questions of interest to policy-makers and other audiences. For example, in the Improving America's Schools Act of 1994, Congress directed the Commissioner of NCES to study methods to gather information about spending for administration at the school and district levels. In another example, the international Organization for Economic Cooperation and Development (OECD) requests NCES to report the total amount spent per year on elementary and secondary education in the United States, including spending in both public and private schools. Another frequently asked question of NCES concerns how much is spent on instructional technology. Improved resource and expenditure data are required to answer these and other inquiries directed to NCES.

Two Approaches to the Collection of School-Level Resource Data

During 1997, researchers at the American Institutes for Research (AIR) were asked by NCES to develop two approaches to collecting data about the allocation of resources in public and private schools—a Resource Cost Model (RCM) approach and a traditional finance approach. Work on the RCM approach was undertaken by a team of AIR researchers in Palo Alto, California, under the leadership of Jay Chambers, while work on the traditional finance approach was undertaken by a team of researchers at AIR's Pelavin Research Center in Washington, D.C., under the leadership of Joel Sherman. Each of these two approaches is summarized below, first in general terms, and then as a specific data collection strategy developed by the AIR research teams.

Overview of the Resource Cost Model (RCM) Approach

The RCM approach is essentially a bottom-up approach to the analysis of school resources. In contrast to the more traditional accounting systems that study resources by dividing a total budget down into fine-grained spending categories, the RCM approach starts at the level of service delivery and builds up to total costs by aggregating specific resources used in an educational program. It requires four basic steps: 1) specifying the types of physical ingredients (teachers, books, etc.) used in an educational program; 2) measuring the intensity of these resources by quantifying them; 3) assigning prices to the specific physical ingredients; and 4) using the price data to aggregate resources across the entire program to determine overall program costs. The four steps in this process are illustrated in the four columns of table 2, which shows how staff resource costs could be measured in Rosemont School, a hypothetical elementary school serving 400 students. Although in this example the educational program under analysis is an entire school, the RCM approach also can be used very effectively to study resources associated with a specific program within a school, such as a special education program or compensatory education program.

The RCM approach is essentially a bottom-up approach to the analysis of school resources . . . [It] starts at the level of service delivery and builds up to total costs by aggregating specific resources used in an educational program.

The level of detail and scope of data collection required by the RCM approach depends to a large extent upon decisions made during the first step outlined above: determining the categories of resources under study. In the example shown in table 2, data are collected for staffing resources only, across a broad range of staff ranging from teachers to custodians. A more streamlined model might split staff

Table 2.—Staff resources at Rosemont School: physical ingredients, quantities, prices, and total costs

Physical ingredient	Quantity (Full-time equivalent)	Price per unit, in dollars*	Total cost, in dollars
Principal	1.0	$96,000	$96,000
Assistant principal	1.0	74,000	74,000
Instructional coordinator	0.0	73,000	0
Librarian	0.8	64,000	51,200
Library aide	0.5	28,000	14,000
Teacher 1 (MA)	12.0	54,000	648,000
Teacher 2 (BA)	6.0	49,000	294,000
Music/arts teacher	1.2	49,000	58,800
Physical education teacher	1.5	44,000	66,000
Special education aide	2.0	20,000	40,000
Bilingual English as a Second Language aide	2.0	23,000	46,000
Other teacher aide	4.0	20,000	80,000
Counselor	1.0	50,000	50,000
Nurse	0.4	67,000	26,800
Social worker	0.2	53,000	10,600
Psychologist	0.1	65,000	6,500
Speech therapist	0.1	64,000	6,400
Health aide	1.0	33,000	33,000
Secretary/Clerical staff	5.0	33,000	165,000
Lunch-room attendant	1.0	28,000	28,000
Custodian	2.0	28,000	56,000
Total	**42.8**	—	**1,850,300**

— Not applicable.

*Prices include salaries at Rosemont School, multiplied by a 0.28 fringe benefit rate.

NOTE: Rosemont is a hypothetical elementary school with 400 students. Staffing costs per student are $1,850,300 ÷ 400, or $4,626.

SOURCE: American Institutes for Research.

among fewer categories than the 21 categories shown in table 2, or might be limited to instructional and administrative staffing resources under the assumption that variations in intensity of these staffing resources have the most significant effect on educational outcomes. An expanded model might list more categories of teachers (i.e., bilingual teachers, special education teachers, general education teachers by subject matter, Title I teachers, or reading specialists), or might collect data for more categories of staff (i.e., physical/occupational therapists, audiologists, or maintenance workers). Ideally, one might want to compile a detailed listing of all individuals working in a school (or all staff involved with in a specific educational program under study), and collect selected data about each staff member. In addition, a full-fledged model would include data on textbooks, computers, science equipment, facilities, and other non-staffing resources.

Once the staff and other resources under study are identified, the next step is to measure the intensity of resources used. Staffing resources in each staffing category can be measured in a variety of ways: numbers of full-time and part-time staff, full-time equivalents (as in table 2), hours of labor, days of service, etc. Quantifying staff contributions can be complicated when staff are shared among several schools. An itinerant music teacher, for example, who works 3 days in Rosemont school and 2 days in Greenwood school is a full-time employee, but should be counted as an 0.6 full-time equivalent (FTE) employee when measuring staff resources at Rosemont. (In the 8th line of table 2, Rosemont is reported as having 1.2 in FTE music/arts teachers, including the itinerant music teacher and an itinerant art teacher who also teaches 3 days a week). In Rosemont, as in many schools, staff who perform student support functions are most likely to be shared among several schools. For example, the nurse is at Rosemont 2 days a week (0.4 in FTEs), the social worker 1 day a week (0.2 in FTEs), and the school psychologist and speech pathologist are each assigned to Rosemont for only half a day per week (0.1 in FTEs). As this example demonstrates, the measurement of staff in full-time equivalents, though difficult for some respondents to do, provides a more accurate measure of staff resources than simpler measures, such as the number of part-time staff.

The final challenge involves attaching prices to each resource. Attaching prices to resources allows the analyst to aggregate resources across categories. One approach is to take actual prices, based on salary *and benefit* information for staff, and actual prices paid for non-staff resources. The example in table 2 assumes that actual salaries are used in the analysis of staffing resources at Rosemont schools, with a 28 percent fringe benefit rate used to allocate employee benefits across all categories of staff. An alternative approach is to assign a standard set of prices, drawn from national data on salaries, benefits, and prices. The advantage to this latter approach is that it allows researchers to compare the intensity (quantity) of resources used across educational settings, measured separately from variations caused by differences in local prices.[4] Such a comparison is critical to answering the question, do variations in quantities of services make a difference?

Initial Proposal for Collecting RCM Data through SASS

A set of specific recommendations for collecting RCM data as part of SASS are set forth by Levine, Chambers, Duenas, and Hikido (1998) in a recently published NCES Working Paper (NCES 97–42). In their proposal, Levine et al. focus primarily on the collection of staffing resources at the school site level.[5] Specifically, they recommend that data on staffing resources be collected through Staff Listing Forms, to be filled out by the school principal or school secretary. Their proposed Staff Listing Forms would collect information for all individuals in the school, including information on the number of hours per week spent in various teaching, administrative, and support positions. This would allow fairly accurate measures, in hours per week, of the intensity of staffing resources devoted to various school-related activities. The proposed forms represent a substantial expansion over the existing Teacher Listing Forms, which collect more limited data on teaching assignments and which have been used to generate the sample of teachers surveyed through SASS, but have not been used for analytical purposes.

Levine et al. focus primarily on the collection of staffing resources at the school site level . . . [by collecting data] through Staff Listing Forms, to be filled out by the school principal or school secretary.

Because of their interest in comparing the intensity of resources across schools while controlling for variations in local prices, Levine et al. propose attaching national prices to the data on staffing resources collected through the Staff Listing Forms. The national price data would be drawn from the samples of teacher and administrator salaries that are already collected through other components of

[4] For example, assume a teacher with a master's degree and 5 years experience and training in mathematics receives $33,000 in compensation (salaries and benefits) in small, rural school districts in Idaho, $44,000 in large, urban districts in California, and $39,000 nationally. Use of the national price of $39,000 in analyzing resource costs in schools in Idaho and California will allow better measurement of the real differences in staff resources across different schools.

[5] In recognition of the potential burden posed by the collection of detailed resource data, Levine et al. do not recommend collecting data about non-staffing resources at the school or about any resources at the central administrative offices.

the SASS, supplemented through a small amount of additional salary and benefit data collected in a new, short survey to school business officers in the SASS sample.

Overview of the Traditional Finance Approach

The traditional finance approach relies on expenditure data collected through the accounting system of the public school district or private school. Expenditure data are typically collected and analyzed by function, object, and program. While the accounting systems differ across districts and states, many systems employ a core set of functions, including instruction, administration, student and instructional support, and operations and maintenance. Accounting systems frequently record a small set of objects, including salaries, supplies, and contracted services. In addition, many accounting systems account for expenditures by program—for example, regular education, vocational education, and community programs.

While the accounting systems differ across districts and states, many systems employ a core set of functions, including instruction, administration, student and instructional support, and operations and maintenance.

The function/object/program framework forms the basis of a number of existing school finance surveys. For example, the National Public Education Finance Survey (NPEFS), requires all states to report expenditures across a function by object by program matrix. To guide states in the collection of these data, NCES has developed a national accounting guide, *Fundamentals of Financial Accounting for Local and State School Systems* (NCES, 1990).

There are a number of challenges involved in collecting traditional expenditure data at the district and school levels. First, many districts do not follow the NCES accounting handbook, and there is considerable variation across districts in the ways particular expenditures are treated. For example, principals' salaries are classified as administrative expenditures by NCES, but as instructional expenditures in many school districts. NCES also makes a distinction between "instructional support services," such as library services and professional development, and "student support services," such as health, counseling, and attendance services; however, some states ask districts to classify all such services into one general "support services" category. One challenge in collecting district finance data, therefore, is to align the local accounting system with the standard NCES definitions.

The effort to collect and report finance data at the school-level must confront a second challenge as well: the collection of school-level finance data requires districts to report data associated with a selected school—despite the fact that the district-wide accounting systems of many districts do not directly track expenditures to specified school sites. The district/school problem is not an issue for private schools, although in some private schools, analogous difficulties may arise distinguishing school expenditures from expenditures for an affiliated church.

One approach to resolving these challenges to the collection of school-level finance data is to use software packages, such as "In$ite, The Finance Analysis Model for Education" ™ developed by Coopers and Lybrand, to reclassify the data gathered in local school accounting systems to fit with a standard set of accounting categories (Cooper, Sampiere, and Speakman, 1994). Under such software packages, districts are provided with an array of algorithms that can be used to allocate centrally-billed expenditures (such as centrally-billed utilities or itinerant teachers) to specific school sites. For example, expenditures could be allocated by square foot of building space, student enrollment, number of students transported, etc. Such a system has the advantage of drawing from a district's existing administrative records, but requires participating school districts to purchase the software package, and take the time in the first year of use to translate or "map" the data in a local accounting system to the predefined functions, programs and school site locations used in the software package.

Initial Proposal to Collect Finance Data through SASS

Isaacs, Best, Cullen, Garet, and Sherman (1998) have developed a proposal for collecting public school expenditure data using a mailed survey conducted as part of SASS, along with a comparable proposal for collecting expenditures data for private schools (Isaacs, Garet, and Sherman, 1997). In both questionnaires, respondents are asked to report expenditures across a simplified set of functions and objects, as shown in table 3. To reduce burden on respondents, the functions are fairly broad. For example, data on expenditures for instructional support and student services are collected as one broad category, rather than as two separate categories as in the NCES accounting manual. Likewise, three different NCES administrative functions (General, Business, and Central Support Services) have been collapsed into one overarching administrative function. Finally, expenditures for equipment, benefits, and long-term debt are not collected in as much detail as the expenditures for salaries and other current operating expenditures.

For the public schools, the function by object data are collected in three parts: expenditures for the district as a whole; central-office expenditures; and school-based expenditures at a selected school in the SASS sample . . .

The questionnaires would be sent to the public school district business officer and the private school principal (who, in large schools, would forward it to the business manager). For the public schools, the function by object data are collected in three parts: expenditures for the district as a whole; central-office expenditures; and school-based expenditures at a selected school in the SASS sample (e.g., expenditures at school sites). Central office expenditures include general administrative expenditures (e.g., the superintendent's office), business administration, and coordination of support services, operations and maintenance, etc. All expenditures other than expenditures for central-office operations are defined as school-based expenditures.

To accommodate the diverse capabilities of district accounting systems, school-based expenditures are reported in two sections:

- **Section A: Actual Expenditures at Selected School.** Districts are asked to report actual expenditures for the selected school in Section A to the extent that such expenditures are known, and tracked to that specific school site. Respondents are instructed to report zeros in Section A if the district's accounting system does not track any expenditures to specific school locations.
- **Section B: School-level Expenditures Not Assigned to Any Specific School.** Districts are to use Section B to report any expenditures for school-based services that are not assigned to any particular school or location. This might include itinerant staff (e.g., itinerant music teachers), personnel or materials used in schools on an "as-needed" basis (e.g., psychologists, maintenance workers), or personnel or materials associated with school-based services but which are accounted for under a central office location (e.g., nurses coded to central location, centrally-billed utilities). Section B includes all expenditures other than central-office expenditures if a district's accounting system does not track any expenditures to specific school locations.

An estimate of the operating expenditures for each school in the district's sample may be obtained by summing the reported expenditures under Section A: *Actual Expenditures at the Selected School* and the school's proportional share of overall district expenditures under Section B: *School-level Expenditures Not Assigned to Any Specific School.* To ease response burden and maintain data comparability, the questionnaire does not ask districts to carry out the calculations necessary to allocate a share of Section B: *School-level Expenditures Not Assigned to Any Specific School* to each target

Table 3.—Collection of expenditure data by functions and objects

Functions	Objects				
	Salaries	Supplies and contracted services	Equipment	Benefits	Facilities, debt
Instruction	[1]✔	✔	[2]✔		
Instructional support and student services	[1]✔	✔			
Administration	✔	✔			
Plant/maintenance	✔	✔			
Food service	✔	✔	✔		
Transportation	✔	✔	✔		
Other	✔	✔	✔		
Total	✔	✔	✔	✔	✔

[1] In the public school expenditure survey, salaries for instruction and support services are reported separately for special education and regular education.

[2] Instruction-related computers.

SOURCE: American Institutes for Research.

school. Instead, enrollment and other basic data for the district and the selected school are collected in another item in the survey, allowing NCES to perform the necessary calculations during data cleaning and analysis.[6]

An example of how expenditure data might be reported appears in table 4, which displays expenditures for Rosemont School, the hypothetical school with the staffing resources shown in table 2. In this example, as in many schools, instructional expenditures primarily consist of salary and benefit expenditures for teachers at the school, but also include some centrally-billed salary and benefit expenditures (a $50,000 allocation for Rosemont's itinerant music and art teachers), as well as expenditures for instructional supplies.[7] In total, instructional expenditures account for 60 percent of all school expenditures.

Expenditures for instructional support and student services are much lower (14 percent of the total, as shown in the second row of table 4), but include significant expenditures for staff who are shared among several schools and accounted for centrally (i.e., Rosemont's allocation of salaries and benefits for the shared librarian, nurse, social worker psychologist, and speech pathologist). In this example, all administrative expenditures are tracked to the specific school.

The $1,848,000 total in expenditures for salaries and benefits shown in the last row of table 4 is within $2,300 of the salaries and benefits calculated under the RCM approach illustrated in table 2. The $2,300 difference reflects differences in shared staff—in this example, the finance model reports lower expenditures for shared teachers and support staff, but higher expenditures for centrally-billed mainte-

[6] Depending on the purpose of the analysis, central-office expenditures can also be allocated to target schools based on student enrollment or other criteria.

[7] Note that the $50,000 allocation for Rosemont's itinerant music/art teachers differs from the $58,500 resource cost figure derived from staff FTEs reported in table 2. The $50,000 allocation is a proportion of the district's total spending on itinerant teachers: in this example, 10 percent of total district spending of $500,000 for itinerant teachers because Rosemont student enrollment is 10 percent of the district's total enrollment. The precision of the reported expenditure data is diminished by this need to use estimated allocations for centrally-billed expenditures such as itinerant teachers. The data would be more accurate if Rosemont's accounting system tracked all expenditures to the school—including itinerant teacher salaries, prorated to each school on the basis of time spent at the school—but few accounting systems can do so at this time.

nance staff (who are not listed as school staff in table 2 because maintenance staff in this district are not assigned to specific schools).[8] Differences between salary expenditures calculated under the RCM approach and the finance approach would probably be larger with actual data collected under normal circumstances—and the comparison could not be made very easily if the RCM resource estimates were calculated with national prices rather than actual salaries for each staff member. A final difference between the two approaches is that the finance data reported in table 4 include expenditures for supplies and contracted services, which were, of course, not included in the staff resources reported in table 2.[9]

Recommendations for Collection of School-Level Data through SASS

In January 1998, a group of education finance experts met with staff from NCES and AIR to discuss the RCM and traditional finance approaches to the collection of school-level data.[10] During a day-long meeting devoted to analyzing both approaches, the technical work group recommended to NCES that both types of data be collected as part of the 1999–2000 SASS: staffing data in line with the RCM approach and expenditure data in line with the traditional finance approach.

Traditional finance data provide basic information on differences in total expenditures and expenditures per pupil across schools . . . [while the] RCM data would move the emphasis closer to the point of the instruction.

The two types of data are expected to serve complementary purposes. Traditional finance data provide basic information on differences in total expenditures and expenditures per pupil across schools, as well as information to address basic resource allocation questions, such as the allocation of expenditures across functions (i.e., between instruction and administration) and between the school site and the central office. To answer more detailed questions regarding how dollars are spent, and how services are delivered, researchers would like the more detailed staffing data collected under the RCM approach. RCM data would move the emphasis closer to the point of the instruction and allow an analysis of differences in resource use between different educational programs, such as special education or compensatory education.

An integrated collection of traditional finance data and staffing data collected under the RCM approaches provides certain analytical benefits. For example, using expenditure data, analysts might estimate differences between public and private schools in per-pupil spending for instructional salaries. If differences in per-pupil spending are observed, RCM staffing data might then be used to determine how much of the observed difference in spending can be explained by differences in the intensity of staff resources (i.e., by the number of regular and special education teachers, special education aides, bilingual/ESL teacher aides, and other teacher aides).[11] Public/private differences in staff quality or staff pricing (salaries and benefits) would also need to be examined; and such differences could begin to be explored, at least for teachers and administrators, through other components of the SASS.

[8] More specifically, the finance model allocates expenditures for itinerant teachers, shared support staff, and other centrally-billed expenditures on the basis of school enrollment, building square footage, or other such parameters, while the resource cost model allocates expenditures on the basis of time spent in the school (measured in table 2 in terms of full-time equivalents (FTEs), but in simpler models, simply as counts of full-time and part-time staff).

[9] Because the proposed public school expenditure survey also collects data on central-office expenditures, these can also be included in reports of per-pupil expenditures. For example, the note to table 4 suggests that per-pupil expenditures at Rosemont school are $5,285 when limited to school-based expenditures, and $5,730 per student when including a share of central-office expenditures.

[10] See footnote 2 for a list of meeting participants.

[11] Such an analysis might be conducted by estimating a regression model predicting per-pupil spending on instructional salaries based on a public/private indicator variable, staff hours per pupil for types of instructional staff, and the interaction of the indicator variable and the measures of staff hours.

Table 4.—Expenditure data for Rosemont School, by function, object, and location

Functions	Salaries and benefits[1]	Supplies and contracted services	Total for function	Function as percent of total school expenditures
Instruction				
Expenditures tracked to Rosemont[2]	1,174,000	45,500		
Rosemont allocation of centrally-billed expenditures[3]	50,000	100		
Subtotals	*1,224,000*	*45,600*	*1,269,600*	*60*
Instructional support and student services				
Expenditures tracked to Rosemont[2]	163,000	25,900		
Rosemont allocation of centrally-billed expenditures[3]	100,000	100		
Subtotals	*263,000*	*26,000*	*289,000*	*14*
Administration				
Expenditures tracked to Rosemont[2]	269,000	1,000		
Rosemont allocation of centrally-billed expenditures[3]	0	0		
Subtotals	*269,000*	*1,000*	*270,000*	*13*
Operations and Maintenance				
Expenditures tracked to Rosemont[2]	56,000	63,000		
Rosemont allocation of centrally-billed expenditures[3]	8,000	60,000		
Subtotals	*64,000*	*123,000*	*187,000*	*9*
Food service				
Expenditures tracked to Rosemont[2]	0	0		
Rosemont allocation of centrally-billed expenditures[3]	28,000	60,200		
Subtotals	*28,000*	*60,200*	*88,200*	*4*
Total school-level expenditures	**1,848,000**	**255,800**	**2,103,800**	**100**

[1] Benefits are allocated across salaries assuming a constant 28 percent fringe benefit rate.

[2] Expenditures tracked to Rosemont are actual expenditures as reported by district accounting system.

[3] Rosemont allocations are based on school:district ratios of students, full-time equivalent teachers, square feet in buildings, and number of meals served.

NOTE: Rosemont is a hypothetical elementary school with 400 students. Per-pupil expenditures are $2,103,800 ÷ 400 students, or $5,260. In addition, per-pupil expenditures for central-office salaries (superintendent, finance, etc.) in Rosemont's district are $283, and per-pupil expenditures for central-office supplies and contracted services are $162, bringing total per-pupil expenditures (including central administration) to $5,705.

SOURCE: American Institutes for Research.

While recognizing the virtue of collecting both expenditure and RCM data as part of SASS, the technical work group was cognizant of the potential burden posed by both types of data collection. NCES staff responsible for overseeing the administration of SASS were particularly concerned that expansions to the existing Teacher Listing Forms might lower response rates and thus endanger the validity of the teacher sample. Moreover, the addition of a separate component on school finances might overwhelm the SASS, both in terms of response burden and budgetary costs. The technical work group therefore recommended that scaled-down versions of the RCM and traditional finance instruments be developed.

Recommendations for Collecting RCM Data Through SASS

The technical work group recommended that improved staffing resource data be collected by making relatively modest modifications to two sets of existing SASS instruments: the Teacher Listing Form and the Public and Private School Questionnaires. The work group recommended that, in a departure from the past, data from the Teacher Listing Form be entered into an analytical database, allowing researchers access to more detailed data about the complete set of teachers at each sampled school (i.e., the grade range taught, subject matter taught, full- or part-time status, ethnicity, status as a new teacher, and status as a teacher of students with limited English proficiency). If field testing is favorable, the Teacher Listing Form will be expanded to collect more information about part-time status (i.e., ¼ time or less, ¼ to ½ time, ½ to ¾ time) and more information about special education teachers (i.e., whether teaching in a self-contained special education classroom or serving as a resource teacher/specialist).

Because of concerns about response burden, the technical work group did not recommend a full-scale expansion of the teacher listing form to cover all staff in the school. Instead, they recommended that items in the existing Public School Questionnaire (and the corresponding items in the Private School Questionnaire) be expanded to ask more detailed questions about various categories of staff. The new categories under consideration are shown in **bold type** in table 5. As in past rounds of SASS, principals will be asked to report the number of staff in full-time and part-time positions for each category, a measure of staffing intensity which simplifies the burden for respondents, but reduces the precision of estimates for itinerant and other part-time staff.

The work group recommended that, . . . data from the Teacher Listing Form be entered into an analytical database, allowing researchers access to more detailed data about the complete set of teachers at each sampled school . . .

The current proposal does not include the collection of additional salary data directly through the SASS. (Some salary, but not benefit, information already is collected for a *sample* of teachers and the principal at each SASS school). This lack of emphasis on the collection of additional price data reflects, in part, the greater interest of researchers at the January 1998 meeting in staffing data than in price data. However, there also are other alternatives for collecting the price data.

For example, the Current Population Survey (CPS) collects annual data on salaries for a national sample of individuals classified by occupation, industry, and type of employer. Using CPS data, it is possible to obtain national estimates of the salaries earned by broad categories of workers that may be used as reasonable approximations of the salaries earned by different types of school staff. The average salary earned by secretaries employed in local governments, for example, might be used as an approximation of the average salary of school secretaries. And the average salary of cleaning and building service occupations employed in local governments might be used as an approximation of the average salary of school custodians.

In addition to information on salaries, information on the dollar value of staff benefits is required to attach appropriate prices to staff resources. The technical work group recommended that SASS explore the possibility of adding a SASS item on fringe benefit rates. This question would be added to the proposed expenditure survey and asked of school business officers, who would be asked, for the first time, to participate in the SASS.

Recommendations for Collecting Finance Data Through SASS

The consensus of the technical work group was that it was important to collect expenditure data in addition to the staffing data discussed above. These data are needed to determine per-pupil expenditures, as well as allocations across functions and between the central office and the school sites.

No concrete recommendations were made by the group regarding specific changes to be made to the public or private school instruments developed by Isaacs et al. Several researchers suggested, however, that the instruments be scaled down, perhaps

Table 5.—Staffing data proposed for SASS Public School Questionnaire

a. Principals
b. Vice principals and assistant principals
c. Instructional coordinators and supervisors
  c1 Special education coordinators, supervisors or administrators*
  c2 Other instructional coord. and supervisors, such as curriculum specialists
d. Library media specialists/librarians
e. School counselors
f. Other student support services professional staff *(in past this was one category; now proposed to be split among the five sub-categories below):*
  f1 Nurses*
  f2 Social workers*
  f3 Psychologists*
  f4 Speech pathologists*
  f5 Occupational or physical therapists, other professional staff*
g. Teachers *(not split among different categories because this information is collected elsewhere)*
h. Aides or Assistants
  h1 Library media center aides
  h2 Health and other non-instructional aides providing student support services*
  h3 Special education aides*
  h4 Bilingual/English as a Second Language teacher aides*
  h5 Other teacher aides such as kindergarten or Title I aides
i. Secretaries and other clerical support staff
j. Food service personnel*
k. Custodial and maintenance personnel, security personnel*
l. Other employees if cannot assign to any category above (formerly included food service, custodial and maintenance, and other)

*New category or sub-category under consideration for the Schools and Staffing Survey (SASS).

SOURCE: American Institutes for Research.

by further reducing the amount of detail collected with regard to equipment and long-term debt.

Work on refining the public school expenditure instrument is still underway. There may be more time for refining the expenditure instruments than time for the staffing instruments, because the expenditure survey should, ideally, be administered during the school year following the administration of the main SASS instruments. That is, if the SASS is administered in the fall of 1999, with questions about school characteristics pertaining to the 1999–2000 school year, the expenditure survey should be administered in the fall of 2000, when financial records of actual expenditures for 1999–2000 are available. In this way, the expenditure data would cover the same school year as the staffing data and other data on school characteristics.

The private school finance survey developed by Isaacs, Garet, and Sherman (1997) is ready for full-scale SASS field-testing, having undergone successful pilot tests in 17 private schools. In addition to collecting data on school expenditures, the private school finance survey includes items on income and contributed resources. In this latter item, respondents are asked to indicate, through simple check-off boxes, an estimate of the quantity of services and materials contributed by public agencies (e.g., student transportation, remedial instruction), religious institutions (e.g., space, custodial services, bookkeeping assistance), and parents and others (e.g., donated supplies or equipment, volunteer labor). Because of confidentiality concerns and distrust of the government, the reaction of the private school universe to this proposed addition to SASS will depend to a large degree on the ability of NCES to work closely with the major private school associations—several of which have been involved in survey development and have expressed interest in supporting NCES efforts to gather more of these types of data.

SASS Resource Data: An Incremental Step Forward

The proposal to include expenditure and resource components in SASS represents an important step forward in improving understanding of how resources are allocated within and among schools. SASS is well-suited to cross-state comparisons of school resource levels and resource utilization patterns, as well as national estimates, because the SASS sample design supports state-reliable estimates.[12] Furthermore, because SASS collects such a rich assortment of data on school characteristics, researchers will be able to analyze how resource allocations vary among schools with different programs and services, alternative forms of school organization, and varying student body characteristics. The SASS sample size is large enough to allow the data to be reported for specific sub-groups. For example, typical resource allocations could be reported for large public high schools in high-poverty urban areas, small public elementary schools in suburban areas, or Catholic elementary schools. SASS data on teacher and principal characteristics can be used to begin to add some understanding of how teacher and administrator quality, as measured by education and years of experience, are related to resource allocations. Finally, the collection of comparable staffing and expenditure data for public and private schools will enable powerful comparisons between the public and private sectors.

SASS is well-suited to cross-state comparisons of school resource levels and resource utilization patterns, as well as national estimates, because the SASS sample design supports state-reliable estimates.

It is important to be aware, however, of the limitations of the proposal for collecting school-level resource and finance data through SASS, and of the need for ongoing work on complementary data collection and data analysis strategies to improve understanding of school-level resources. First, the instruments discussed in this article are still undergoing refinement and have not yet been submitted to full-scale field testing in a large sample of schools. More will be known about the feasibility of this data collection strategy after completion of the SASS field testing scheduled for fall of 1999.

Second, administration of a national survey such as SASS is only one means for NCES to support the collection of school-level resource data. During the technical work group meeting of education finance experts in January 1998, NCES Associate Commissioners Paul Planchon and Martin Orland noted that NCES has been exploring ways to collect school-level resource data through two principal means—a national sample survey and administrative records. Members of the technical work group urged NCES to proceed on both fronts at the same time—the sample survey because it can be accomplished more readily in the short-term, and a collection from administrative records because of its promise to yield more comprehensive data in the longer run.[13]

---

[12] State-reliable estimates are supported for the public sector only. SASS is designed to support estimates at the national and affiliation level for the private sector.

[13] A universe of administrative records is needed, for example, to compare resource allocations across different schools in a district. Such intra-district equity comparisons cannot be done through SASS or other national surveys that sample from a small number of schools in each district.

NCES might play a number of roles in supporting improved administrative records. For example, NCES could encourage standardization across states and districts in methods of collecting staffing data and in methods of allocating finance data to the school level, playing a leadership role similar to the role it has played in standardizing district-level finance records through development of the *Financial Accounting for State and Local School System.* NCES might also provide technical assistance to improve administrative records or to "harvest" the data existing in state systems.

In addition, research on school finance would be improved by the development of additional resource measures beyond those proposed to be included in the 1999–2000 SASS. For example, productivity analyses would benefit from more comprehensive measures of resources (including the adequacy of facilities and other non-staffing resources), as well as more information about teacher quality and student needs. One particularly important area for further work concerns the development of methods to identify the variation in resources used by students who, although enrolled in the same school, participate in different types of educational activities. For example, the resources used by high school students enrolled in laboratory courses or courses with small enrollments may differ substantially from the resources used by students in other types of courses.

Finally, although much can be gained by an improved understanding of the cost structure of schools, the long-run goal of researchers and policymakers is to measure educational productivity, a task that requires measures of outputs (i.e., educational outcomes) in addition to inputs or resources. Although SASS provides some limited measures of outcomes (i.e., reported graduation rates, college-going rates, absenteeism), SASS school-level data are not at this point linked to direct measures of student educational attainment.[14] Linking student outcome data linked to the proposed school-level resource and expenditure to be collected as part of SASS would provide a substantial new opportunity for the analysis of educational productivity.

[14] For discussion of a proposal to add school-level measures of attainment to the SASS, see Wu, G., Royal, M., and McLaughlin, D., *Development of a SASS 1993–94 school-level achievement subfile: Using state assessments and state NAEP. Feasibility study.* NCES Working Paper No. 97–44. Project Officer, Michael Ross. Washington, D.C.: NCES, 1997.

References

Abramson, R., Cole, C., Fondelier, S., Jackson, B., Parner, R., and Kaufman, S., *1993–94 Schools and Staffing Survey: Sample Design and Estimation.* NCES Technical/Methodology Report 96–089. Washington, D.C.: NCES, 1996.

Berne, R., Stiefel, L., and Moser, M. "The coming of age of school-level finance data." *Journal of Education Finance.* Vol. 22(3) Winter 1997: 246–254.

Berne, R. and Stiefel, L. "Student-level school resource measures." *Selected Papers in School Finance 1995.* Washington, D.C.: National Center for Education Statistics.

Busch, C., and Odden, A., "Improving educational policy and results with school-level data: A synthesis of multiple perspectives," *Journal of Education Finance.* Vol. 22(3) Winter 1997: 225–245.

Cohen, M., "Issues in school-level analysis of education expenditure data." *Journal of Education Finance.* Vol. 22(3) Winter 1997: 255–279.

Cooper, B., Sampiere, B., and Speakman, S. T. *The Finance Analysis Model, Linking Resources for Education.* Chicago, IL: Coopers & Lybrand and Washington, D.C.: the Center for Workforce Preparation, 1994.

Goertz, Margaret. "The challenge of collecting school-based data." *Journal of Education Finance.* Vol. 22(3) Winter 1997: 291–302.

Hanushek, E.A. "Assessing the Effects of School Resources on Student Performance: An Update." *Educational Evaluation and Policy Analysis.* Vol. 19(2) Summer 1997.

Hedges, L.V., Laine, R. D., and Greenwald, R. "Does money matter? A Meta-analysis of studies of the effects of differential school inputs on student outcomes." *Educational Researcher.* April 1994 Vol. 23 (3):5–14

Isaacs, J., Best, C., Garet, M., Sherman, J., and Cullen, A., *Collection of Public School Expenditure Data: Development of a Questionnaire,* NCES Working Paper No. 98–01. Project Officer, Stephen Broughman. Washington, D.C.: NCES, 1998.

Isaacs, J., Garet, M., and Sherman, J., *Collection of Private School Finance Data: Development of a Questionnaire,* NCES Working Paper No. 97–22. Project Officer, Stephen Broughman. Washington, D.C.: NCES, 1997.

Levine, R., Chambers, J., Duenas, I., and Hikido, C. *Improving the Measurement of Staffing Resources at the School Level: The Development of Recommendations for NCES for the Schools and Staffing Survey (SASS).* NCES Working Paper No. 97–42. Project Officer, Mary Rollefson. Washington, D.C.: NCES, 1998.

Monk, D. H., "Challenges surrounding the collection and use of data for the study of education finance and productivity." *Journal of Education Finance.* Vol. 22(3) Winter 1997: 303–316.

Monk, D. H., Brian, O.B., and Roelke, C.F., "Teacher resource use within New York state secondary schools." *Selected Papers in School Finance 1996.* Washington, D.C.: National Center for Education Statistics.

National Center for Education Statistics. *Fundamentals for Financial Accounting for State and Local Accounting Systems*. (Washington, D.C.: National Center for Education Statistics, 1990).

Picus, L. O. "Does Money Matter in Education?" *Selected Papers in School Finance 1995.* Washington, D.C.: National Center for Education Statistics.

ISBN 0-16-049659-4
90000
9 780160 496592

United States
Department of Education
Washington, DC 20208–5651

Official Business
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G–17

Standard Mail (B)

EA 029205

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC®

NOTICE

REPRODUCTION BASIS

- [ ] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

- [x] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").