NCES
National Center for
Education Statistics

Developments in School Finance: 2001–02

Fiscal Proceedings From the Annual State Data Conferences of July 2001 and July 2002

U.S. Department of Education
Institute of Education Sciences
NCES 2003–403

NCES
National Center for
Education Statistics

U.S. Department of Education
Institute of Education Sciences
NCES 2003–403

Developments in School Finance: 2001–02

Fiscal Proceedings From the Annual State Data Conferences of July 2001 and July 2002

June 2003

William J. Fowler, Jr.
Editor
National Center for Education Statistics

U.S. Department of Education
Rod Paige
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Val Plisko
Associate 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
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006–5651

June 2003

The NCES World Wide Web Home Page address is *http://nces.ed.gov*
The NCES World Wide Web Electronic Catalog is *http://nces.ed.gov/pubsearch*
The NCES education finance World Wide Web Home Page address is *http://nces.ed.gov/edfin*

Suggested Citation

U.S. Department of Education, National Center for Education Statistics. *Developments in School Finance: 2001–02,* NCES 2003–403, William J. Fowler, Jr., Editor. Washington, DC: 2003.

For ordering information on this report, write:

U.S. Department of Education
ED PUBS
P.O. Box 1398
Jessup, MD 20794–1398

call toll free 1–877–4ED–PUBS, or order online at *www.edpubs.org*

Content Contact
William J. Fowler, Jr.
(202) 502–7338
William.Fowler@ed.gov

Dedication

In memory of

Charles W. Foster III, 1918–2002
James E. Gibbs, 1910–2001
William P. McLure, 1910–2002

In March 2003, at the First General Session of the 28th Annual Conference of the American Education Finance Association, Eugene P. McLoone, Past President of the AEFA, requested a moment of silence in memory of Charles W. Foster III, James E. Gibbs, and William P. McLure. This year's *Developments in School Finance* is dedicated to these individuals in recognition and appreciation of their contributions to the field of education finance.

Charles W. Foster III was the second full-time employee of the Association of School Business Officials (ASBO) International and its first executive secretary, a position he held from 1955 until his retirement in 1978. During his tenure as executive secretary, ASBO International flourished, growing in membership from under 2,000 to more than 5,000. He was instrumental in bringing research to education business practices as well as improving professional education for school business officials. He received his doctorate of education in business management from Northwestern University in 1954.

James E. Gibbs was the first chief of the State Branch, Elementary and Secondary Division, U.S. Office of Education, after it was entrusted with improving data collection in state departments of education under Title X of the National Defense Education Act (NDEA). He continued in that position, with additional responsibilities under the Elementary and Secondary Education Act of 1965 (ESEA), until his retirement. He was the federal official most responsible for advancing state departments of education into the electronic data processing era. He also served as the fourth president of the American Education Finance Association (AEFA). He received his doctorate in 1954 from Peabody College of Vanderbilt University.

William P. McLure spent his life as an educator, analyst, and specialist in education finance and administration. He was AEFA's second president, in 1977. For many years he was professor of Educational Administration at the University of Illinois, Champaign-Urbana, and director of the Bureau of Education Research and Service. Studies he conducted improved understanding of the relationships among administrative, cost, and school performance factors. He received his doctorate from Columbia University in 1948.

Foreword

Jeffrey A. Owings

Associate Commissioner
NCES Elementary/Secondary and Libraries Studies Division

At the 2001 National Center for Education Statistics (NCES) Summer Data Conference, scholars in the field of education finance addressed the theme "Making Data Work." Discussions and presentations dealt with topics such as the effective display of finance data, assessing the financial condition of school districts, and the economic efficiency and funding adequacy of school districts. The theme for the 2002 Summer Data Conference was "Common Data, Common Goals" and the topics of education finance addressed included teacher pay, vouchers, measuring the cost of education, and the school district bond rating process.

Developments in School Finance: 2001–02 contains papers presented at the 2001 and 2002 annual NCES Summer Data Conferences. These Conferences attracted several state department of education policymakers, fiscal analysts, and fiscal data providers from each state, who were offered fiscal training sessions and updates on developments in the field of education finance. The presenters are experts in their respective fields, each of whom has a unique perspective or interesting quantitative or qualitative research regarding emerging issues in education finance. It is my understanding that the reaction of those who attended the Conferences was overwhelmingly positive. We hope that will be your reaction as well.

This volume is the seventh education finance publication produced from papers presented at the NCES Summer Data Conferences. The papers included 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 Data Conferences.

Acknowledgments

The editor gratefully acknowledges the suggestions and comments of the reviewers at the National Center for Education Statistics (NCES): Jeffrey Owings, Associate Commissioner for the Elementary/Secondary and Libraries Studies Division, who provided overall direction, and Karen O'Conor, who provided early guidance on table format. At the Education Statistics Services Institute (ESSI), Leslie Scott provided technical review of the entire publication. Also at ESSI, Tom Nachazel proofread and coordinated production of the publication, with assistance from other members of the ESSI editorial team; Heather Block and Jennie Jbara performed the desktop publishing; and Elina Hartwell designed the cover.

Contents

Introduction

William J. Fowler, Jr.
National Center for Education Statistics

This introduction is divided into two sections. The first section provides a brief overview of each of the 10 papers included in this volume. The second section describes two new features of the search function of the National Center for Education Statistics (NCES) education finance (edfin) web site. These two sections are followed by the complete text of the 10 papers.

Section I—Overview of Papers

For the benefit of the reader, this section provides brief overviews of the 10 papers selected for this volume of fiscal proceedings. These papers were presented by education finance experts at the July 2001 and July 2002 NCES Summer Data Conferences. The presenters were invited to address "cutting edge" research in public school education finance. The following paragraphs present an overview of the papers in this volume, in the order in which they appear. For each paper, the title (in bold) and list of authors and their affiliations introduce the paper summary.

What We Know and What We Need to Know About Vouchers and Charter Schools. In this paper, Brian P. Gill, P. Michael Timpane, Karen E. Ross, and Dominic J. Brewer of RAND focus on competition in education, by such means as vouchers and charter schools, and explore the scant empirical evidence about these forms of competition. They raise empirical questions that they believe have not been addressed, examine whether the questions can be addressed in future work, and provide recommendations for policymakers planning to enact programs to promote competition.

The authors begin by examining empirical studies of the effects of competition on academic achievement, choice, access, integration, and civic socialization. Their empirical investigation demonstrates that many of the questions they believe should be addressed remain unanswered. For example, there appears to be a modest achievement benefit for African American children after 1 to 3 years in voucher schools, compared with continuation in local public schools; but the reason for the benefit and how long those effects last are unknown. Also, for other racial/ethnic groups, there is no consistent evidence. The little evidence that is available regarding the effects of a voucher program on the students who remain in public schools suggests that competition from vouchers may improve academic performance in public schools.

Gill et al. find that what we don't know about competition vastly exceeds what is known. There is little information comparing the achievement efficacy of different reforms, such as class size reduction, teacher professional development, high-stakes accountability, and the effects of programs to promote competition on those students who remain in public schools.

One aspect of competition that is likely to have a large impact on empirical outcomes is the program scale of such efforts. The authors remind us that nearly all of the existing evidence comes from tiny "escape valve" interventions in which competition evolves from efforts to assist high-risk children. It may be that only in comparison with underperforming public schools are vouchers and charter schools effective. The authors note that the economic costs of larger scale competition programs are unknown and warn that "scaling-up" often causes unexpected difficulties.

Gill et al. conclude by providing guidance to policymakers about the intelligent design of programs that promote competition, to maximize program benefits and mitigate harm.

Getting the Biggest Bang for the Educational Buck: An Empirical Analysis of Public School Corporations as Budget-Maximizing Bureaus. Anthony Rolle of Vanderbilt University presents, in this paper, an empirical analysis of whether public schools are budget-maximizing institutions. He undertook this examination because recent trends in education seem to be characterized by continued increases in organizational size and fiscal resources and by decreases in educational outcomes. The goal of the paper is to create a common understanding about the efficient uses of public education dollars.

In examining the literature, Rolle finds there is ample evidence that bureaucrats systematically request larger budgets regardless of the level of organizational output. While most of traditional economics deals with the behavior of profit-seeking firms, Rolle turns to a 1971 theory of supply for public bureaus developed by Niskanen, who concludes that bureaucrats attempt to maximize their agency's total budget during their tenure. Rolle selects Indiana, from 1981 through 1997, as an ideal setting to examine whether Niskanen's theory applies to public schools.

Efficiency in public schools, Rolle asserts, is concerned with how much education or knowledge is delivered to, and acquired by, students and at what cost. He uses a quadriform, or four quadrants, to assess school district efficiency outcomes. Inefficient school districts generate lower than expected outcomes with higher than expected expenditures. Effective school districts have higher than expected outcomes but also higher than expected expenditures. Efficient school districts have high outcomes and lower than expected expenditures. Rolle hypothesizes that if Niskanen is correct, variables measuring educational outcomes should not be statistically significant predictors of total expenditures per pupil.

Rolle concludes that Indiana public school districts cannot be designated as budget-maximizing agencies as defined by Niskanen's theory. However, they produce educational outcomes in a manner that is economically inefficient.

Occupational Choices and the Academic Proficiency of the Teacher Workforce. Authors Dan D. Goldhaber of the University of Washington and the Urban Institute and Albert Yung-Hsu Liu of the Urban Institute seek to identify the characteristics of teacher candidates in the "teacher pipeline" and examine the effects of compensation in different occupations on a teacher candidate's progress. They do this by examining whether respondents to the NCES Baccalaureate and Beyond Longitudinal Study (B&B) have taught, have trained as a teacher, are currently considering entering teaching, and have applied for at least one teaching job.

The rationale for this analysis is the authors' assertion that among schooling characteristics, teacher effectiveness has been shown to explain the largest share of the variation in student achievement. Although researchers have been unable to reach a consensus on which teacher characteristics correlate with student achievement, it is apparent that the teacher workforce tends to consist of college graduates with weaker academic skills. In addition, teachers with strong academic and specialized skills tend to migrate to schools with high socioeconomic and high-achieving students. Thus, it is important to identify other characteristics correlated with interest in teaching.

Exploring the B&B data, Goldhaber and Liu find that students who have considered teaching are more likely to have had a mother employed as a teacher and to come from a low-income family. As might be expected, males are far less likely than females to consider teaching. Also, students who attend more selective colleges are less likely to actually apply for a teaching job.

The researchers find that the teacher labor market primarily rewards experience and advanced degrees. In contrast, the non-teacher labor market, while rewarding experience and advanced degrees, also rewards college selectivity and technical major. In the non-teacher labor market, wages are predicted to be higher for students with higher SAT scores; however, there is no similar premium for SAT scores (nor for college selectivity) in the teacher labor market. For most students, Goldhaber and Liu find, earnings are predicted to be higher outside of teaching than in the teaching profession and can be as much as $10,000 more for males with technical majors who graduate from more selective colleges. The authors conclude by cautioning that teachers may be particularly sensitive to non-pecuniary job characteristics when deciding to become teachers and choosing schools in which to work.

Variation in the Rewards for a Teacher's Performance: An Application of Quantile Regressions. In this paper, Sherrilyn M. Billger of Illinois State University explores the pay rewards for a private secondary school teacher's performance, using data from the NCES Schools and Staffing Survey (SASS) for 1990–91. She explores private school salaries because they exhibit substantial variation and a greater use of incentives than public school salaries.

Incentive pay, Billger asserts, is regaining popularity, even among teachers, and some 10 percent of public and private schools have such incentives. Using quantile regressions, Billger provides a fuller understanding of the relationship between salary and experience, suggesting that the returns to experience are greatest for the highest performing teachers.

The subject taught affects private secondary school teacher compensation, but compensation is not related to teaching the same subject as the teacher's college major. Incentive programs also affect salary; private secondary school teachers at schools with a merit pay program earn 6 percent higher salaries than do private secondary school teachers at schools without such a program.

National Evidence on Racial Disparities in School Finance Adequacy. Ross Rubenstein of Syracuse University explores, in this paper, the NCES Common Core of Data (CCD) to examine racial disparities in the adequacy of school financing across the United States. Rubenstein quantifies differences in adequacy across states, and across racial groups within states, and estimates the cost to bring all students' schools to selected levels of adequacy. A great deal of research has explored school finance equity within states, but much less research has examined adequacy. While equity analyses compare school districts to each other, adequacy measures education funding relative to an absolute standard.

Rubenstein asserts that an adequate funding level is one that provides all students the opportunity to achieve specified benchmarks and goals. Three methods of measuring adequacy have typically been used to determine adequate funding levels for different types of students. One method is a "professional expert" approach, in which experienced educators identify preferred instruction and estimate the price of the necessary components. A second method is the "exemplary district" approach, in which school districts with higher performance and lower spending are identified and set as the standard for each type of district. A third method uses an econometric approach, in which expenditures are related to various measures of student performance, and needs are used to construct a "cost index" that measures differences across districts in the resource levels required to produce a given level of student performance. Rubenstein uses the Odden-Picus Adequacy Index (OPAI) to quantify how far a given finance system is from achieving adequacy. Generally, achieving adequacy involves raising spending in all districts to the national median to provide adequate funding.

Rubenstein finds that most states have a higher proportion of students in schools below the adequacy benchmark than of districts below the benchmark. Most states with higher proportions of African American students in districts below the national benchmark also have lower proportions of minority students in school districts that have lower spending. In other words, minority children within states do not appear to be concentrated in lower spending districts.

He concludes that additional spending of $14–$16 billion would be needed to raise all districts in the country to the national median. The most consistent disparities across states are regional, with northeastern states generally having high levels of adequacy and southeastern states having low levels of adequacy. Interstate racial disparities in adequacy are generally greater than intrastate disparities. Urban and urban fringe districts are more likely to be below the median.

Competing Perspectives on the Cost of Education. Several presenters' papers addressed geographic cost variations and how differing approaches yielded considerably different estimates of the costs of education for school districts in Texas and New York. This paper, presented by Lori L. Taylor and Harrison Keller of the University of Texas at Austin, offers a brief discussion of current theory and practice regarding geographic cost adjustments, followed by discussion on the costs of public education in Texas, using different indexing strategies.

Taylor and Keller assert that there are two basic strategies for reflecting differences in school districts' geographic costs: cost-of-living (COL) and cost-of-education (COE). The basic premise of COL is that areas with relatively higher costs of living have to pay higher salaries to attract and retain school employees, which increases the cost of operating schools and school districts. COL estimates use either a "market basket" of goods and services (much like the Consumer Price Index) or a "comparable wage" strategy. The latter approach involves comparing the salaries of educators and non-educators. An advantage of both approaches is that these costs are beyond the control of school administrators. However, there are at least two limitations to these approaches, aside from the expense of data collection. Different communities may select different "market baskets," which would have different costs. For example, some school districts might select only teachers with advanced degrees and previous teaching experience. Another disadvantage is that high-cost communities may have amenities that make them desirable places to work. Finally, estimates typically are available only for large metropolitan areas, and many school districts with different costs may have the same estimate from a COL strategy.

COE estimates, Taylor and Keller maintain, use data on district expenditures to estimate either the costs of providing comparable levels of educational services (by estimating the cost to hire a typical teacher) or the costs of producing comparable educational outcomes. COE estimates can be applied to specific school districts, rather than a large metropolitan area, taking into account the cost variations within labor markets. COE estimates might also be obtained at a lower cost than COL estimates. COE estimates based on the cost of achieving educational outcomes can estimate both for variations in the prices paid for school employees and for deploying those employees to attain better student outcomes. Disadvantages of the COE approach include the possibility of missing a variable in the equation that increases costs for a school district. Perhaps most troubling, there is some evidence that certain school districts do not exhibit cost-minimizing behavior. If school officials can manipulate expenditures, or if the COE reflects inefficient school district operation, those districts that appear to have higher costs may simply reflect these local actions.

Taylor and Keller's examination of seven Texas geographic cost indexes finds little agreement across indexes regarding characteristics of high- and low-cost districts; it attributes these differences across indexes to differences in methodology. Within-market variations in labor markets are relatively small compared to between-market variations, favoring the COL approach. The authors conclude that the cost of educational inputs is a poor proxy for the cost of educational outcomes. It is precisely this last finding of Taylor and Keller that interests the authors of the following paper.

Financing an Adequate Education: A Case Study of New York. In this paper, William Duncombe and John Yinger of Syracuse University and Anna Lukemeyer of the University of Nevada, Las Vegas, seek to develop a school finance system that supports students and school districts trying to reach higher performance standards in New York State. They focus on the problem that schools with disadvantaged students must spend more than other schools to meet any given standard. The authors develop estimates of a district's cost for achieving an adequacy standard, and they propose funding such costs through a "foundation aid" formula.

As a standard of performance, Duncombe, Lukemeyer, and Yinger select a measure of performance set by the New York State Education Department. They find wide disparities in student achievement across districts in New York State, tied closely to school district size and urbanization. The five large city school districts have performance levels well below the current state average. Although only 5 percent of school districts do not reach a modest standard, they serve close to half the students in the state.

Using a COE index, the authors find that teacher costs differ between upstate and downstate districts, with downstate districts having above-average costs and upstate school districts having below-average costs. They suggest that wealthier school districts may be less efficient than poor school districts. The cost function they include uses the share of district enrollment of limited English proficient (LEP) students and the percentage of district children living below the poverty line. They calculate that each student in poverty requires a district to spend between $7,000 and $9,000 in additional resources to maintain the average performance level in New York.

Having determined the cost of adequacy, Duncombe, Lukemeyer, and Yinger devise a "cost-adjusted foundation aid" funding system with a minimum local tax rate requirement to achieve student outcome adequacy. The spending levels in the high-need New York urban school districts, the authors find, would have to rise to levels seldom achieved in large cities anywhere in the nation to bring students up to any reasonable standard. Such a substantial increase in state aid to high-need districts might increase inefficiency, they warn. However, they conclude that it is time to implement state aid systems that explicitly recognize that some districts must spend more than others to achieve any performance standard.

Bond Ratings and Bond Insurance: Market and Empirical Analysis for School Districts. Few studies have explored the decisionmaking process for school district officials when they are faced with the prospect of issuing bonds to fund extensive capital expenditures. In this paper, Mary H. Harris of Cabrini College empirically explores the difficult decision of school officials to have a bond issue rated by an independent rating agency. Few readers may be aware that the district must pay the independent agency a fee to cover the cost of conducting the credit rating. Once the bond is rated, the school district officials must then decide whether to purchase insurance as a credit enhancement to improve the rating. Harris examines 148 bond issues in 10 states from July 1993 through June 1994, where the proceeds were used for capital expenditures.

The rating agencies were originally developed to assist investors in comparing different bond issues with standard letter ratings. Moody's Investors Service (Moody's) focuses more on debt, and Standard & Poor's (S&P) focuses more on the economic base of the issuing entity. The fee these agencies charge is usually based on time and effort and averaged $7,000 per rating in 1993–94. The majority (58 percent) of rated bonds are rated by both Moody's and S&P. Larger bond issues typically receive ratings from two or three rating companies. Often, school districts choose to stay with a particular rating agency, as the cost of updating their information is lower than the cost of switching agencies.

If a school district receives a high bond rating on its issue, the results will be a lower bond financing cost and the ability to market the issue to a larger pool of investors. A high bond rating also reduces the price of the bond, thereby reducing the total financing cost. However, a school district may choose not to obtain a rating. For example, if the school district official anticipates that the issue will receive a poor rating and decides that not having any rating at all is just as attractive, the official may choose not to obtain a rating. If the issue is to be marketed locally, there is also little need for an agency rating. Another reason not to obtain an agency rating is if the amount of debt is small enough that the interest savings from the good rating are not large enough to offset the cost of obtaining a rating. Harris finds interesting regional differences and calls for more research in an often-neglected area.

GASB Update. The Governmental Accounting Standards Board (GASB) has adopted many accounting changes that will affect school districts, and Randal Finden of GASB addresses several in this brief article. He begins with a mention of the new financial reporting model for school districts and other public governmental entities (Statement No. 34), which includes a required "management's discussion and analysis" (MD&A) that describes a school district's financial events in layman's language. GASB Statement No. 39 requires the inclusion in a school district's financial statements of school district "affiliated organizations," such as parent-teacher organizations (PTOs), parent-teacher associations (PTAs), and foundations, provided their resources are "significant" to the school district.

Finden reports that GASB has also recently issued an exposure draft of a proposed Statement, to be effective after June 15, 2004, regarding the review of existing deposit and investment disclosure requirements. Investments must be reported at fair value, and as such things as interest rates change, investment values vary. The proposed Statement suggests methods a small government may use to reveal such investment risk, the simplest method being to list the investment, its maturity, and any call options. Credit risk outlines the debt obligations of a local government. Custodial credit risk involves deposits in financial institutions that might fail. GASB is seeking comment on this proposed Statement.

Finally, Finden discusses a current project of GASB, Other Post Employment Benefits (OPEB). OPEB refers to postemployment benefits other than retirement benefits, such as medical, dental, vision, and hearing benefits, and, when they are provided separately from a pension plan, life insurance and long-term care. GASB has begun to consider OPEB a part of compensation, deferred until after employment, and has tentatively decided to require recognition of OPEB costs over an employee's years of service, similar to current pension reporting requirements. GASB is still working on a method for small employers to calculate OPEB liability and expense without the use of an actuary. These and other changes in financial reporting will require our attention for years to come.

High Performance of Minority Students in DoDEA Schools: Lessons for America's Public Schools. The average academic achievement of all students and of African American and Hispanic students in Department of Defense (DoD) schools is among the highest in the nation according to the National Assessment of Educational Progress (NAEP). The Department of Defense enrolls approximately 112,000 students in schools in the United States and overseas, about the same number as the enrollment in the Charlotte-Mecklenburg, North Carolina, school district, with roughly 40 percent minority students. Claire E. Smrekar and Debra E. Owens of Vanderbilt University explore in their paper why DoD students outperform their peers, using a sample of 15 middle schools located in 10 school districts across the United States, Germany, and Japan. The paper is based on approximately 130 interviews conducted over a 4-month data collection period, focusing on financial support, resource allocation, personnel recruitment and selection, teacher quality, accountability, leadership styles, program diversity, and academic policy priorities.

Smrekar and Owens find that the DoD assesses the achievement of DoD students every year through standardized testing. Every school and district is provided with detailed assessment results, including performance by grade level, gender, and race. This intensive testing assists in school improvement. Schools that do not meet the DoD standard are targeted for intervention and enhanced resources. The cost per pupil in DoD schools is higher than the national average cost per pupil for U.S. public schools. Teacher salaries are competitive and schools are well staffed. Instruction, Smrekar and Owens find, is enhanced by state-of-the-art equipment and well-maintained facilities. The salary schedules in school districts of comparable size and demographics are reviewed regularly by the DoD to maintain a competitive salary schedule. All districts in the study reported extensive staff training linked to school goals.

Smrekar and Owens conclude with a variety of policy recommendations that flow from their findings. For example, they find that a larger proportion of middle and high schools in the DoD system have small enrollments compared to comparable public schools. They suggest that creating smaller "learning communities" may facilitate educational benefits for minority students in civilian schools.

Section II—New Features of the NCES Education Finance Web Site

Many readers are aware of the NCES education finance web site (*http://nces.ed.gov/edfin*) as a source of information, publications, and data on elementary and secondary education school finance and as a means of comparing school districts. In addition to the Standard search feature previously available for comparing school districts, two new search features have been added: 1) Geographic and 2) Create Your Own Group. Each of the three search features is described below. The descriptions are followed by examples.

- *Standard* peer search uses the characteristics of a named school district, such as total students, student/teacher ratio, percentage of children in poverty, district type, and locale code, to select similar districts and compare their spending.
- *Geographic* peer search permits users to select school districts a specified distance from a particular zip code, listing all the school districts within, say, 10 miles of the zip code.
- *Create Your Own Group* permits you to choose only those school districts you wish to compare.

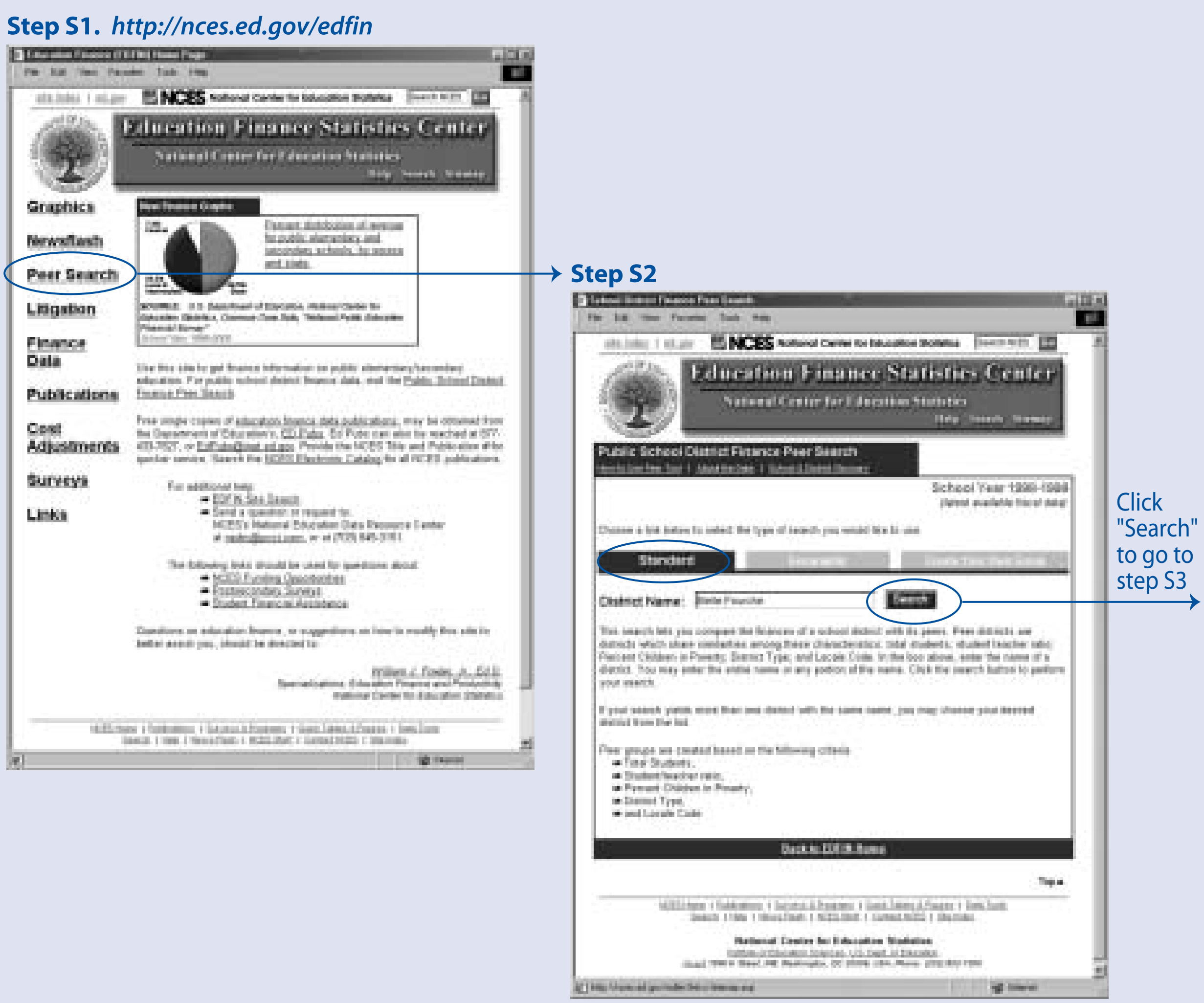
Standard peer search

For this example, assume that a school district official wants to compare the Belle Fourche school district to similar districts throughout the country.

Step S1 (figure 1)—
Go to the education finance web site. Click on "Peer Search."

Step S2 (figure 1)—
With "Standard" selected, specify school district (Belle Fourche). Click on "Search" to get a list of peer districts.

Figure 1. Starting a Standard peer search



Step S3 (figure 2)—
With peer districts listed, select one of the tabs at the top of the page (Revenues, Expenditures, Other Expenditures, Characteristics, Other Characteristics) for the information you want.

The step S3 graphic shows 3 of the 146 peer districts that were found, with "Revenues" selected. We see that the South Dakota school districts Belle Fourche, Hot Springs, and Vermillion have revenue per student of $6,741, $5,670, and $6,080, respectively.

Step S4 (figure 2)—
To get a web page of information about a school district, as shown in the step S4 graphic, click on the name of one of the districts listed in step S3.

Geographic peer search

For this example, assume a school district official wants to compare the Belle Fourche school district to districts that are located within 15 miles of Belle Fourche.

Step G1 (figure 2)—
To conduct a geographic peer search, return to the peer search home page. (One way to do this is to click on "New Search" at the bottom of a peer search web page, as shown in figure 2 below the step S3 graphic.)

Step G2 (figure 3)—
Select the "Geographic" tab. Enter the zip code of the district you want to compare (e.g., 57717 for Belle Fourche). Select the distance you want to search—15

Figure 2. Getting results of a Standard peer search and starting a new search

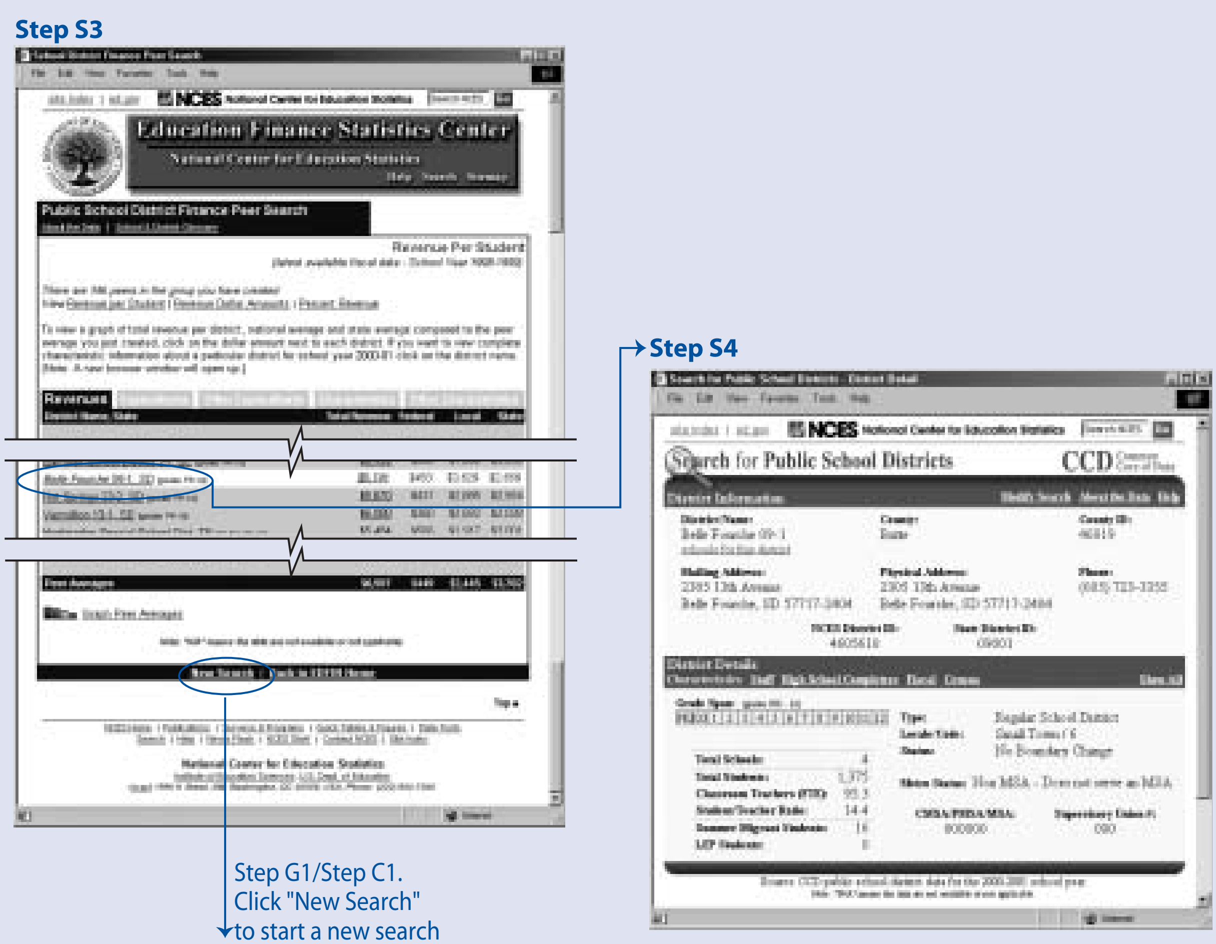
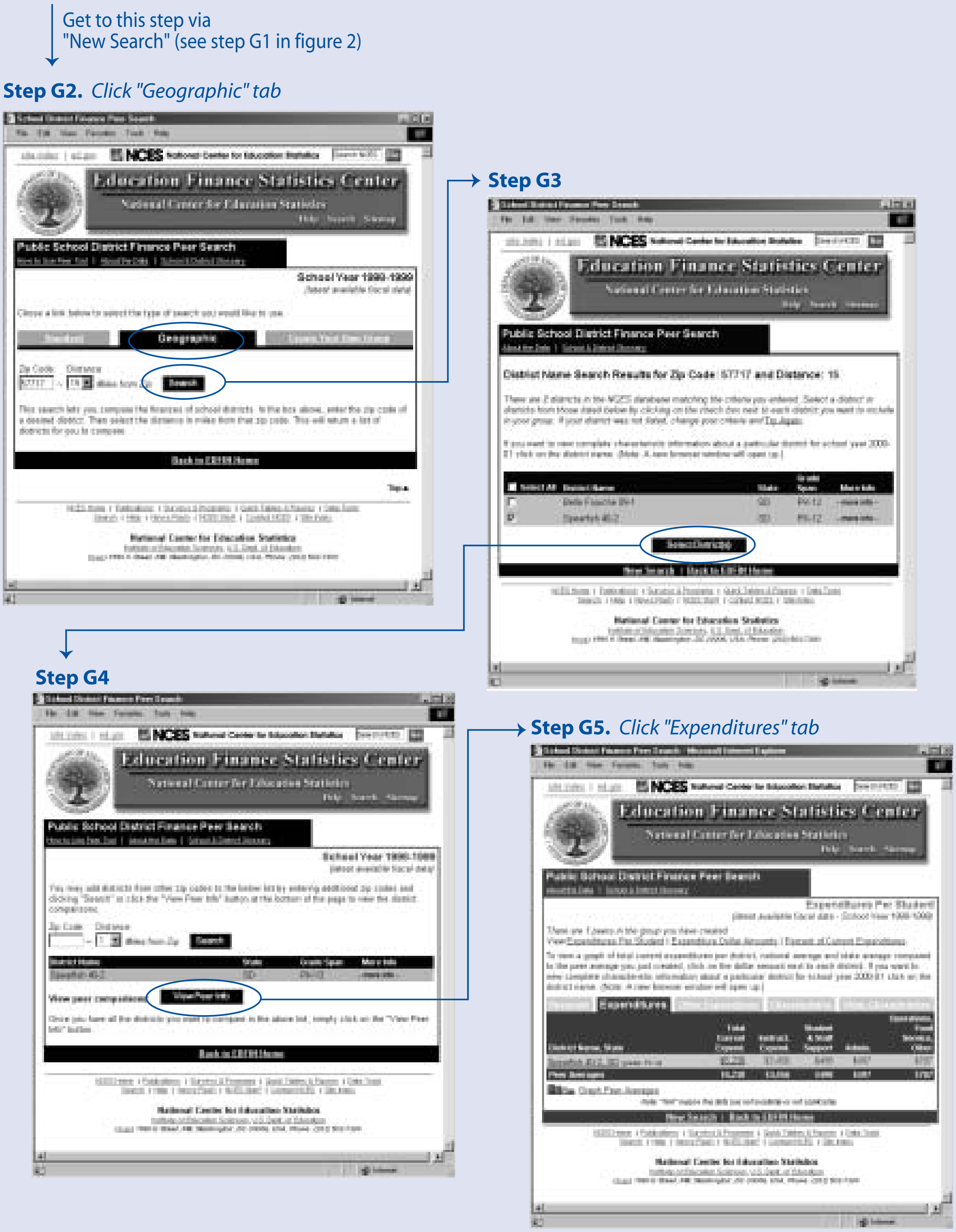


Figure 3. Conducting a Geographic peer search



miles—from the drop-down mileage selector. Click on "Search."

Step G3 (figure 3)—
The geographic search found Spearfish, South Dakota. Check Spearfish and click on "Select District" to get to step G4.

Step G4 (figure 3)—
Click on "View Peer Info" to get to the step G5 page, where you can select one of the following tabs: Revenues, Expenditures, Other Expenditures, Characteristics, or Other Characteristics.

Step G5 (figure 3)—
Click on "Expenditures" to get the information about Spearfish shown in the step G5 graphic.

Create Your Own Group peer search

For this step, assume a Belle Fourche school district official has identified two other districts with which to compare Belle Fourche.

Step C1 (figure 2)—
To compare Belle Fourche with other districts of your choice, return to the peer search home page. (One way to do this is to click on "New Search" at the bottom of a peer search web page, as shown in figure 2 below the step S3 graphic.)

Step C2 (figure 4)—
Click on "Create Your Own Group." Enter the name of each school district you want to compare, along with the state name from the drop-down list, clicking on "Search" after each selection. In this example, Belle Fourche, Hot Springs, and Vermillion (all in South Dakota) are selected.

Step C3 (figure 4)—
With three districts selected as a result of step C2, click on "View Peer Info."

Step C4 (figure 4)—
As a result of step C3, five tabs appear above the list of districts: Revenues, Expenditures, Other Expenditures, Characteristics, and Other Characteristics. Click on "Characteristics" to access the information shown in the step C4 graphic.

Conclusion

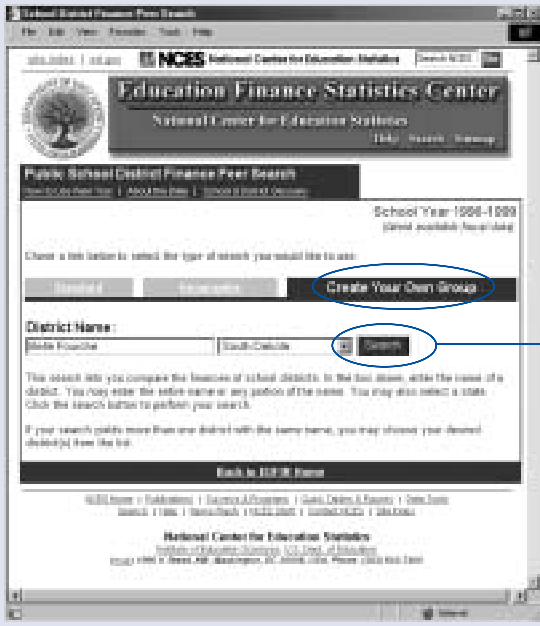
An "advanced" function is still under construction. It is anticipated that this upgraded function will provide a greater choice of school district criteria, such as confining searches to a single state, or to schools with a certain percentage of students in poverty.

NCES hopes you will try out the new "peer search" features and welcomes comments and suggestions to enhance this function, and that of the edfin web site. Comments may be e-mailed to *William.Fowler@ed.gov*.

Figure 4. Conducting a Create Your Own Group peer search

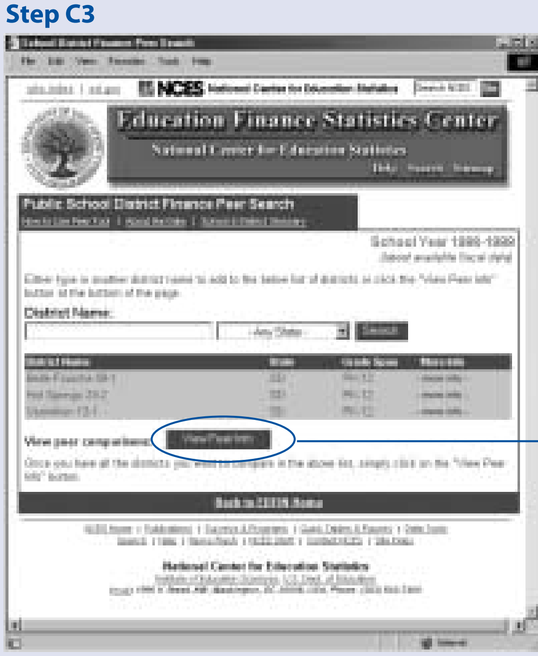
Get to this step via
"New Search" (see step C1 in figure 2)

Step C2. *Click "Create Your Own Group" tab*

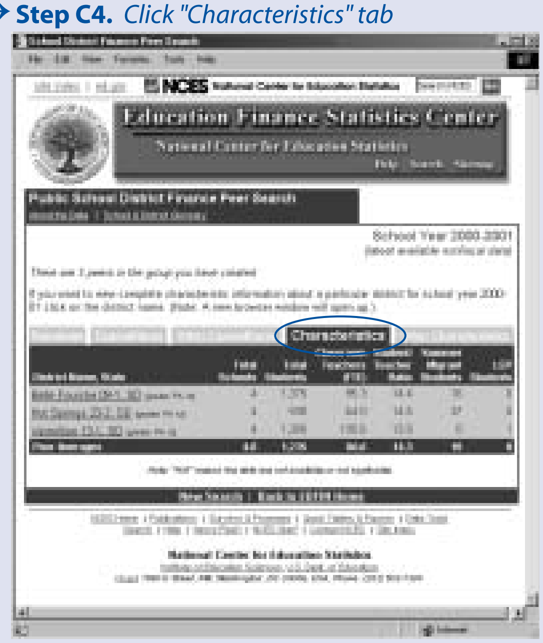


Enter each district name and click "Search" until you have created your desired group

Step C3



Step C4. *Click "Characteristics" tab*



What We Know and What We Need to Know About Vouchers and Charter Schools

Brian P. Gill
P. Michael Timpane
Karen E. Ross
Dominic J. Brewer
RAND

About the Authors

Brian P. Gill (Ph.D., Jurisprudence and Social Policy, University of California, Berkeley; J.D., Boalt Hall School of Law, University of California, Berkeley) is a social scientist at RAND who studies educational choice, the structures of school governance, and accountability to both parents and government. He is lead author of *Rhetoric Versus Reality: What We Know and What We Need to Know About Vouchers and Charter Schools* (2001). Dr. Gill is co-principal investigator in an ongoing RAND evaluation of Edison Schools, the nation's largest for-profit operator of public schools, and he is a participant in RAND's state-commissioned evaluation of charter schools in California. He is also a member of the Gates National Commission on School Choice at the Brookings Institution.

P. Michael Timpane (M.A., History, Catholic University; Master's in Public Administration, Harvard University) is co-chair of the Aspen Institute Program on Education in a Changing Society and a Senior Advisor for Education Policy at RAND. His assignments span the range of education policy, from pre-K to postgraduate studies, and emphasize the relationships among education and other realms of social and economic policy. As Vice President and Senior Scholar at the Carnegie Foundation for the Advancement of Teaching prior to joining RAND, Mike was involved in developing all aspects of the Foundation's program and in his own research assessing the progress and problems of contemporary national education reform. He is Professor of Education and a former president of Teachers College, Columbia University, and has served as Dean of Teachers College and Deputy Director and Director of the federal government's National Institute of Education.

Karen E. Ross (currently pursuing a Ph.D. in Sociology and Public Policy, University of Michigan; M.A., Sociology, Johns Hopkins University) was a co-author of *Rhetoric Versus Reality* and a participant in RAND's evaluation of Edison Schools. Her dissertation examines the effect of charter schools and educational choice on racial integration in the state of Michigan, and she has recently been commissioned to contribute her expertise on the topic to the Gates National Commission on School Choice at the Brookings Institution.

Dominic J. Brewer (Ph.D., Economics, Cornell University) is Senior Economist and RAND Education Unit Director, and was a co-author of *Rhetoric Versus Reality*. Dr. Brewer has led numerous projects for

RAND for a wide variety of clients, including federal government agencies and foundations dealing with the effectiveness of educational programs and policies. He has published more than 40 articles and books, including many specifically directed toward a policy or lay audience (for example, a recent article on class size reduction appeared in *Scientific American*). As the Director of the Education Unit, he is responsible for the management of an $8 million research program in education policy, with over 40 staff members.

The papers in this publication were requested by the National Center for Education Statistics, U.S. Department of Education. They are intended to promote the exchange of ideas among researchers and policymakers. The views are those of the authors, and no official endorsement or support by the U.S. Department of Education is intended or should be inferred. This publication is in the public domain. Authorization to reproduce it in whole or in part is granted. While permission to reprint this publication is not necessary, please credit the National Center for Education Statistics and the corresponding authors.

What We Know and What We Need to Know About Vouchers and Charter Schools

Brian P. Gill
P. Michael Timpane
Karen E. Ross
Dominic J. Brewer
RAND

Introduction

In today's context of widespread dissatisfaction with our nation's public education system, a variety of reforms have been proposed to improve educational outcomes. One of the most controversial proposals is to provide parents with a financial grant, or "voucher," for use at any public or private school. Proponents argue that children using the voucher would be able to attend more effective and efficient schools; that the diversity of choices available would promote parental liberty and, if properly designed, benefit poor and minority youth; and that the competitive threat to public schools would induce them to improve. Thus, all would benefit from the use of vouchers. In what has become a fiercely contentious and highly political debate, opponents claim that vouchers would destroy public schools, exacerbate inequities in student outcomes, increase school segregation, breach the constitutional wall between church and state, and undermine the fabric of democracy by promoting narrow, particularistic forms of schooling.

Another proposal for education reform, less controversial among policymakers and the public, is to establish "charter" schools that are funded by public money and approved by a public agency but are schools of choice that operate outside the traditional system of public-school governance. A charter school is a self-governing institution, operating under a quasi-contract, or "charter," that has been issued by an agency of government such as a school district or a state education authority. A few voices have been raised in opposition to charter schools, expressing concerns about the possibility that they could lead to stratification in student placement and balkanization in curriculum. For the most part, however, charter schools have achieved considerable popularity across the political spectrum, with policy arguments centering on the terms and conditions of public oversight, such as collective bargaining provisions, applicability of assessment and accountability programs, and admissions policies. Charter school advocates argue that these schools will serve as laboratories for pedagogical innovation, provide havens for children who have been poorly served by traditional public schools, promote parental involvement and satisfaction, improve academic achievement, and save public education.

Conceptually and structurally, vouchers and charter schools challenge the "common school" model that has been the basis for the American system of public education for most of the nation's history. Opponents fear that privatizing the governance and operation of schools will undermine their public purposes; supporters believe that autonomously operated voucher and charter schools can serve the public purposes of the educational system even though they are not owned and operated by government. ("Voucher schools" are schools that students with vouchers choose to attend.) Policymakers need empirical information on the effects of vouchers and charter schools to assess their merits and resolve this dispute.

This essay summarizes findings of a recent RAND book (Gill, Timpane, Ross, and Brewer 2001) that examines the empirical evidence on vouchers and charter schools. The aims of the book and this essay are fourfold. First, we identify and articulate the range of empirical questions that ought to be answered to fully assess the wisdom of policies that promote vouchers or charter schools, establishing a theoretical framework that accounts for the multiple purposes of public education. Second, we examine the existing empirical evidence on these questions, providing a broad assessment of what is currently known about the effects of vouchers and charter schools, in terms of academic achievement and otherwise. Third, we discuss the important empirical questions that are as yet unresolved and consider the prospects for answering them in the future. Fourth, we explore the details of the design of voucher and charter policies, concluding with recommendations for policymakers who are considering their enactment.

This essay summarizes findings of a recent RAND book that examines the empirical evidence on vouchers and charter schools.

The empirical evidence discussed in the book is derived from an exhaustive review of the existing literature on vouchers and charter schools, from studies of other forms of school choice in the United States and abroad, and from comparative studies of public and private schools.

Defining the Relevant Policy Questions

We seek to define the full range of policy questions about the empirical effects of school choice. An assessment of the wisdom of a voucher or charter law requires a full understanding of the varied goals that a system of schooling should promote. We divide the major policy questions into five broad categories, constructed to reflect the goals that are explicit or implicit in the arguments of both the supporters and opponents of educational choice, and more generally in the philosophical positions of those who have supported a public role in education over the last two centuries:

1. *Academic achievement*. Will vouchers/charter schools promote the academic skills, knowledge, and attainment of their students? How will they affect the achievement of those who remain in assigned public schools, as well as those who move to voucher/charter schools?
2. *Choice*. What is the parental demand for vouchers and charter schools? Will vouchers/charter schools induce a supply response that makes a variety of desirable school options available? What do voucher/charter parents think of their children's schools?
3. *Access*. Will voucher/charter programs be available to those who presently lack such options, notably low-income (frequently minority) residents of the inner city? Will they provide any options for students with special needs?
4. *Integration*. Will vouchers/charter schools increase or reduce the integration of students across and within schools and communities, by race/ethnicity and socioeconomic status?
5. *Civic socialization*. Will vouchers/charter schools contribute to the socialization of responsible, tolerant, democratically active citizens, or will they promote intolerance and balkanization?

What We Know From the Existing Empirical Evidence

Our evaluation of the existing evidence indicates that many of the important empirical questions about vouchers and charter schools have not yet been answered. Indeed, it would be fair to say that none of the important empirical questions have been answered definitively. Even the strongest evidence is based on programs that have been operating for only a short period of time with a small number of participants; serious questions about generalizability remain. Nevertheless, the evidence is converging in some areas, as outlined below:

Academic Achievement

- Small-scale, experimental, privately funded programs that are targeted to low-income students suggest a possible (but as yet uncertain) modest achievement benefit (on the order of 0.1 to 0.3 standard deviations after 1 to 3 years) for African American children after 1 to 3 years in voucher schools, as compared with local public schools (Greene 2000; Howell and Peterson 2002; Howell et al. 2002; Mayer et al. 2002).
- For children of other racial/ethnic groups, attendance at voucher schools has not provided consistent evidence of either benefit or harm in terms of academic achievement (Howell and Peterson 2002; Howell et al. 2002; Mayer et al. 2002).
- Achievement results in charter schools are mixed, but they suggest that charter school performance improves after the first year of operation. None of the studies suggest that charter school achievement outcomes are dramatically better or worse than those of conventional public schools, on average (Bettinger 1999; Gronberg and Jansen 2001; Solmon, Paark, and Garcia 2001).
- Minimal evidence is available to assess the effect of vouchers and charter schools on the achievement of students who remain in conventional public schools. One of the few studies assessing the issue systematically (examining vouchers in Milwaukee and charter schools in Arizona and Michigan) suggests the possibility that competition from vouchers or charter schools may improve academic performance in conventional public schools (Hoxby 2002).

Even the strongest evidence is based on programs that have been operating for only a short period of time with a small number of participants; serious questions about generalizability remain.

Choice

- Parental satisfaction levels are high in virtually all voucher and charter programs studied, indicating that parents are happy with the school choices made available by the programs (see, e.g., Beales and Wahl 1995; Horn and Miron 1999; Howell and Peterson 2002; Mulholland 1999; Pioneer Institute 1998; Texas Education Agency 2000; Weinschrott and Kilgore 1998). In the experimental voucher programs that have been studied for 2 successive years, levels of parental satisfaction decline slightly in the second year but remain substantially higher than those of public school comparison groups (Howell and Peterson 2002; Myers et al. 2000).

Access

- Programs that have been explicitly designed with income qualifications have succeeded in placing low-income, low-achieving, and minority children in voucher schools (Howell and Peterson 2002; Metcalf et al. 1999; Myers et al. 2000; Wisconsin Legislative Audit Bureau 2000; Witte 2000).
- On the other hand, in most choice programs (whether voucher or charter), children with disabilities and children with poorly educated parents are somewhat underrepresented (see Beales and Wahl 1995; Howell and Peterson 2002; Metcalf et al. 1999; Myers et al. 2000; Peterson, Howell, and Greene 1999; Peterson, Myers, and Howell 1999; Wolf, Howell, and Peterson 2000; Young 2000).
- Education tax subsidy programs are disproportionately used by middle- and upper-income families (see Catterall 1983; Catterall and Levin 1982; Darling-Hammond, Kirby, and Schlegel 1985).

Integration

- In communities where public schools are highly stratified, targeted voucher programs may modestly increase racial integration by putting minority children into voucher schools that are less uniformly minority, without reducing integration in the public schools (see, e.g., Fuller and Mitchell 1999; Fuller and Mitchell 2000; Howell and Peterson 2002).

- Limited evidence suggests that, across the nation, most charter schools have racial/ethnic distributions that probably fall within the range of distributions of local public schools. In some states, however, many charter schools serve populations that are racially homogeneous (North Carolina Department of Public Instruction 1998; RPP International 2000).
- Evidence from other school-choice contexts, both in the United States and abroad, suggests that large-scale, unregulated choice programs are likely to lead to some increase in stratification (Ladd and Fiske 2001; McEwan and Carnoy 1999; Willms 1996).

Civic Socialization

- Virtually nothing is yet known empirically about the civic socialization effects of voucher and charter schools.

What We Don't Know

The brevity of the above list of findings should send a note of caution to policymakers and to supporters and opponents of choice. For most of the key questions, direct evaluations of vouchers and charter schools have not yet provided clear answers, and the list of unknowns remains substantially longer than the list of knowns, as summarized below:

For most of the key questions, direct evaluations of vouchers and charter schools have not yet provided clear answers, and the list of unknowns remains substantially longer than the list of knowns.

Academic Achievement

Unknowns in the realm of academic achievement include, first, an explanation for the possible voucher advantage for African American children. In addition, the academic effectiveness of charter schools must be examined in a larger number of states over a longer period of time. Long-term effects on both achievement and attainment in both voucher and charter programs are as yet unexamined. Moreover, we have little information that would permit us to compare the effectiveness of vouchers and charter schools with other, more conventional reforms, such as class-size reduction, professional development, high-stakes accountability, and district-level interventions. Finally, the systemic effects—positive or negative—of both voucher and charter programs have yet to be clearly identified. Whether the introduction of vouchers/charter schools will help or harm the achievement of students who stay in conventional public schools remains largely uncertain, although a recent study suggests favorable effects (Hoxby 2002). This is perhaps the most important achievement issue, because most students are likely to be "nonchoosers" and remain in conventional public schools.

Choice

The most important unknown related to parental liberty concerns the quality and quantity of the supply of schools made available by voucher and charter programs. The number of high-quality alternatives that different varieties of voucher and charter programs will produce is for the moment highly speculative.

Access

Critical unanswered questions about access to voucher and charter schools relate to the variability that would result from different kinds of programs. The characteristics of voucher students in existing programs differ from those of charter students, and the characteristics of charter students vary across states. Other programs might differ further still in terms of the access they provide to different groups of students. In particular, many varieties of vouchers may be used disproportionately by middle- and upper-income families.

Integration

The effects of voucher and charter programs on the sorting of students across schools have not been well explored thus far. Although studies have produced extensive amounts of demographic data on the students participating in voucher and charter programs, very few studies have provided school-level information on both voucher/charter schools and local public schools, linked to information on individual students—which is essential to understand dynamic integration effects. Even a direct comparison of school-level integration in voucher/charter schools and conventional public schools does not tell us how the introduction of a voucher/charter policy changes levels of integration across schools. A full understanding of integration effects requires a clear assessment of all possible counterfactuals. Where would students of different racial/ethnic groups be in the absence of vouchers/charter schools? Different answers to this question imply very different effects for vouchers and charter schools. Would the students attend local public schools, would they pay tuition at racially homogeneous private schools, would their families move to the suburbs to enable them to attend racially homogeneous public schools, or would they be schooled at home? Unfortunately, no studies of vouchers or charter schools have provided the kind of dynamic analysis that would produce clear answers.

Civic Socialization

Despite the fact that civic socialization is commonly recognized as a critical public purpose of the educational system, next to nothing is known about the relative effectiveness of voucher, charter, and conventional public schools in socializing students to become responsible citizens. The best available evidence is far short of that which is available to assess each of the other outcome dimensions. The slim evidence that is available provides little support for the view that existing private schools do any worse than public schools, on average, at socializing citizens (Campbell 2001).

Specific variations in the details of voucher/charter policies are likely to make a big difference to many of the empirical outcomes. Program scale is one variable that is likely to be especially important.

Implications for Policy

The Significance of Scale

Specific variations in the details of voucher/charter policies are likely to make a big difference to many of the empirical outcomes. Program scale is one variable that is likely to be especially important.

Nearly all of the existing empirical evidence on the effects of vouchers and charter schools comes from relatively small-scale programs. Many of the existing voucher programs are "escape valves" that are targeted to a small number of at-risk children. For these programs, most of the existing evidence is neutral or somewhat favorable: they provide valued new choices to low-income families and may provide achievement benefits to African American children. Although we know little about empirical effects in other dimensions—including integration, civic socialization, and cost—it seems unlikely that escape-valve programs would result in major harms on any of these dimensions. In brief, in some contexts—such as high-poverty cities with substantial African American populations, or communities that have underperforming public schools—targeted voucher programs may produce discrete benefits. Such programs will not be the silver bullet that will rescue urban education, but they are unlikely to produce the negative consequences that voucher opponents fear.

Evidence on existing charter laws is harder to summarize because variation across states is dramatic, in terms of both the provisions of the laws and the observed empirical effects. Existing charter schools frequently satisfy a parental demand, and they are producing academic results that are mixed but show signs of promise. Other effects are ambiguous or unknown.

The implications of the findings for larger scale choice programs, however, are unclear. Generalizing from evidence on small voucher/charter programs to infer the outcomes of large-scale choice programs is not easy, for several reasons. First, the voucher experiments that provide some of the best evidence on achievement effects are "black boxes"—they do not allow a look "inside" to explain the mechanisms that produce the (apparent) achievement advantage for low-income African American children who use vouchers. The range of possible explanations for the observed achievement difference is wide, and different explanations have profoundly different implications for whether the effect would be reproduced in a larger scale program. If, for example, African American voucher students have benefited only because the voucher program put them in classrooms with high-achieving peers, then the effect might disappear in a larger scale program in which large numbers of low-achieving students end up in voucher classrooms together. Similarly, if the experimental advantage is attributable to a context of underperforming public schools, then a universally available alternative might show no advantage when compared to a broader range of higher performing public schools. Other mechanisms that could explain the experimental findings may be more easily duplicated on a larger scale. Until we understand the source of the experimental findings, however, we cannot know whether they will apply to larger scale programs.

Similar issues arise with respect to the studies of achievement in charter schools. The existing charter studies show mixed results, with some agreement that academic performance is lowest in the first year of a charter school's existence. Programs that seek to open large numbers of new charter schools should not expect high achievement in the short term.

Even if small-scale programs are theoretically generalizable, programs in the process of scaling up often encounter unexpected difficulties.

Empirical effects on the dimensions of access and integration will almost certainly differ for large-scale programs. Most existing voucher programs serve low-income or other at-risk children because they are explicitly designed to do so, with eligibility tied to income or to the performance of the local public school. Universally available voucher programs, by contrast, may disproportionately benefit highly educated and upper-income families who have the means to take advantage of them, particularly if the programs are funded at low levels and permit supplemental tuition payments. Similarly, large-scale choice programs (whether vouchers or charters) are more likely to undermine school-level integration than are "escape-valve" vouchers that put low-income children in existing private schools.

The economic costs of large-scale voucher/charter programs are also highly unpredictable. They depend not only on the details of policy design, but also on the "takeup rate": the number of students who switch schools to participate in the program. Costs will go up if students switch into higher cost schools, but costs could actually decline if students switch from higher cost to lower cost schools. Escape-valve programs provide little guidance about the takeup rate of universally available programs.

Even if small-scale programs are theoretically generalizable, programs in the process of scaling up often encounter unexpected difficulties. Scale-up often results in a distortion of the original conditions by which the program was effective. Newly established voucher/charter schools may or may not be as effective as preexisting private schools. High-quality, nonprofit providers (including religious institutions) may lack the capacity and incentive to expand, and the supply may be filled largely by for-profit school operators—whose effectiveness is as yet unknown.

On the other hand, vouchers and charter schools may in some respects be easier to scale up because they can be uniquely sensitive to local needs and desires. They are chosen and implemented at the school level, rather than imposed from above, which makes them fully compatible with all school-level programmatic reforms. In consequence, they may bypass at least a few of the implementation and scale-up problems that have undermined a wide variety of educational reforms over the last 30 years. Whether they will succeed in doing so—and in producing the achievement, access, liberty, integration, and socialization outcomes desired from our schools—remains to be seen.

A Note on Universal-Choice Systems

The most ambitious voucher/charter programs would replace the existing system of educational governance and finance with an entirely new system in which all schools are autonomous and every family must choose a school. Direct evidence on these highly ambitious proposals is very limited, because they have never been fully implemented in the United States.

Universal-choice systems would, of course, encounter many of the implementation challenges described above. In addition, however, because such proposals would directly change the entire educational system, they have the potential to create larger effects—both positive and negative—than other varieties of programs. Systemic effects would not merely be the indirect result of competition or creaming, but would follow directly from the changes to all public schools. These proposals, therefore, could create either the greatest benefits or the greatest harms. Care in the details of design might permit the construction of a universal-choice program which could avoid negative consequences and perhaps produce substantial benefits; but predicting such benefits depends, for now, on theory rather than existing evidence.

Considerations in Policy Design

Despite the large number of uncertainties that remain about the empirical effects of vouchers and charter schools, it is possible to provide some guidance about the intelligent design of the details of voucher/charter programs. Policymakers who are considering voucher or charter laws can maximize program benefits and mitigate harms through thoughtful policy design. Here we consider a series of questions that address the relationship between policy details and empirical effects in each of the five key outcome dimensions. Because tradeoffs among desired outcomes may sometimes be necessary, the ideal design depends to some extent on how policymakers value and rank the various outcomes promoted by the educational system. Nevertheless, the relationship among outcomes is sometimes complementary rather than competitive; a few of the same policy prescriptions can serve multiple purposes. The prescriptions below should be considered tentative rather than definitive; they are promising policy options based on plausible extrapolation from the available evidence:

These prescriptions should be considered tentative rather than definitive; they are promising policy options based on plausible extrapolation from the available evidence.

How might policymakers maximize the likelihood that voucher/charter schools will be academically effective?

- Include existing private and parochial schools.
- Enforce requirements for testing and information dissemination.
- Don't skimp on resources.

How might policymakers maximize the likelihood that systemic effects on non-choosers are positive rather than negative?

- Establish communication among schools.
- Impose consequences on schools that do not perform at acceptable levels.
- Give the public schools the autonomy to act competitively.
- Require open admissions.
- Require all students to choose.

How can policymakers ensure that a substantial number of autonomous schools are available?

- Permit existing private and parochial schools to participate.
- Provide generous funding.
- Avoid overregulation.
- Do not make the local school district the exclusive chartering authority.

How can policymakers ensure that autonomous schools serve low-income and special-needs children?

- Actively disseminate information about schools.
- Target specific children.
- Forbid tuition add-ons.
- Provide generous funding.
- Use an equitable funding method.
- Provide supplemental funding for students with special needs.
- Require open admissions.

How can policymakers promote integration in programs of autonomous schooling?

- Require open admissions.
- Target communities with racially homogenous public schools.
- Include existing private and parochial schools.

- Reward integration financially.

How can policymakers ensure that voucher/charter schools are effectively socializing their students to become responsible citizens of our democracy?

- Disseminate information about mission, values, curriculum, and outcomes.

Conclusion

Our review of the evidence leaves us without a crisp, bottom-line judgment of the wisdom of voucher and charter programs. Prudent observers would note that, at the current scale of such efforts, many important questions cannot be answered at all, notably those concerning total demand, supply response, school characteristics and performance at scale, or final impact on public schools in the new equilibrium. Moreover, in important respects—notably civic socialization—the effects of current or proposed autonomous schools are virtually unknown. And design is crucial: autonomous school policy can be targeted or not, regulated or not, generously funded or not, inclusive of existing providers or not. Each of these policy levers has important implications for student outcomes. A program of vigorous research and experimentation is called for, but not one confined to choice programs: better information is needed on the performance of conventional public schools and alternative reform models as well. In the meantime, political decisions will undoubtedly be made, for and against vouchers and charter programs. They will be informed by good evidence, one hopes, but not fully justified by it for many years to come.

References

Beales, J.R., and Wahl, M. (1995). Private Vouchers in Milwaukee: The PAVE Program. In T.M. Moe (Ed.), *Private Vouchers*. Stanford, CA: Stanford University, Hoover Institution Press.

Bettinger, E. (1999). *The Effect of Charter Schools on Charter Students and Public Schools* (Occasional Paper No. 4). National Center for the Study of Privatization in Education. Available: *http://www.tc.columbia.edu/ncspe/* (retrieved February 8, 2001).

Campbell, D.E. (2001). Making Democratic Education Work. In P.E. Peterson and D.E. Campbell (Eds.), *Charters, Vouchers, and Public Education*. Washington, DC: Brookings Institution.

Catterall, J.S. (1983). Tuition Tax Credits: Issues of Equity. In T. James and H.M. Levin (Eds.), *Public Dollars for Private Schools: The Case of Tuition Tax Credits* (pp. 130–150). Philadelphia: Temple University Press.

Catterall, J.S., and Levin, H.M. (1982). Public and Private Schools: Evidence on Tuition Tax Credits. *Sociology of Education, 55*(April/July): 144–151.

Darling-Hammond, L., Kirby, S.N., and Schlegel, P.M. (1985). *Tuition Tax Deductions and Parent School Choice: A Case Study of Minnesota* (R-3294-NIE). Santa Monica, CA: RAND.

Fuller, H.L., and Mitchell, G.A. (1999). *The Impact of School Choice on Racial and Ethnic Enrollment in Milwaukee Private Schools*. Marquette University, Institute for the Transformation of Learning. Available: *http://www.schoolchoiceinfo.org/* (retrieved October 20, 2000).

Fuller, H.L., and Mitchell, G.A. (2000). *The Impact of School Choice on Integration in Milwaukee Private Schools*. Marquette University, Institute for the Transformation of Learning. Available: *http://www.schoolchoiceinfo.org/servlets/SendArticle/4/integ1299.pdf* (retrieved October 20, 2000).

Gill, B.P., Timpane, P.M., Ross, K.E., and Brewer, D.J. (2001). *Rhetoric Versus Reality: What We Know and What We Need to Know About Vouchers and Charter Schools*. Santa Monica, CA: RAND.

Greene, J.P. (2000). *The Effect of School Choice: An Evaluation of the Charlotte Children's Scholarship Fund* (Civic Report 12). New York: Center for Civic Innovation at the Manhattan Institute.

Gronberg, T.J., and Jansen, D.W. (2001). *Navigating Newly Chartered Waters: An Analysis of Texas Charter School Performance*. San Antonio, TX: Texas Public Policy Foundation.

Horn, J., and Miron, G. (1999). *Evaluation of the Michigan Public School Academy Initiative*. Kalamazoo, MI: Western Michigan University, The Evaluation Center.

Howell, W.G., and Peterson, P.E. (2002). *The Education Gap: Vouchers and Urban Schools*. Washington, DC: Brookings Institution.

Howell, W.G., Wolf, P.J., Campbell, D.E., and Peterson, P.E. (2002). School Vouchers and Academic Performance: Results From Three Randomized Field Trials. *Journal of Policy Analysis and Management, 21*, 191–217.

Hoxby, C.M. (2002). *School Choice and School Productivity (Or, Could School Choice Be a Tide That Lifts All Boats?)* (Working Paper No. 8873). Cambridge, MA: National Bureau of Economic Research.

Ladd, H.F., and Fiske, E.B. (2001). The Uneven Playing Field of School Choice: Evidence From New Zealand. *Journal of Policy Analysis and Management, 20*(1): 43–64.

Mayer, D.P., Peterson, P.E., Myers, D.E., and Tuttle, C.C. (2002). *School Choice in New York City After Three Years: An Evaluation of the School Choice Scholarships Program* (8404-045). Princeton, NJ: Mathematica Policy Research.

McEwan, P.J., and Carnoy, M. (1999). *The Impact of Competition on Public School Quality: Longitudinal Evidence from Chile's Voucher System*. Stanford, CA: Stanford University.

Metcalf, K.K., Muller, P.A., Boone, W.J., Tait, P.A., Stage, F.K., and Stacey, N.M. (1999). *Evaluation of the Cleveland Scholarship Program: Second-Year Report*. Bloomington, IN: Indiana University, Junior Achievement Evaluation Project.

Mulholland, L.A. (1999). *Arizona Charter School Progress Evaluation*. Phoenix, AZ: Morrison Institute for Public Policy.

Myers, D., Peterson, P., Mayer, D., Chou, J., and Howell, W.G. (2000). *School Choice in New York City After Two Years: An Evaluation of the School Choice Scholarship Program* (Interim Report 8404-036). Princeton, NJ: Mathematica Policy Research, Inc.

North Carolina Department of Public Instruction. (1998). *Selected Characteristics of Charter Schools, Programs, Students, and Teachers 1997–98* (Report I of the Charter Schools Evaluation). Durham: North Carolina Department of Public Instruction, Division of Accountability Services.

Peterson, P.E., Howell, W.G., and Greene, J.P. (1999). *An Evaluation of the Cleveland Voucher Program After Two Years.* Cambridge, MA: Harvard University, John F. Kennedy School of Government, Program on Education Policy and Governance.

Peterson, P.E., Myers, D., and Howell, W.G. (1999). *An Evaluation of the Horizon Scholarship Program in the Edgewood Independent School District, San Antonio, Texas: The First Year*. Princeton, NJ: Mathematica Policy Research.

Pioneer Institute. (1998). *Poll Finds Higher Satisfaction Rate Among Charter School Parents*. Pioneer Institute. Available*: http://www.pioneerinstitute.org* (retrieved July 29, 2000).

RPP International. (2000). *The State of Charter Schools: Fourth-Year Report*. U.S. Department of Education. Washington, DC: U.S. Government Printing Office.

Solmon, L., Paark, K., and Garcia, D. (2001). *Does Charter School Attendance Improve Test Scores? The Arizona Results*. Phoenix, AZ: Goldwater Institute Center for Market-Based Education.

Texas Education Agency. (2000). *Texas Open-Enrollment Charter Schools: Third Year Evaluation, March 2000*. Texas Education Agency, Division of Charter Schools. Available: *http://www.tea.state.tx.us/charter/eval99/index.html* (retrieved June 2, 2000).

Weinschrott, D.J., and Kilgore, S.B. (1998). Evidence From the Indianapolis Voucher Program. In P.E. Peterson and B.C. Hassel (Eds.), *Learning From School Choice*. Washington, DC: Brookings Institution.

Willms, J.D. (1996). School Choice and Community Segregation: Findings From Scotland. In A.C. Kerckhoff (Ed.), *Generating School Stratification: Toward a New Research Agenda*. Boulder, CO: Westview Press.

Wisconsin Legislative Audit Bureau. (2000). *Milwaukee Parental Choice Program: An Evaluation*. Madison, WI: Legislative Audit Bureau.

Witte, J.F. (2000). *The Market Approach to Education*. Princeton, NJ: Princeton University Press.

Wolf, P., Howell, W., and Peterson, P. (2000). *School Choice in Washington, D.C.: An Evaluation After One Year*. Cambridge, MA: Harvard University, John F. Kennedy School of Government, Program on Education Policy and Governance.

Young, B.A. (2000). *Characteristics of the 100 Largest Public Elementary and Secondary School Districts in the United States: 1998–1999* (NCES 2000–345). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Getting the Biggest Bang for the Educational Buck: An Empirical Analysis of Public School Corporations as Budget-Maximizing Bureaus

Anthony Rolle
Vanderbilt University

About the Author

Dr. Anthony Rolle has conducted educational finance and policy research for the University of Washington's Institute for Public Policy & Management, the Washington State Legislature and Democratic House Majority Whip, the Indiana Education Policy Center, the National Education Association, and North Carolina State University. Currently an Assistant Professor of Public Policy and Education in the Department of Leadership, Policy, and Organizations at Vanderbilt University's Peabody College, Dr. Rolle's academic interests include

- Continuing research and development of alternative measures of economic productivity in public schools;
- Continuing research examining levels of educational equity provided through state funding mechanisms in public schools; and
- Research examining the economic importance of public schooling in Tennessee and in the new global economy.

Dr. Rolle received a Bachelor of Science in Political Science from Santa Clara University; a Master's Degree in Public Administration from the University of Washington's Graduate School of Public Affairs; and a Ph.D. in Educational Policy Analysis from the School of Education at Indiana University.

Getting the Biggest Bang for the Educational Buck: An Empirical Analysis of Public School Corporations as Budget-Maximizing Bureaus

Anthony Rolle

Vanderbilt University

Acknowledgments

First, I would like to thank my family: my parents, brother (Jeff), sister (Amber), uncles (Allan, Livingston, Robert, Samuel, and William), aunts (Lillian, Mary, and Shirley), and cousins—especially Nichol (who helped me compose and correct many, *many* regression tables). They each supported me in their own way—thank you all.

I also would like to thank my dissertation committee. Without the guidance, wisdom, and opportunities provided by these men, my journey would have been much longer and far less gratifying. I will do my best to continue making all of you proud.

I also would like to thank Bob Arnove, Bill Becker, Steve Franz, Gayle Hall, Martha McCarthy, and Eduardo Rhodes, as well as other faculty, staff, and students in the Education Policy Studies Program, the Indiana Education Policy Center, and the Student Academic Center. Each of these people and programs have contributed to my growth as a person and a professional in ways too numerous to count. Thank you.

This list would not be complete if I did not extend it to include Richard Brandon, Julius Debro, Caroline Maillard, and Cynthia Morales of the University of Washington; Martin Cooke, Elena Dorabji, and Peggy Downes of Santa Clara University; and Patricia Franklin, Wilma Hooper, Lucille Sapiens, Edward Sims, Mike Therrell, Jose Valdes, and Susan Wong from the East Side Union High School District in San Jose, California. Your image of me still is what I hope to be—thank you all.

Introduction

Profit-seeking organizations—deriving substantial proportions of their revenues from consumer purchases—generate dollars in a different manner than public bureaus, which derive a large portion of their revenues from taxing authorities. Consequently, the economic incentives faced by public bureaus differ from those facing private firms in ways that do not generate expected cost-effective behaviors (Barnett 1994). Still, despite the differences in organizational and economic incentives between managers of public sector organizations and private sector firms, educational researchers (e.g., Hanushek 1986; Kirst 1983; Mann and Inman 1984; Rossmiller 1987; Walberg and Fowler 1987; Walberg and Walberg 1994) are committed to improving the results of schooling through the use of traditional economic analyses. Yet this type of research has shown that examining educational bureaus using the cost-minimizing assumptions of traditional economic theory provides mixed results at best.

The inconsistent analytical results generated by the examination of educational finance issues when using traditional economic assumptions—a relationship where financial inputs are assumed to be minimized while desired outputs are maximized—leaves a number of important questions about the nature, productivity, and efficiency of public school districts unanswered (Hentschke 1988):

- What type of incentives and constraints influence the expenditure behavior of public school districts?
- What organizational objectives are pursued by public school districts? Among these, which are maximized or optimized?
- To what extent, and under what circumstances, are individual or bureaucratic desires reflected in the organizational outcomes generated by public school districts?

The goal of this research paper is to create a common understanding about the efficient uses of public education dollars.

In light of commonly used economic analyses, these questions become particularly important given that trends in education seem to be exemplified by continued increases in organizational size, fiscal resources, and decreases in educational outcomes (Bennett 1992; Hanushek 1995; Sowell 1993).

Within the traditional economic framework, public school districts generally are labeled as economically inefficient organizations. These assertions are supported primarily by the absence of strong production function relationships. Little is known, though, about the efficiency of educational organizations when examined outside of the traditional cost-minimization framework. Consequently, the purpose of this research is to contribute to our knowledge about the efficiency of public schools by examining empirically the theory of budget-maximizing bureaucratic behavior (Niskanen 1968, 1971, 1973, 1975, 1991, 1994). Specifically, this research will examine Indiana public school corporations (i.e., school districts) to determine if they act as budget-maximizing bureaucracies; and if so, examine whether these school corporations produce educational outcomes in a manner that is economically inefficient.

Ultimately, the goal of this research paper is to create a common understanding about the efficient uses of public education dollars. Using a budget-maximizing economic theory to analyze educational bureaus provides a different, but important, perspective when examining

1. The cost-minimization assumption applied commonly to economic models of education;
2. The nature of the input-output relationship assumed to apply to educational bureaus;
3. The concepts of efficiency as they apply to educational bureaus;
4. The time-lagged effects of prior years' educational outcomes on current year budgets; and
5. The pervasive use of simple linear production functions to predict educational outcomes.

With this increased level of understanding, policymakers and the public can begin to address the more complex issue of improving the use of public resources to produce higher levels of organizational outcomes.

Background

Prior to Friedman's (1962) assertion that "educational free markets" would be more efficient at allocating educational resources than a system of public sector bureaucracies, Mises (1944) provided some of the earliest critical insights to economic theory as it related to public organizations:

> [Public] bureaus specialize in the supply of those services whose value cannot be exchanged for money at a per-unit rate . . . Consequently, bureaus cannot be managed by profit goals and the traditional economic incentives (pp. 47–49).

According to Niskanen (1971), Tullock (1965) developed the theory for a "maximizing bureaucrat" that examined personal relations and advancement procedures within public bureaus; but he did not use the general applicability of his theory to investigate budget and output behavior. Similarly, Downs (1967) focused primarily on behavior within public bureaus. He developed a comprehensive theory of internal management processes but also stopped short of investigating the consequences of maximizing behavior as it relates to budget and output performance.

Although the ideas of Mises, Tullock, and Downs form the basis for a theory that addresses budget-maximizing bureaucratic behavior, Niskanen (1968) was the first to specifically address the questions answered by traditional economic theory:

- Given differing levels of demand and supply, how much output is produced at what cost?
- How do output levels and costs vary under these changing economic conditions?
- Is the output produced efficiently or inefficiently?

The theory of budget-maximizing bureaucratic behavior states that bureaucrats attempt to maximize the agency's total budget during their tenure.

Specifically, the theory of budget-maximizing bureaucratic behavior states that subject to a budget constraint greater than or equal to the costs of supplying the output expected by a public bureau's sponsors, bureaucrats attempt to maximize the agency's total budget during their tenure (for a more complete explanation, see *Appendix*). In other words, lacking the lure of performance-based salaries and benefits as rewards for increasing organizational efficiency, public sector managers—acting as self-interested individuals seeking to maximize their own welfare—attempt to maximize their nonpecuniary benefits (e.g., prestige, scope of activities, or perquisites) through the pursuit of larger agency budgets. As a result of this budget-maximizing behavior, Niskanen hypothesizes that public bureaus generate

- budgets that are larger than optimal;
- output that may be too low when compared to expenditure levels; and
- output that is produced inefficiently.

Niskanen advances two arguments in support of his assertion that public managers act as budget-maximizing bureaucrats: *rationality* and *survival*. He claims that by personality or preparation, public administrators strive to serve their perception of the public interest. However, they cannot acquire all the information on individual preferences and production opportunities necessary to determine the public interest. In addition, he claims public administrators do not have the authority to order an action contrary to the different perceptions of the public interest held by other bureaucrats or higher level government officials. Rationally, therefore, bureaucrats must pursue preferences through the acquisition of larger and larger budgets in order to maximize their personal utility (Niskanen 1973, p. 23).

Niskanen's survival argument also suggests bureaucrats seek to maximize budgets. Two groups of people significantly influence an administrator's tenure in public office: the employees of the agency and the sponsors of the agency (e.g., taxpayers, municipal government officials, and state legislators). He claims individuals employed by public organizations not only desire budget-maximization for reasons similar to those of the bureau administrators but also can influence the agencies to seek increased budgets. Employees of public bureaus can be cooperative, responsive, and work effectively; or, they can deny information to the public sector manager in order to undermine directives. Niskanen believes the behavior of this group depends greatly on their perceived rewards for employment with the bureau. Consequently, the bureaucrat who seeks operating efficiencies without budget increases will have difficulty "buying the cooperation" of employees (Niskanen 1973, p. 24).

Niskanen asserts that at each stage of the budgetary review process, sponsors—due to a lack of time, information, and staff necessary to monitor programs—depend on public managers to propose new programs while advocating for the maintenance of existing programs. He believes this dependency is due to the fact that the total activities and budget of most bureaus are beyond the comprehensive understanding of people who are not involved directly with the organization. Consequently, sponsors focus most of their attention advocating for budget changes and reveal their spending priorities by approving of—or disapproving of—different portions of budgets. When the preferences of sponsors are realized, bureau executives are nominated, confirmed, and supported by them repeatedly; on the other hand, bureau executives are forced to resign when sponsor preferences are not realized. Therefore, bureaucrats must attempt to realize sponsor preferences through the acquisition of larger and larger budgets or face resignation (Boyd and Hartman 1988).

There is ample evidence that bureaucrats systematically request larger budgets regardless of the level of organizational output generated. This idea has been relatively unchallenged since the claim was made (Wildavsky 1964). Wildavsky claimed bureaucrats request moderate annual budget increases in order to maximize long-term budget goals. Bush and Denzau (1977)—supported later by Lynn (1991)—found evidence that a majority of bureaucrats want and ask for increased budgets. Similarly, Blais and Dion (1991) found that bureaucrats tend to vote for political parties that favor state intervention. Young (1991) found that there is little relationship between growth of bureaus and bureaucrats' salaries. Aucoin (1991) found that budget controls are put into place by sponsors and legislators because there exists a belief that bureaus always will attempt to increase their budgets. Kiewiet (1991) found that school superintendents hold tax rates as high as possible without having to obtain voter approval. Finally, Campbell and Naulls (1991) found that bureaucrats seek larger budgets because of their values—regardless of self-interest in salary. It is this type of research literature—historical and contemporary work supporting the idea that bureaucrats believe it is in their best interest to obtain increased budgets—that affirms the necessity for exploring economic models and theories employing budget-maximizing frameworks as opposed to the cost-minimization assumptions of traditional economic analyses.

There is ample evidence that bureaucrats systematically request larger budgets regardless of the level of organizational output generated.

Research Methodology

Examining Niskanen's theory of budget-maximizing bureaus within the state of Indiana should be extremely enlightening because of the state's history of preferring low state tax rates combined with prudence in budget allocations. Prior to 1973, the state used a foundation program approach to allocate monies to its 294 school corporations. Under this approach, the state provided one-third of total education funding, with local property taxes providing the remaining two-thirds. The state formula was calculated as the difference between a common revenue level and the yield of a common property tax. In order for school corporations to generate extra revenue, high property tax levies were needed to obtain the desired revenue (Lehnen and Johnson 1989).

In 1973, the state legislature took action to control property taxes by freezing property tax levies. This action resulted in a decreased reliance on support provided locally and an increased reliance on state funds. The property tax reform program also had an unanticipated effect: It increased inequities in revenues that had been developing prior to 1973, because the property tax legislation froze an inequitable funding system in place (Lehnen and Johnson 1989). Between 1979 and 1986, the Indiana General Assembly used a combination of flat grants, percentage increases, and combinations of the two methods to increase revenues for school corporations. Despite these revenue generating measures, disparities continued at the local levels even though state aid became the dominant source of school support, providing approximately two-thirds of education funding (Bauer 1992). In 1986, recognizing that there were large disparities in the per-pupil revenues of school corporations, the state legislature provided "bottom-up" support through per-pupil allocations to ensure that low-spending corporations could increase their revenues.

Tired of waiting for the General Assembly to take meaningful action on school finance issues, the Lake Central School District filed suit in 1987 alleging the following:

1. The state of Indiana failed to provide "for a general, and uniform system of Common Schools" as outlined in the state constitution (Ind. Const. art. 8, § 1); and
2. The state of Indiana violated the Equal Protection clause outlined in the state constitution (Ind. Const. art. 1, § 23) by granting property-rich school corporations the privilege or option of generating more revenue than property-poor.

Later, forming a group called Schools Allied for Funding Equity (SAFE), the Lake Central plaintiffs also charged that state property tax limitations take fiscal control away from local school districts while aggravating funding disparities across districts.

In 1991, the House Select Committee on Primary and Secondary Education began meeting to "create legislation addressing both improvements in finance equity and general education reform" (Bauer 1992). The key component of the proposed legislation—to be enacted in the 1994 academic year—was the creation of a new funding formula that achieves the joint goals of reducing disparities in spending and tax effort. In 1993, SAFE agreed to have their lawsuit "dismissed without prejudice" with promises from then Governor Evan Bayh and the General Assembly that the year's legislative session would seek a more equitable education funding formula (Rolle 1994).

It is within this sociopolitical context—Indiana's economically conservative one—that Niskanen's theory will be examined.

Indiana's 1993 "Reward for Effort" funding system—which was to be phased in over a period of at least 6 years—seeks to provide revenue to school corporations in a more equitable manner. This approach is designed to provide low-revenue school corporations with access to higher assessed valuations per student. Despite the formula's emphasis on increasing revenues for low-spending schools, a number of revenue and tax rate limitations constrain the ability of the new school funding formula to achieve high levels of interdistrict equity quickly (Theobald and Rolle 1995). It is within this sociopolitical context—Indiana's economically conservative one—that Niskanen's theory will be examined.

Data Description

The data obtained and examined originally in this research span 10–20 academic years, 1981 through 1997, depending upon the variable examined. The variables are defined, calculated, and reported in the *Accounting and Uniform Compliance Guidelines Manual for Indiana Public School Corporations* (Indiana State Board of Accounts 1998), published jointly by the Indiana State Board of Accounts and the Indiana Department of Education's Center for Administration & Fiscal Measurement. Additionally, the study uses expenditure, output, and demographic variables specified by the state of Indiana's Performance-Based Accreditation System legislation (Ind. Code § 20-1-1.2 [1998]) and the Indiana Department of Education's Center for Assessment, Research, and Information Technology. Due to the embryonic state of Indiana's still-developing comprehensive reporting system, the number and type of school corporation variables examined are not as expansive as desired in the theoretical construction of this research. For example, college attendance rates and Scholastic Aptitude Test (SAT) scores were not available in order to provide more objective measures of school outcomes.

A few clarifications are necessary regarding some of the variables examined. The measurement of the variable representing teacher experience included *all* certified employees prior to 1986—not only classroom teachers. In lieu of actual statistics on remediation, summer school enrollment as a percent of the annual district enrollment is used as a proxy measure for the percent of students remediated. Dollars of Indiana School Incentive Awards (ISIA) received at the district level is used as a proxy for overall quality of a school corporation. More specifically, in 1989, the ISIA program began granting monetary compensation to schools showing improvement over their prior 3-year average in academic performance and attendance. Each school competes only against its own performance averages. Cash awards—based on an annual appropriation from the state assembly—go to individual schools that improve in at least two of four areas: Indiana Statewide Test for Educational Progress (ISTEP) total battery scores, ISTEP language scores, ISTEP mathematics scores, and average daily attendance.

A special comment needs to be made regarding the exclusion of student achievement test scores as a primary outcome variable in this analysis. Despite the emphasis placed on standardized achievement test scores as a key outcome variable in education finance research, Berlin and Sum (1988)—re-emphasized in Levin and Kelly (1994)—used multivariate analyses to show that completing high school (typically measured by graduation rates) is more important economically than a student gaining an additional grade equivalent on a standardized test. This research evidence—combined with concerns that standardized exams are biased against both low-income and minority students—was the major factor that led to the exclusion of test scores as a primary outcome variable. Additionally, standardized achievement data—California Achievement Test (CAT), California Test for Basic Skills (CTBS), and ISTEP scores—were not reported by the Indiana Department of Education in a manageable (or malleable) form. Nevertheless, a form of standardized achievement still is being measured in this analysis because standardized achievement scores are included as part of Indiana's incentive grant calculation.

Completing high school (typically measured by graduation rates) is more important economically than a student gaining an additional grade equivalent on a standardized test.

Operationalization of Data

The data reviewed above form the bases for the operationalization of the variables examined in this research. After careful consideration, some variables remained unchanged; others were combined, modified, or excluded from the analysis based on the integrity of the data. Ultimately, the operational variables were assigned to one of three data categories: expenditure, outcome, or demographic. Expenditure variables represent the budget amounts that sponsors are willing to spend on education in a particular school corporation. The number of expenditure variables were reduced to include only total expenditures. Outcome variables represent the student results that sponsors are monitoring to determine how well educational services are being delivered by a school corporation. The number of outcome variables were reduced to include (1) average daily attendance rates; (2) percent of 12th-graders graduating; (3) percent of students remediated; and (4) school corporation quality. The average daily attendance rate was calculated by dividing the average daily attendance by the number of students enrolled.

The graduation rate was calculated by dividing the number of high school graduates by the number of students enrolled in the 12th grade. The percentage of students remediated was calculated by dividing summer school enrollment by the number of students enrolled during the regular academic year. School corporation quality was measured by dividing the total amount of ISIA grants allocated to a school corporation by the number of students enrolled in the district.

Demographic variables represent socioeconomic conditions that affect budget expenditures, delivery of educational services, and student outcomes. Along with variables representing city type and student enrollment, the number of demographic variables were reduced to include (1) average years of teacher experience; (2) average age of teachers; (3) number of single parent households per student; (4) student/teacher ratio; (5) median family income; (6) percent of population without high school diplomas; and (7) percent of Asian, Black, Hispanic, Native American, and White students enrolled.

Only regular school corporations are included in the analysis—special education, vocational, and cooperative school districts were excluded due to lack of data. The percentages of Asian, Black, Hispanic, and Native American students were calculated by dividing the number of each type of student by the total number of students enrolled. The percentage of White students was calculated by subtracting the total percentage of minority students from 100 and will be used in lieu of individual race categories. Unfortunately, the percentages of single parent households in a school corporation were unavailable. In the absence of this data, the number of single parent households per student is used as a substitute.

The basic budget-output model—consisting of an expenditure variable regressed onto linear and quadratic organizational outcome variables—is determined by Niskanen's original theory. The use of demographic variables is not; therefore, before a final operationalization of these variables was made, correlational analyses were used to reduce the amount of collinearity between them. Teacher age and teacher experience were correlated highly $(r \approx 0.9)$; consequently, teacher age was removed from further analysis. Also, a high correlation $(r \approx -0.7)$ existed between median family income and percent of individuals without a high school diploma. Finally, there was a moderate interaction $\{r : r \in [0.3, 0.6]\}$ across years between amount of teacher experience and number of single parent households per student.

Before factor analyses were conducted, the demographic variables were separated into two categories: community and school characteristics. Only the variables measuring community characteristics were subjected to principal component factor analyses with varimax rotation to develop standardized scales that measure specific combinations of demographic influences. For the 1981–85 school years, one community risk factor—family status—emerged from the analyses as a combination of median family income, percent of individuals without a high school diploma, and number of single parent households per student. For the years 1986 and beyond, when statistics on student race began to be recorded by the Indiana Department of Education, two community risk factors emerged from the analyses: (1) *family status,* consisting of inversed median family income and percent of individuals without a high school diploma; and (2) *family type,* consisting of the percentage of minority students in a school corporation and number of single parent households per student.

A major component of this research is determining whether or not Niskanen's original budget-output function represents the economic behavior of Indiana public school corporations accurately.

Therefore, four demographic variables are used in the analysis: student/teacher ratio, teacher experience, family status, and family type.

Model Specification

A major component of this research is determining whether or not Niskanen's original budget-output function represents the economic behavior of Indiana public school corporations accurately. In his presentation, Niskanen uses a simple mathematical function to represent the maximum budget that sponsors are willing to grant a bureau for an expected level of output. The function has the following properties: (1) the first derivative is a positive monotonic function over the relevant range (i.e., over some range of expected output, the sponsor is willing to grant a higher budget for a larger expected output); and (2) the second derivative is a negative monotonic function over the relevant range (i.e., over some range, the sponsor is willing to grant a higher budget for a smaller expected output).

Several types of mathematical functions share these two properties, but Niskanen uses a quadratic production function to represent the budget-output equation.

In lieu of examining every type of mathematical function that satisfies—or does not satisfy—Niskanen's mathematical assumptions, it seems appropriate to apply Weierstrass's Theorem of Polynomial Approximation (Johnson 1984). Weierstrass proved mathematically that over a closed interval any continuous function may be approximated over the whole interval by a polynomial of suitable degree. Therefore, only polynomial functions will be examined in this research. Regardless of the type (or types) of budget-output function (or functions) found to describe the economic behavior of Indiana public school corporations, they also will need to be examined while taking into account the influences of various demographic characteristics that affect education. The addition of these variables is necessary to control for different levels of student preparation. Without such controls, comparisons between districts would be unfair—similar to comparing the quality of car production between the Ford and Mercedes-Benz corporations. Therefore, the basic functional form of the four types of quadratic budget-output functions mentioned previously becomes slightly more complex and will be represented mathematically as:

$$B_t(x) = Q_{t-1}(x) ¤ D_{t-1}(x)$$

where $B_t(x)$ and $Q_{t-1}(x)$ represent the basic functional components of a budget-output function, and $D_{t-1}(x)$ represents demographic characteristics that influence both the endogenous and exogenous variables in the budget-output function.

Measuring Efficiency

An examination of average total costs and marginal costs is warranted in the private sector because most production functions are known—this phenomenon does not exist in the public sector. Efficiency in public schools, by contrast, is concerned with how much education or knowledge is delivered to—and acquired by—students, and at what cost. Similar to the private sector, being "more efficient" means one of two things when discussing education finance and economic issues:

1. increasing output levels while using the same amounts of input, or
2. maintaining output levels while using lesser amounts of input.

Additionally, public school spending is conducted such that no student's educational situation is made worse in order to improve the situation of another student. Therefore, a different measure of production efficiency is warranted. In this research, economic efficiency is measured using the modified quadriform method (Anderson 1996; Genge 1991; Genge 1992; Hickrod et al. 1989; Hickrod et al. 1990). A quadriform is an abstract tool devised to allow a hypothesized relationship to be viewed both graphically and quantitatively (figure 1).

Unlike average-marginal cost analyses, the modified quadriform examines expenditure and output variations relative to other school corporations within the state. This method seems more appropriate to apply to public schools given that a production function for education has not yet been determined. The following four terms define the economic relationships shown by the quadriform:

- *Inefficient* school corporations are those that generate lower than expected outcomes with higher than expected expenditures (quadrant 1).

Figure 1. Basic quadriform diagram

SOURCE: Adaptation from Anderson, D.M. (1996). *Stretching the Tax Dollar: Increasing Efficiency in Urban and Rural Schools.*

- *Effective* school corporations are those that generate higher than expected outcomes using higher than expected expenditures (quadrant 2);
- *Ineffective* school corporations are those that generate lower than expected outcomes using lower than expected expenditures (quadrant 3); and
- *Efficient* school corporations are those that generate higher than expected outcomes using lower than expected expenditures (quadrant 4).

In this research, expenditures are measured along the vertical axis; output is measured across the horizontal axis.

Quantitatively, the modified quadriform is constructed as a two-stage model that (1) captures the input-output relationship as two separate regressions; and (2) uses discriminant analysis to identify the "alterable" characteristics that distinguish efficient school corporations from inefficient school corporations. (This research is concerned primarily in determining the efficiency levels of public school corporations; therefore, only the first stage of modified quadriform method will be utilized.)

If Niskanen is correct, variables measuring educational outcomes should not be statistically significant predictors of total expenditures per pupil.

Mathematically, the two regression equations are of the form

$$Z_i = \alpha + \sum B_{t-i} W_{t-i} + u_{t-i}$$

where Z_i represents the expected values—expenditure or outcome—for each school corporation and W_{t-i} represents the unalterable values for each school corporation. Consequently, the Z_i's forming the expenditure and outcome regressions create the axes of the quadriform.

The regression residuals determine which of the four quadriform categories a school corporation is assigned. More specifically, the outcome regression residual values are associated with values on the *x*-axis of a Cartesian plane. At the same time, the expenditure regression residual values are associated with values on the *y*-axis of the same Cartesian plane. Each corresponding (*x,y*) pairing of residuals represents where in the quadriform (i.e., in which of the four quadrants) a specific school corporation will be designated. Theoretically, approximately 25 percent of all school corporations should fall into each quadrant.

Unfortunately, the modified quadriform discussed shows only annual efficiency categorizations among Indiana school corporations—it does not account for school corporations remaining in or changing categories over time. In order to determine the longitudinal nature of efficiency among Indiana public school corporations, an additional layer of analysis was appended to the original modified quadriform analysis. After the initial modified quadriform analysis was completed, each school corporation was given an annual value of 1 for the category within which it fell (e.g., a school corporation may be "ineffective" in 1986 and "efficient" in 1987) and annual values of 0 for the remaining three categories. Subsequently, an arithmetic mean—which will have a value between 0 and 1 for any of the quadriform categories—was taken across all years examined. As a result, a school corporation was defined as a *perennially* efficient, effective, ineffective, or inefficient producer of educational outcomes if its school corporation average was greater than 0.50 in any category. School corporations with averages less than or equal to 0.50 were excluded from further analyses. Finally, these perennially categorized school corporations were reanalyzed within a new set of quadriforms.

Data Analysis

Consistent with Niskanen's theory, the state of Indiana's K–12 public school corporations are public bureaus that promise a bundle of activities—and expected outputs based on these activities—in exchange for a tax-supplied budget. In order to test empirically whether or not the total expenditures per pupil of public school corporations can be represented by a quadratic budget-output function that is concave-down, a series of regression equations were modeled to conform to Niskanen's original hypothesis. If Niskanen is correct, variables measuring educational outcomes should not be statistically significant predictors of total expenditures per pupil. Further, variations in total per-pupil expenditures should not be explained by variations in outcome variables (i.e., the adjusted r-squared statistics of the regression models should be close to 0).

Analysis of Original Budget-Output Regressions

For 15 of the 17 years covering 1981–97, less than 10 percent of the variation in total expenditures per student can be explained by variations in average daily attendance (table 1). In 1984 and 1990, though, independent variables in the hypothesized function explained 15.9 and 33.9 percent, respectively, of the variation in total expenditures per student. The standardized coefficient for the linear term (a_{std}) of average daily attendance was positive for 6 of the 17 years while being a statistically significant predictor of total expenditures per student for 4 years. The standardized coefficient for the quadratic term (b_{std}) of average daily attendance was negative for 5 of the 17 years while being a statistically significant predictor for 4 years. Over the time

Table 1. Original form budget-output regressions for total expenditures per student and average daily attendance: 1981–97

Year	Coefficients for Q_{t-1}			Coefficients for $(Q_{t-1})^2$			Power
B_t	a	a_{std}	p-value	b	b_{std}	p-value	$AdjR^2$
1981	5,714.7	1.062	.001	–1802.8	–0.999	.002	0.031
1982	2,829.0	0.301	.468	–212.5	–0.049	.906	0.057
1983	–438.1	–0.045	.912	1,027.3	0.218	.587	0.024
1984	–314.1	–0.042	.882	1,279.8	0.447	.114	0.159
1985	–166.8	–0.029	.919	450.6	0.241	.398	0.039
1986	–3468.4	–0.255	.551	2,950.4	0.445	.298	0.032
1987	–463.3	–0.034	.938	1,635.2	0.245	.572	0.038
1988	–2989.3	–0.243	.538	2,995.4	0.514	.193	0.070
1989	–8284.5	–0.727	.003	5,738.0	0.950	.000	0.082
1990	–32881.6	–1.848	.000	20,655.9	2.286	.000	0.339
1991	10,501.1	0.667	.085	–4146.6	0.552	.154	0.015
1992	15,944.5	0.807	.188	–6582.9	–0.62	.312	0.033
1993	–2196	–0.098	.883	3,339.4	0.274	.679	0.025
1994	–6796.6	–0.246	.662	5,947.8	0.722	.471	0.020
1995	8,642.9	0.744	.013	–2707.1	–0.722	.015	0.015
1996	5,048.1	0.200	.733	–146.6	–0.011	.985	0.029
1997	–6949.6	–0.243	.728	6,517.8	0.424	.545	0.027

where

B_t = current year's total expenditures per student on public education
Q_{t-1} = previous year's average daily attendance
a = regression coefficient for Q_{t-1}
a_{std} = standardized regression coefficient for Q_{t-1}
$(Q_{t-1})^2$ = previous year's average daily attendance squared
b = regression coefficient for $(Q_{t-1})^2$
b_{std} = standardized regression coefficient for $(Q_{t-1})^2$
t = time in academic school years
$AdjR^2$ = amount of variation in B_t explained by variations in Q_{t-1} and $(Q_{t-1})^2$

SOURCE: Calculated from Indiana Department of Education data, 1981–97.

period examined, average daily attendance is a weak predictor of total expenditures per student, but the strongest predictive models represent functions that are concave-up.

For 16 of the 17 years covering 1981–97, less than 2 percent of the variation in total expenditures per student can be explained by variations in graduation rates (table 2). In 1994, though, independent variables in the original budget-output function explained 15.4 percent of the variation in total expenditures per student. The standardized coefficient for the linear term of the graduation rate was positive for 7 of the 17 years, while this term was a statistically significant predictor for 2 years. The standardized coefficient for the quadratic term of the graduation rate was negative for 8 of the 17 years, while this term was a statistically significant predictor for 2 years. Therefore, over the time period examined, graduation rates are a weak predictor of total expenditures per pupil, but the strongest predictive model represents a function that is concave-up.

Table 2. Original form budget-output regressions for total expenditures per student and graduation rate: 1981–97

Year	Coefficients for Q_{t-1}			Coefficients for $(Q_{t-1})^2$			Power
B_t	a	a_{std}	p-value	b	b_{std}	p-value	$AdjR^2$
1981	–1055.6	–0.292	.140	875.4	0.407	.040	0.017
1982	630.6	0.192	.359	–409.3	–0.219	.297	–0.003
1983	698.4	0.197	.338	–302.3	–0.146	.478	–0.002
1984	419.7	0.112	.594	–57.9	–0.028	.895	0.000
1985	951.6	0.291	.168	–380.1	–0.231	.274	0.002
1986	260.5	0.071	.741	–5.4	–0.003	.988	–0.002
1987	1,604.3	0.445	.048	–738.5	–0.424	.059	0.007
1988	569.7	0.152	.501	–324.7	–0.173	.443	–0.005
1989	–127.3	–0.02	.917	195.2	0.039	.838	–0.007
1990	–942.6	–0.13	.303	663.4	0.212	.093	0.006
1991	–1188.2	–0.171	.354	744.5	0.141	.443	–0.004
1992	–70.1	–0.01	.956	–60	–0.011	.953	–0.007
1993	–695.9	–0.091	.582	444.2	0.084	.611	–0.006
1994	–2817.8	–0.344	.005	2,620.4	0.678	.000	0.154
1995	–1446.5	–0.169	.298	1,056.2	0.182	.264	–0.003
1996	–1318.2	–0.149	.414	1,125.4	0.172	.346	–0.004
1997	–392.4	–0.038	.825	32.9	0.004	.980	–0.006

where

B_t = current year's total expenditures per student on public education
Q_{t-1} = previous year's average daily attendance
a = regression coefficient for Q_{t-1}
a_{std} = standardized regression coefficient for Q_{t-1}
$(Q_{t-1})^2$ = previous year's average daily attendance squared
b = regression coefficient for $(Q_{t-1})^2$
b_{std} = standardized regression coefficient for $(Q_{t-1})^2$
t = time in academic school years
$AdjR^2$ = amount of variation in B_t explained by variations in Q_{t-1} and $(Q_{t-1})^2$

SOURCE: Calculated from Indiana Department of Education data, 1981–97.

For 9 of the 10 years covering 1988–97, less than 5 percent of the variation in total expenditures per student can be explained by variations in remediation rates (table 3). In 1990, independent variables in the original budget-output function explained 6.3 percent of the variation in total expenditures per student. The standardized coefficient for the linear term of the remediation rate was positive for 6 of the 10 years, while this term was a statistically significant predictor for only 1 year. The standardized coefficient for the quadratic term of the remediation rate was negative for 4 of the 10 years, while this term was a statistically significant predictor for only 1 year. Therefore, over the time period examined, remediation rates are a weak predictor of total expenditures per pupil, but the strongest predictive model represents a budget-output function that is concave-up.

For 6 of the 8 years covering 1990–97, less than 1 percent of the variation in total expenditures per student can be explained by variations in school quality (table 4). In 1990 and 1994, though, independent variables in the original budget-output function explained 38.6 and 15.2 percent, respectively, of the

Table 3. Original form budget-output regressions for total expenditures per student and remediation rate: 1981–97

Year	Coefficients for Q_{t-1}			Coefficients for $(Q_{t-1})^2$			Power
B_t	a	a_{std}	p-value	b	b_{std}	p-value	$AdjR^2$
1981	—	—	—	—	—	—	—
1982	—	—	—	—	—	—	—
1983	—	—	—	—	—	—	—
1984	—	—	—	—	—	—	—
1985	—	—	—	—	—	—	—
1986	—	—	—	—	—	—	—
1987	—	—	—	—	—	—	—
1988	12,645.6	0.115	.483	516,370.9	0.131	.423	0.050
1989	35,046.8	0.305	.068	–487658.2	–0.126	.452	0.033
1990	–2872.3	–0.229	.176	16,619.2	0.467	.006	0.063
1991	–828	–0.084	.601	8,062.0	0.284	.079	0.036
1992	117.7	0.012	.941	1,964.7	0.071	.669	–0.001
1993	–48.5	–0.008	.974	2,993.1	0.124	.412	0.007
1994	2,248.7	0.181	.234	–1556.4	–0.047	.754	0.012
1995	209.3	0.187	.184	–4249.1	–0.155	.273	–0.001
1996	2,668.5	0.303	.003	–2421.2	–0.211	.04	0.024
1997	–543.7	–0.047	.756	8,035.3	0.258	.088	0.040

where

B_t = current year's total expenditures per student on public education
Q_{t-1} = previous year's average daily attendance
a = regression coefficient for Q_{t-1}
a_{std} = standardized regression coefficient for Q_{t-1}
$(Q_{t-1})^2$ = previous year's average daily attendance squared
b = regression coefficient for $(Q_{t-1})^2$
b_{std} = standardized regression coefficient for $(Q_{t-1})^2$
t = time in academic school years
$AdjR^2$ = amount of variation in B_t explained by variations in Q_{t-1} and $(Q_{t-1})^2$

— Not available.

SOURCE: Calculated from Indiana Department of Education data, 1981–97.

variation in total expenditures per student. The standardized coefficient for the linear term of school quality was positive for 2 of the 8 years, while this term was a statistically significant predictor for 2 years. The standardized coefficient for the quadratic term of school quality was negative for 1 of the 8 years, while this term was a statistically significant predictor for 2 years. Therefore, over the time period examined, school quality is a weak predictor of total expenditures per pupil, but the strongest predictive models represent budget-output functions that are concave-up.

Budget-Output Regressions With Demographic Variables

Niskanen's original budget-output function also needs to be examined while taking into account the influences of various demographic characteristics that affect education. The addition of variables representing these characteristics is necessary to control for different types of community and family characteristics that affect education. Without such controls, comparisons between districts with different types of external in-

Table 4. Original form budget-output regressions for total expenditures per student and school quality: 1981–97

Year	Coefficients for Q_{t-1}			Coefficients for $(Q_{t-1})^2$			Power
B_t	a	a_{std}	p-value	b	b_{std}	p-value	$AdjR^2$
1981	—	—	—	—	—	—	—
1982	—	—	—	—	—	—	—
1983	—	—	—	—	—	—	—
1984	—	—	—	—	—	—	—
1985	—	—	—	—	—	—	—
1986	—	—	—	—	—	—	—
1987	—	—	—	—	—	—	—
1988	—	—	—	—	—	—	—
1989	—	—	—	—	—	—	—
1990	–10984	–0.237	.031	102,647.2	0.832	.000	0.386
1991	–2428.5	–0.067	.649	5,379.8	0.057	.698	–0.006
1992	–5479	–0.091	.525	37,599.8	0.093	.516	–0.005
1993	–13510.3	–0.198	.209	86,359.2	0.171	.278	–0.001
1994	–50091.3	–0.604	.000	438,397.0	0.849	.000	0.152
1995	–12313.2	–0.176	.243	52,470.1	0.072	.632	0.006
1996	1.3	0.008	.964	–0.2	–0.025	.888	–0.007
1997	5.0	0.024	.891	0.3	0.040	.820	–0.003

where

B_t = current year's total expenditures per student on public education
Q_{t-1} = previous year's average daily attendance
a = regression coefficient for Q_{t-1}
a_{std} = standardized regression coefficient for Q_{t-1}
$(Q_{t-1})^2$ = previous year's average daily attendance squared
b = regression coefficient for $(Q_{t-1})^2$
b_{std} = standardized regression coefficient for $(Q_{t-1})^2$
t = time in academic school years
$AdjR^2$ = amount of variation in B_t explained by variations in Q_{t-1} and $(Q_{t-1})^2$

— Not available.

SOURCE: Calculated from Indiana Department of Education data, 1981–97.

fluences would be unfair. For each of the 17 years covering 1981–97, more than 20 percent of the variation in total expenditures per student can be explained by variations in average daily attendance and demographic characteristics (table 5). In 1984, 1985, and 1990, independent variables in the modified budget-output function explained 42.5, 38.7, and 46.7 percent, respectively, of the variation in total expenditures per student. The standardized coefficient for the linear term of average daily attendance was positive for 4 of the 17 years, while this term was a statistically significant predictor for 5 years. The standardized coefficient for the quadratic term of average daily attendance was negative for 4 of the 17 years, while this term was a statistically significant predictor for 8 years. Among the demographic variables examined, family status, family type, and student/teacher ratio were statistically significant predictors of total expenditures every year examined while teacher experience was a statistically significant predictor for 7 years. Therefore, when controlling for variations in demographic characteristics over the time period examined, average daily attendance is a moderately strong predictor of total expenditures per pupil, but the strongest predictive models represent functions that are concave-up.

For each of the 17 years covering 1981–97, more than 18 percent of the variation in total expenditures per

Table 5. Original form budget-output regressions for total expenditures per student with average daily attendance and demographic variables: 1981–97

Year	Standardized coefficients for Q_{t-1}, $(Q_{t-1})^2$, and demographic variables						Power
B_t	a_{std}	b_{std}	Family status	Family type	Student/teacher ratio	Teacher experience	$AdjR^2$
1981	*0.694	*–0.755	*–0.139	—	*–0.400	*0.178	0.205
1982	–0.166	0.310	*–0.236	—	*–0.453	*0.254	0.335
1983	–0.611	*0.680	*–0.241	—	*–0.498	*0.222	0.335
1984	–0.455	*0.785	*–0.264	—	*–0.401	*0.248	0.425
1985	*–0.528	*0.666	*–0.269	—	*–0.465	*0.318	0.387
1986	–0.667	*0.781	*–0.238	—	*–0.347	*0.152	0.218
1987	–0.138	0.256	*–0.221	*–0.235	*–0.357	0.040	0.278
1988	–0.325	0.523	*–0.297	*–0.265	*–0.290	0.030	0.316
1989	*–0.570	*0.736	*–0.251	*–0.293	*–0.271	–0.071	0.283
1990	*–1.797	*2.179	*–0.175	*–0.245	*–0.229	*–0.107	0.467
1991	0.473	–0.451	*–0.229	*–0.292	*–0.371	–0.002	0.281
1992	0.373	–0.303	*–0.176	*–0.282	*–0.396	0.036	0.299
1993	–0.419	0.512	*–0.201	*–0.290	*–0.369	–0.04	0.287
1994	–0.507	0.598	*–0.247	*–0.325	*–0.270	–0.078	0.252
1995	*0.653	*–0.666	*–0.194	*–0.354	*–0.170	–0.089	0.205
1996	–0.076	0.200	*–0.246	*–0.361	*–0.270	–0.09	0.296
1997	–0.094	0.171	*–0.263	*–0.413	*–0.296	–0.1	0.352

where

B_t	= current year's total expenditures per student on public education
Q_{t-1}	= previous year's average daily attendance
$(Q_{t-1})^2$	= previous year's average daily attendance squared
a_{std}	= standardized regression coefficient for Q_{t-1}
b_{std}	= standardized regression coefficient for $(Q_{t-1})^2$
Family status	= measures influence of income and education level
Family type	= measures influence of single parent household and race
Student/teacher ratio	= proportion of students to teachers
Teacher experience	= years of teaching
t	= time in academic school years
$AdjR^2$	= amount of variation in B_t explained by variations in Q_{t-1} and $(Q_{t-1})^2$

* Denotes statistical significance at the .05 level or better.

— Not available.

SOURCE: Calculated from Indiana Department of Education data, 1981–97.

student can be explained by variations in graduation rates and demographic characteristics (table 6). In 1981 and 1994, independent variables in the modified budget-output function explained 20.9 and 37.6 percent, respectively, of the variation in total expenditures per student. The standardized coefficient for the linear term of the graduation rate was positive for 8 of the 17 years, while this term was a statistically significant predictor for 2 years. The standardized coefficient for the quadratic term of the graduation rate was negative for 9 of the 17 years, while this term was a statistically significant predictor for 2 years. Among the demographic variables examined, family status, family type, and student/teacher ratio were statistically significant predictors of total expenditures every year examined, while teacher experience was a statistically significant predictor for 8 years. Therefore, when controlling for variations in demographic characteristics over the time period examined, the graduation rate is a weak predictor of total expenditures per pupil, but the strongest predictive models represent budget-output functions that are concave-up.

For each of the 10 years covering 1988–97, more than 18 percent of the variation in total expenditures per student can be explained by variations in remediation rates and demographic characteristics (table 7). In 1990, independent variables in the modified bud-

Table 6. Budget-output function regressions for total expenditures per student with graduation rates and demographic variables: 1981–97

Year	Standardized coefficients for Q_{t-1}, $(Q_{t-1})^2$, and demographic variables						Power
B_t	a_{std}	b_{std}	Family status	Family type	Student/teacher ratio	Teacher experience	$AdjR^2$
1981	*–0.388	*0.475	*–0.135	—	*–0.389	*0.207	0.209
1982	0.082	–0.139	*–0.227	—	*–0.489	*0.254	0.318
1983	0.165	–0.192	*–0.228	—	*–0.507	*0.224	0.325
1984	0.112	–0.134	*–0.237	—	*–0.455	*0.260	0.306
1985	0.206	–0.27	*–0.269	—	*–0.488	*0.302	0.362
1986	0.018	–0.011	*–0.226	—	*–0.360	*0.159	0.196
1987	0.335	–0.356	*–0.203	*–0.236	*–0.387	0.047	0.273
1988	0.009	–0.053	*–0.270	*–0.295	*–0.333	0.029	0.277
1989	–0.069	0.094	*–0.237	*–0.336	*–0.306	–0.099	0.235
1990	–0.039	0.053	*–0.141	*–0.331	*–0.279	*–0.152	0.186
1991	–0.219	0.229	*–0.236	*–0.290	*–0.381	0.006	0.281
1992	0.005	–0.035	*–0.173	*–0.282	*–0.419	0.040	0.295
1993	–0.019	–0.003	*–0.192	*–0.293	*–0.389	–0.038	0.278
1994	*–0.356	*0.648	*–0.235	*–0.308	*–0.273	–0.068	0.376
1995	–0.131	0.119	*–0.195	*–0.347	*–0.191	–0.067	0.190
1996	–0.08	0.083	*–0.251	*–0.361	*–0.297	–0.069	0.282
1997	–0.046	0.013	*–0.264	*–0.417	*–0.313	*–0.111	0.347

where

B_t = current year's total expenditures per student on public education
Q_{t-1} = previous year's average daily attendance
$(Q_{t-1})^2$ = previous year's average daily attendance squared
a_{std} = standardized regression coefficient for Q_{t-1}
b_{std} = standardized regression coefficient for $(Q_{t-1})^2$
Family status = measures influence of income and education level
Family type = measures influence of single parent household and race
Student/teacher ratio = proportion of students to teachers
Teacher experience = years of teaching
t = time in academic school years
$AdjR^2$ = amount of variation in B_t explained by variations in Q_{t-1} and $(Q_{t-1})^2$

* Denotes statistical significance at the .05 level or better.

— Not available.

SOURCE: Calculated from Indiana Department of Education data, 1981–97.

get-output function explained 23.2 percent of the variation in total expenditures per student. The standardized coefficient for the linear term of the remediation rate was positive for 5 of the 10 years, while never being a statistically significant predictor. The standardized coefficient for the quadratic term of the remediation rate was negative for 3 of the 10 years, while this term was a statistically significant predictor for only 1 year. Among the demographic variables examined, family status, family type, and student/teacher ratio were statistically significant predictors of total expenditures every year examined, while teacher experience was a statistically significant predictor for 2 years. Therefore, when controlling for variations in demographic characteristics over the time period examined, the remediation rate is a weak predictor of total expenditures per pupil, but the strongest predictive model represents a budget-output function that is concave-up.

For each of the 8 years covering 1990–97, more than 19 percent of the variation in total expenditures per student can be explained by variations in school quality and demographic characteristics (table 8). In 1990 and 1994, independent variables in the modified budget-output function explained 52.7 and 37.3 percent, respectively, of the variation in total expenditures per student. The standardized coeffi-

Table 7. Budget-output function regressions for total expenditures per student with remediation rates and demographic variables: 1981–97

Year	Standardized coefficients for Q_{t-1}, $(Q_{t-1})^2$, and demographic variables						Power
B_t	a_{std}	b_{std}	Family status	Family type	Student/teacher ratio	Teacher experience	$AdjR^2$
1981	—	—	—	—	—	—	—
1982	—	—	—	—	—	—	—
1983	—	—	—	—	—	—	—
1984	—	—	—	—	—	—	—
1985	—	—	—	—	—	—	—
1986	—	—	—	—	—	—	—
1987	—	—	—	—	—	—	—
1988	–0.107	0.228	*–0.270	*–0.280	*–0.325	0.006	0.290
1989	0.201	–0.116	*–0.225	*–0.303	*–0.313	–0.101	0.241
1990	–0.216	*0.407	*–0.132	*–0.308	*–0.275	*–0.167	0.232
1991	–0.127	0.236	*–0.224	*–0.270	*–0.375	0.003	0.291
1992	0.033	0.058	*–0.163	*–0.274	*–0.428	0.044	0.301
1993	–0.042	0.181	*–0.178	*–0.290	*–0.403	–0.034	0.296
1994	0.070	0.049	*–0.219	*–0.319	*–0.306	–0.07	0.254
1995	0.085	–0.077	*–0.188	*–0.342	*–0.200	–0.071	0.189
1996	0.153	–0.112	*–0.231	*–0.344	*–0.307	–0.062	0.288
1997	–0.033	0.163	*–0.239	*–0.405	*–0.313	*–0.099	0.363

where

B_t = current year's total expenditures per student on public education
Q_{t-1} = previous year's average daily attendance
$(Q_{t-1})^2$ = previous year's average daily attendance squared
a_{std} = standardized regression coefficient for Q_{t-1}
b_{std} = standardized regression coefficient for $(Q_{t-1})^2$
Family status = measures influence of income and education level
Family type = measures influence of single parent household and race
Student/teacher ratio = proportion of students to teachers
Teacher experience = years of teaching
t = time in academic school years
$AdjR^2$ = amount of variation in B_t explained by variations in Q_{t-1} and $(Q_{t-1})^2$

* Denotes statistical significance at the .05 level or better.

— Not available.

SOURCE: Calculated from Indiana Department of Education data, 1981–97.

cient for the linear term of school quality was positive for 1 of the 8 years, while this term was a statistically significant predictor for 2 years. The standardized coefficient for the quadratic term of school quality was negative for 2 of the 8 years, while being a statistically significant predictor for 2 years. Among the demographic variables examined, family status, family type, and student/teacher ratio were statistically significant predictors of total expenditures every year examined, while teacher experience was a statistically significant predictor for 2 years. Therefore, when controlling for variations in demographic characteristics over the time period examined, school quality is a weak predictor of total expenditures per pupil, but the strongest predictive models represent budget-output functions that are concave-up.

Niskanen and Modified Quadriform Analyses

Figure 2 shows total expenditures per student within quadriforms of perennially categorized school corporations for average daily attendance, graduation rates, remediation rates, and school quality. Just under 13 percent of school corporations in Indiana are inefficient producers of average daily attendance compared to 27 percent that are efficient producers. Twelve percent of Indiana school corporations are inefficient producers of graduation rates com-

Table 8. Budget-output function regressions for total expenditures per student with school quality and demographic variables: 1981–97

Year	Standardized coefficients for Q_{t-1}, $(Q_{t-1})^2$, and demographic variables						Power
B_t	a_{std}	b_{std}	Family status	Family type	Student/teacher ratio	Teacher experience	*AdjR*²
1981	—	—	—	—	—	—	—
1982	—	—	—	—	—	—	—
1983	—	—	—	—	—	—	—
1984	—	—	—	—	—	—	—
1985	—	—	—	—	—	—	—
1986	—	—	—	—	—	—	—
1987	—	—	—	—	—	—	—
1988	—	—	—	—	—	—	—
1989	—	—	—	—	—	—	—
1990	*–0.245	*0.798	*–0.142	*–0.247	*–0.268	*–0.097	0.527
1991	–0.039	–0.021	*–0.230	*–0.290	*–0.387	0.004	0.279
1992	–0.001	–0.026	*–0.177	*–0.280	*–0.421	0.039	0.295
1993	–0.08	0.033	*–0.204	*–0.280	*–0.382	–0.035	0.280
1994	*–0.531	*0.762	*–0.218	*–0.331	*–0.254	–0.077	0.373
1995	–0.075	0.014	*–0.185	*–0.343	*–0.193	–0.074	0.192
1996	–0.02	0.000	*–0.245	*–0.362	*–0.301	–0.072	0.282
1997	0.058	0.019	*–0.276	*–0.410	*–0.313	*–0.120	0.352

where

B_t = current year's total expenditures per student on public education
Q_{t-1} = previous year's average daily attendance
$(Q_{t-1})^2$ = previous year's average daily attendance squared
a_{std} = standardized regression coefficient for Q_{t-1}
b_{std} = standardized regression coefficient for $(Q_{t-1})^2$
Family status = measures influence of income and education level
Family type = measures influence of single parent household and race
Student/teacher ratio = proportion of students to teachers
Teacher experience = years of teaching
t = time in academic school years
*AdjR*² = amount of variation in B_t explained by variations in Q_{t-1} and $(Q_{t-1})^2$

* Denotes statistical significance at the .05 level or better.

— Not available.

SOURCE: Calculated from Indiana Department of Education data, 1981–97.

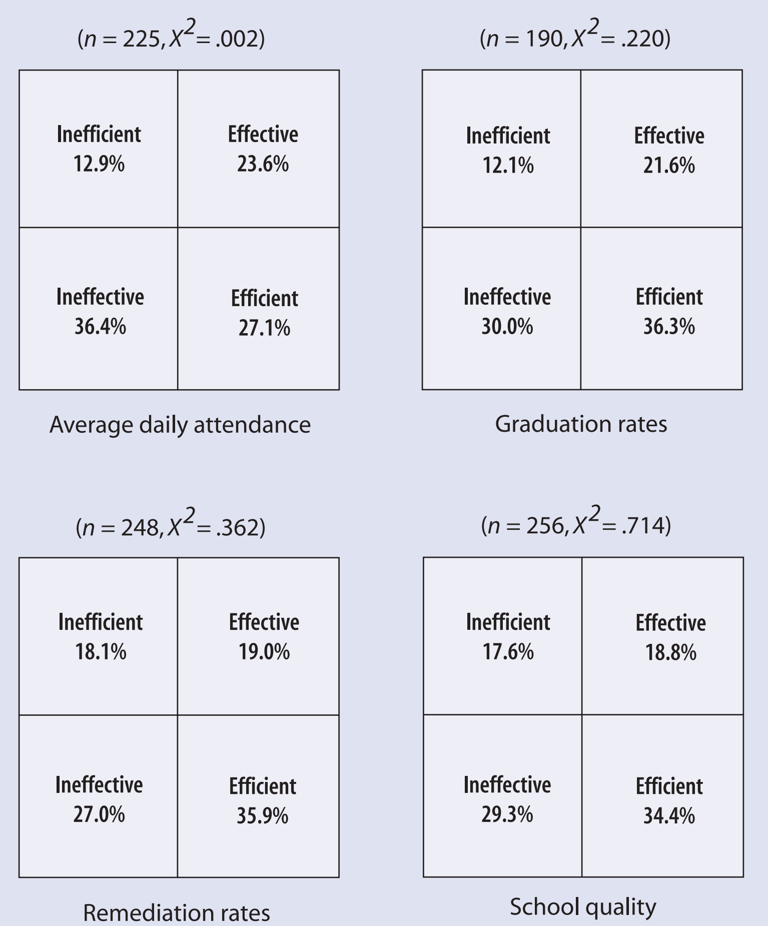
pared to 36 percent that are efficient producers. Eighteen percent of Indiana school corporations are inefficient producers of low remediation rates compared to 36 percent that are efficient producers. Over 17 percent of Indiana school corporations are inefficient producers of school quality compared to 34 percent that are efficient producers.

Statistically significant differences in expected values versus actual values within the "cell" categories were found only in the quadriform measuring average daily attendance. Therefore, in contrast to Niskanen's analytical conclusion that the entire system of Indiana school corporations is an inefficient producer of educational outcomes, modified quadriform analyses of perennially categorized corporations show that approximately 15 percent of the state's school corporations produce educational outcomes in a manner that is economically inefficient. Additionally, more than 30 percent of Indiana public school corporations were found to be economically efficient producers of educational outcomes. Finally, it is necessary to note one unexpected observation: *On average, more Indiana school corporations produce educational outcomes in a manner that is economically ineffective than at levels that are economically inefficient.*

Conclusion

In contrast to Niskanen's hypothetical assertions, expenditure-outcome relationships for Indiana public school corporations were found to be represented best by budget-output relationships that are *concave-up*. Further, analytical evidence in this research suggests that some statistically significant relationships *do exist* between the current year's total expenditures per stu-

Figure 2. Modified quadriform analysis for total expenditures per student across years 1981–97



SOURCE: Calculated from Indiana Department of Education data, 1981–97.

dent and previous year's organizational outcomes. Average and marginal total expenditures per student for the production of education were not found to be equal. On the other hand, modified quadriform analyses showed that the Indiana public school system has twice as many economically efficient school corporations as inefficient ones. Therefore, Indiana public school corporations cannot be designated as budget-maximizing agencies in the context defined by Niskanen's theory.

Budget-Output Functions

Niskanen's theory of budget-maximizing bureaus hypothesizes that expenditure data for Indiana public school corporations should be represented best by quadratic budget-output functions that are concave-down. Additionally, he claims no statistically significant relationships should exist between the current year's total expenditure per student and previous year's organizational outcomes. As a result of these two occurrences, the system of Indiana public schools will be inefficient producers of educational outcomes. After examining Niskanen's original hypothesis empirically, evidence suggests that expenditure data for Indiana public school corporations are represented best by budget-output functions that are concave-up. Moreover, statistically significant relationships do exist between current year total expenditures per pupil and previous year educational outcomes—notably when considering average daily attendance.

Indiana public school corporations cannot be designated as budget-maximizing agencies in the context defined by Niskanen's theory.

After controlling for various demographic characteristics, stronger evidence exists that expenditure data for Indiana public school corporations are represented best by budget-output functions that are concave-up. Moreover, some statistically significant—and stronger—relationships do exist between current year total expenditures per pupil and previous year educational outcomes when controlling for demographic characteristics—especially when considering average daily attendance. It is important to note that every year demographic variables were included in the predictive models that family status, family type, and student/teacher ratio had negative—but statistically significant—relationships with total expenditures per student. In other words, school corporations with large percentages of low-income families, individuals without high school diplomas living in the community, minority students, single-parent households, and high student-teacher ratios tended to receive less money per student—on average—than school corporations without these characteristics. Still, despite the constant statistical significance of the demographic variables, educational outcome measures maintained stronger relative predictive strength among all variables in the budget-output models.

Economic Efficiency

The final portion of Niskanen's theory charges that average and marginal total expenditures per student for the production of educational outcomes will not be equal; therefore, Indiana public school corporations will produce educational outcomes inefficiently. After examining this hypothesis empirically, the average total expenditures per student for Indiana public school corporations were found to be unequal to their marginal costs. This lack of equality existed across all years examined regardless of which of the four outcome variables—average daily attendance, graduation rates, remediation rates, and school quality—were employed in the analysis. Further, over the years examined, remediation rates and school quality had mean differences that were statistically significant more than 50 percent of the time. Therefore, according to Niskanen's criteria, Indiana public school corporations produce educational outcomes in a manner that is economically inefficient.

Unlike the average-marginal cost analyses, the modified quadriform examines expenditure and output variations among individual corporations. When examining total expenditures per student across all educational outcome categories (i.e., total expenditures per student within quadriforms of perennially categorized school corporations for average daily attendance, graduation rates, remediation rates, and school quality), modified quadriform analyses of perennially categorized corporations show that approximately 15 percent of the state's school corporations produce educational outcomes in a manner that is economically inefficient. On the other hand, more than 30 percent of Indiana public school corporations were found to be economically efficient producers of educational outcomes. It is important to note one unexpected observation: On average, more Indiana school corporations produce educational outcomes in a manner that is economically ineffective (30.7 percent) than at levels that are economically inefficient.

Implications for Indiana Public School Corporations

Even though the conclusions based in this research are well-grounded, there still is a need to replicate the statistical and mathematical modeling presented using more comprehensive data. This reanalysis should be conducted both for total expenditures per pupil as well as other expenditure categories (e.g., instructional expenditures per student). The analyses should focus specifically on two aspects of the theory of budget maximization: (1) Determining stronger predictive relationships between expenditures as outputs and educational outcomes as inputs; and (2) Determining accurate mathematical models of sponsor preferences. With further research into these two types of models, a determination of whether bureaus spend in accordance to sponsor preferences will be easier to discern.

Given the recent changes in the state education funding formula that were designed to improve the availability of revenue for low-wealth school corporations, it is surprising that the effects of demographic variables remain so prevalent.

Similar to discovering stronger predictive relationships between expenditures and educational outcomes, research conducted to improve the use of quadriform analyses also is warranted. Though the method is sound analytically, the use of regression analyses to determine the axes of the quadriform may lead to "questionable" placement of school corporations. It is important to remember that the axes of the quadriform are being developed while attempting to control for demographic characteristics that exist in school corporations. It may be better to develop only the expenditure axes based on demographic characteristics while basing the educational outcome axes on *desired* outcomes as opposed to actual averaged outcomes. This small change will show which school corporations are efficient at producing educational outcomes relative to others in Indiana. Moreover, this change also will show which school corporations are efficient at achieving the educational goals desired by parents, teachers, and policymakers in the state.

With demographic characteristics being statistically significant—and having a consistently negative influence on expenditures—in most of the analyses conducted in this research, a more thorough examination of their influences on education expenditures in the state of Indiana is warranted. Given the recent changes in the state education funding formula that were designed to improve the availability of revenue for low-wealth school corporations, it is surprising that the effects of demographic variables remain so prevalent. It will be important to discover what types of school districts—suburban or rural corporations versus large or small corporations—are affected by these demographic characteristics and what causes the consistently negative relationship with expenditures per student.

Finally, the most surprising conclusion from the analysis in the research is the large percentage of schools classified as ineffective in producing educational outcomes. On average, this grouping of districts was second only to efficient school corporations. Initially, a descriptive analysis of this group of school corporations would provide insight to what factors lead to their perennially low expenditure levels, educational outputs, and potential underservicing of their student populations. Further, if these schools are underfunded due primarily to changes in the state's school funding formula, a re-examination of the formula is necessary as questions about the type of equity goals desired in Indiana resurface.

Implications for Budget Maximization Theory

Niskanen claimed that sponsor preferences should be represented best by budget-output equations that are concave-down, quadratic functions. Niskanen's first and second derivative criteria assume that sponsors are willing to pay for services up to a certain expenditure level. Evidence from Indiana public school corporations shows that statistically significant relationships between expenditures and outputs are represented best by quadratic functions that are concave-up. Concave-up functions represent economic behaviors to the converse of Niskanen's derivative assumptions: A reluctance to pay for services until a given impetus is received (small budgetary increases over time). This empirical evidence is consistent with the conservative nature of the Indiana legislature's fiscal policy. Therefore, a change to Niskanen's theory may be first to examine the fiscal history of a particular sponsor's expenditure preferences before determining the concavity of their budget-output functions.

Further, if the concavity assumptions of Niskanen's theory are dependent on particular sponsor preferences at specific times, it is doubtful that bureaus attempt to pursue the spending preferences of their sponsors. Given the nature of state politics, individuals managing bureaus and sponsors—in conflict over differences with respect to values, preferences, beliefs, perceptions of reality, and access to information—struggle for power and the capacity to distribute scarce resources. As a result of this conflict, the ability to bargain, negotiate, and compromise becomes the most important asset utilized by actors in the system. The resulting web of compromises generates a confusing multiplicity of objectives—many in opposition to one another—that emerge as organizational and political goals. Therefore, it seems more likely that bureaus attempt to spend money on programs that will achieve organizational goals while also spending money on programs that appease sponsors. In short, bureaus may not pursue sponsor preferences specifically due to their complex nature and a desire to balance technical and allocative efficiency objectives.

Given the difficulty of defining sponsor preferences, it will be equally difficult to define these complex expenditure-outcome relationships mathematically. For Indiana school corporations, strongly predictive budget-output relationships were found generally around specific years: 1984, 1990, and 1994. Not coincidentally, these also are years when the Indiana legislature made changes in the state's education finance formula. At these particular times, it seems appropriate that strong mathematical relationships should exist between expenditures and educational outcomes. At other times, when various and multiple sponsor preferences are being pursued, it becomes more difficult to find specific mathematical relationships. As such, it seems inappropriate for a bureau's level of efficiency to be measured as a pursuit of what could be the unattainable: an accurate and logical mathematical representation of a state's legislative process. The difficulty in developing this mathematical relationship may be a primary reason that a primary educational production function (or production functions) is yet to be found.

Finally, the attempted coupling of a bureau's expenditure patterns to sponsor preferences also raises questions about the appropriateness of using traditional average-marginal costs analyses as the primary determinant of efficiency within an organization. Average–marginal expenditure analyses usually are reserved for production activities that are well defined. Government spending for public services—such as education—is not one of these activities. Therefore, the use of a method like the modified quadriform analyses seems more appropriate. Here, acknowledging that a primary production function does not exist, efficiency is based on those organizations that are efficient producers of outcomes relative to those that are not. At this point, instead of pursuing a specific economic or mathematical relationship—that very likely will not be found—a series of "best practices" can be developed by examining what the efficient bureaus are doing and what the inefficient bureaus are not.

Appendix: Theory of Budget-Maximizing Bureaucratic Behavior

Most of traditional economics deals with the behavior of profit-seeking firms, owners of production factors, and consumers. Since most economic activity is organized through profit-seeking firms by the voluntary exchange of production factors for capital, and of capital for consumer goods, this methodology is appropriate for market exchanges. Economists also have developed an elaborate structure of widely accepted propositions about what public goods and services ought to be supplied. However, even the theory of public goods rests on the assumption that public agencies—even though financed by government—will behave similarly to those in a competitive industry. Public choice theory—the field of economics that encompasses budget-maximization theory—offers an alternative framework to traditional economic analyses (Buchanan and Tollison 1984; Downs 1998; Peacock 1992). Using this alternative framework challenges not only cost minimization assumptions, but also allows for a discussion on the structure of educational bureaucracies as well as ideas of efficiency and accountability in education.

In *Bureaucracy and Representative Government* (1971), Niskanen developed a theory of supply for public bureaus that is

> based on a model of purposive behavior by the manager of a single bureau. . . . not to explain the actions of individuals but to generate hypotheses concerning aggregate consequences of the interaction among individuals (p. 5).

This construction is similar to the theory of supply that is based on the model of the profit-seeking firm. In this instance, though, the bureaucrat is the central figure—the "chooser" or the "maximizer"—and is assumed to

- face a set of possible actions;
- have personal preferences among the outcomes of these actions; and
- choose the action within the possible set that is preferred.

The larger political and organizational environment also is believed to influence the behavior of the bureaucrat by constraining sets of possible actions; changing relationships between actions and outcomes; and influencing preferences (Niskanen 1971, p. 7).

The central motivational assumption for a business manager is that when profits of the firm are maximized, personal utility can be maximized in a variety of pecuniary and nonpecuniary ways (Niskanen 1971, pp. 36–37). Similarly, Niskanen believes bureaucrats also need to be recognized as individuals who maximize personal utility and not as those devoted solely to promoting the general welfare or the interests of the state. Consequently, several variables may enter a bureaucrat's personal utility function: salary, perquisites of the office, public reputation, power, patronage, output of the bureau, ease of making changes, and ease of managing the bureau. Moreover, he claims that during a bureaucrat's tenure in office, all of these variables—except the last two—are a positive monotonic function of the bureau's total budget. Niskanen believes it is not necessary that a bureaucrat's utility be strongly dependent on each variable increasing in conjunction with the budget, but only that increases are associated positively with the level of the budget. Consequently, budget maximization becomes an adequate proxy even for those bureaucrats with relatively low pecuniary motivations and relatively high motivations for attending to the public interest (Niskanen 1971, pp. 38–39).

Ultimately, Niskanen gives the most complete definition of the central motivational assumption for budget-maximization theory:

> Subject to the constraint that the budget must be greater than or equal to the minimum total costs of supplying the output expected by the bureau's sponsor, bureaucrats maximize their total budget of their bureau during their tenure.

He adds the "budget constraint" maxim because any bureau, during any budget period, may supply more or less than the expected level of output. However, over time, every bureau will be constrained to supply the output expected by the sponsor. A bureau that consistently promises more than it can deliver is penalized by the discounting of future promises and the receipt of lower budgets. Conversely, a bureau that performs better than expected is likely to be rewarded by higher future budgets (Niskanen 1971, p. 42).

Budget-Output Function

From the vantage point of the bureau, Niskanen asserts that the preferences of the sponsor can be summarized mathematically by a *budget-output function* (Niskanen 1971, p. 25). Any point represented by this function represents the maximum total budget a sponsor is willing to grant to the bureau for a specific expected level of output. The function has the following properties:

- Over some range of expected output, the sponsor is willing to grant a higher budget for a higher expected output. (The first derivative is a positive monotonic function over the relevant range.)
- Over some range, the sponsor is willing to grant a higher budget per unit of output for a smaller than expected output than for a larger than expected output. (The second derivative is a negative monotonic function over the relevant range.)

Several types of equations share these two properties, but Niskanen uses a quadratic function of the following form to represent the concave-down budget-output function:

$$B_t = aQ_{t-1} - bQ_{t-1}^2 \text{ subject to } Q_{t-1} : Q_{t-1} \in [0, a/2b]$$

where,

B_t ≡ maximum total budget sponsor is willing to grant to bureau during a specific time period;

Q_{t-1} ≡ expected level of output by bureau during a specific time period;

t ≡ time in academic years; and,

a, b ≡ the coefficients for Q_{t-1} and Q_{t-1}^2, respectively.

He claims a total budget-output function is a necessary building block for a theory of supply by bureaus because the exchange of promised activities and expected output for a budget is conducted "entirely in total" rather than in "unit" terms. The budget-output function, therefore, should be considered to be the product of two relationships: (1) the relationship between budget and level of service; and (2) the relationship between level of service and output (Niskanen 1971, pp. 25–26).

Furthermore, Niskanen states that a bureaucrat usually can estimate the sponsor's budget-output function fairly accurately from previous budget reviews, from recent changes in the composition of the collective organization, and by the levels of influence different constituencies exert on the sponsor. In addition, he believes a bureaucrat also possesses greater knowledge about the cost and production factors for the services provided than members of the sponsor organization do. In contrast, budgets offered by the bureau reveal little about the minimum budget amount that would be sufficient to supply a given output. Therefore, Niskanen claims, a bureaucrat needs relatively little information—most of which can be estimated by the revealed preferences of the sponsor—to exploit the position as a monopoly supplier of a given service. The members of the collective organization, on the other hand, need a great deal of information—little of which can be estimated from revealed behavior—to exploit their position as a monopsony buyer of services. Therefore, the theory Niskanen developed originally assumed that the sponsor is passive and knows the largest budget it is prepared to grant for an expected level of services. These characteristics are assumed because there is no incentive or opportunity for the sponsor to obtain information on the minimum budget necessary to supply this service (Niskanen 1971, pp. 29–30).

Finally, Niskanen's completed theory of budget-maximizing bureaucratic behavior states that subject to a budget constraint greater than or equal to the costs of supplying the output expected by a bureau's sponsors, bureaucrats attempt to maximize an agency's total budget during their tenure. As a result of this budget-maximizing behavior, Niskanen concludes that bureaus generate output that is produced inefficiently. Therefore, if the general theory of budget-maximizing bureaucratic behavior holds, expenditures for a public bureau should be represented by a quadratic budget-output function that is concave-down, with no statistically significant relationships between the current year's total budget and its previous year's organizational outcomes, and with average costs of production that are not equal to marginal costs of production (Niskanen 1971, pp. 49–50).

References

Anderson, D.M. (1996). Stretching the Tax Dollar: Increasing Efficiency in Urban and Rural Schools. In L.O. Picus and J.L. Wattenbarger (Eds.), *Where Does the Money Go?* (pp. 156–177). Thousand Oaks, CA: Corwin Press.

Aucoin, P. (1991). The Politics and Management of Restraint Budgeting. In A. Blais and S. Dion (Eds.), *The Budget-Maximizing Bureaucrat: Appraisals and Evidence* (pp. 119–142). Pittsburgh, PA: University of Pittsburgh Press.

Barnett, W.S. (1994). Obstacles and Opportunities: Some Simple Economics of School Finance Reform. *Educational Policy, 8*(4): 436–452.

Bauer, B.P. (1992). *Final Report: Indiana House of Representatives Select Committee on Primary and Secondary Education.* Indianapolis, IN: National Conference of State Legislatures.

Bennett, W.J. (1992). *The Devaluing of America: The Fight for Our Culture and Children* (pp. 39–70). New York, NY: Summit Books.

Berlin, G., and Sum, A. (1988). *Toward a More Perfect Union: Basic Skills, Poor Families, and Our Economic Future.* Occasional Paper 3 for the Ford Foundation Project on Social Welfare and the American Future. New York, NY.

Blais, A., and Dion, S. (1991). Conclusions: Are Bureaucrats Budget Maximizers? In A. Blais and S. Dion (Eds.), *The Budget-Maximizing Bureaucrat: Appraisals and Evidence* (pp. 355–361). Pittsburgh, PA: University of Pittsburgh Press.

Boyd, W.L., and Hartman, W.T. (1988). The Politics of Educational Productivity. In D.H. Monk and J. Underwood (Eds.), *Micro-Level School Finance: Issues and Implications for Policy* (pp. 271–310). Cambridge, MA: Ballinger Publishing Company.

Buchanan, J.M., and Tollison, R.D. (Eds.). (1984). *The Theory of Public Choice II*. Ann Arbor, MI: University of Michigan Press.

Bush, W.C., and Denzau, A.T. (1977). The Voting Behavior of Bureaucrats and Public Sector Growth. In T.E. Borcherding (Ed.), *Budgets and Bureaucrats: The Sources of Government Growth* (pp. 90–99). Durham, NC: Duke University Press.

Campbell, C., and Naulls, D. (1991). The Limits of the Budget-Maximizing Theory: Some Evidence From Officials' Views of Their Roles and Careers. In A. Blais and S. Dion (Eds.), *The Budget-Maximizing Bureaucrat: Appraisals and Evidence* (pp. 85–118). Pittsburgh, PA: University of Pittsburgh Press.

Downs, A. (1967). *Inside Bureaucracy*. Boston, MA: Little, Brown, and Company.

Downs, A. (1998). *Political Theory and Public Choice*. North Hampton, MA: Edward Elgar Publishing Company.

Friedman, M. (1962). *Capitalism and Freedom*. Chicago, IL: University of Chicago Press.

Genge, F.C. (1991). *A Further Application of the Quadriform to the Study of Economic Efficiency in K–12 Schools in Illinois*. A paper presented at the Annual Meeting of the American Education Finance Association. Williamsburg, VA.

Genge, F.C. (1992). *Assessing Inefficiencies in Illinois School Districts*. A paper presented at the Annual Meeting of the American Education Finance Association. New Orleans, LA.

Hanushek, E.A. (1986). The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature, 24*(3): 1,141–1,177.

Hanushek, E.A. (1995). Improving School Performances While Controlling Costs. In W.J. Fowler, Jr. (Ed.), *Developments in School Finance, 1995* (NCES 96–344) (pp. 111–122). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Hentschke, G.C. (1988). Budgetary Theory and Reality: A Micro View. In D.H. Monk and J. Underwood (Eds.), *Micro-Level School Finance: Issues and Implications for Policy* (pp. 311–336). Cambridge, MA: Ballinger Publishing Company.

Hickrod, G.A., Lui, C.C., Arnold, R., Chaudhari, R., Frank, L.E., Franklin, D.L., Polite, M.M., Pruyne, G.B., and Ward, J.G. (1989). *The Biggest Bang for the Buck: An Initial Report on Technical Economic Efficiency in Illinois K–12 Schools. With a Comment on* Rose v. The Council. Normal, IL: Center for the Study of Educational Finance.

Hickrod, G.A., Genge, F.C., Chaudhari, R., Lui, C.C., Franklin, D.L., Arnold, R., and Frank, L.E. (1990). *The Biggest Bang for the Buck: A Further Investigation of Economic Efficiency in the Public Schools of Illinois.* Normal, IL: Center for the Study of Educational Finance.

Ind. Const. art. 1, § 23.

Ind. Const. art. 8, § 1.

Indiana State Board of Accounts. (1998). *Accounting and Uniform Compliance Guidelines Manual for Indiana Public School Corporations*. Indianapolis, IN: Author.

Johnson, J. (1984). *Econometric Methods* (3rd ed.) (pp. 343–383). New York, NY: McGraw-Hill Publishing Company.

Kiewiet, D.R. (1991). Bureaucrats and Budgetary Outcomes: Quantitative Analyses. In A. Blais and S. Dion (Eds.), *The Budget-Maximizing Bureaucrat: Appraisals and Evidence* (pp. 143–174). Pittsburgh, PA: University of Pittsburgh Press.

Kirst, M.W. (1983). A New School Finance for a New Era of Fiscal Constraint. In A. Odden and L.D. Webb (Eds.), *School Finance and School Improvement Linkages for the 1980s* (pp. 1–15). Cambridge, MA: Ballinger Publishing Company.

Lehnen, R.G., and Johnson, C.E. (1989). *Financing Indiana's Public Schools: Update 1989*. Indianapolis, IN: Indiana University-Purdue University, School of Public and Environmental Affairs.

Levin, H.M., and Kelly, C. (1994). Can Education Do It Alone? *Economics of Education Review, 13*(2): 97–108.

Lynn, L.E. (1991). The Budget-Maximizing Bureaucrat: Is There a Case? In A. Blais and S. Dion (Eds.), *The Budget-Maximizing Bureaucrat: Appraisals and Evidence* (pp. 59–84). Pittsburgh, PA: University of Pittsburgh Press.

Mann, D., and Inman, D. (1984). Improving Education With Existing Resources: The Instructionally Effective Schools' Approach. *Journal of Education Finance*, *10*(2): 259–69.

Mises, L. (1944). *Bureaucracy*. New Haven, CT: Yale University Press.

Niskanen, W.A. (1968). The Peculiar Economics of Bureaucracy. *American Economic Review*, *58*(2): 293–305.

Niskanen, W.A. (1971). *Bureaucracy and Representative Government*. Chicago, IL: Aldine-Atherton Incorporated.

Niskanen, W.A. (1973). *Bureaucracy: Servant or Master?* London, England: The Institute of Economic Affairs.

Niskanen, W.A. (1975). Bureaucrats and Politicians. *Journal of Law and Economics*, *81*(3): 617–644.

Niskanen, W.A. (1991). A Reflection on “Bureaucracy and Representative Government.” In A. Blais and S. Dion (Eds.), *The Budget-Maximizing Bureaucrat: Appraisals and Evidence* (pp. 13–32). Pittsburgh: University of Pittsburgh Press.

Niskanen, W.A. (1994). *Bureaucracy and Public Economics*. Vermont: Edward Elgar Publishing.

Peacock, A.T. (1992). *Public Choice Analysis in Historical Perspective* (pp. 1–83). Cambridge, England: Press Syndicate of the University of Cambridge.

Rolle, R.A. (1994). Indiana. In N.D. Theobald (Ed.), *The State of School Finance Policy Issues, 1994* (pp. 18–21). Washington, DC: American Education Finance Association.

Rossmiller, R.A. (1987). Achieving Equity and Effectiveness in Schooling. *Journal of Education Finance*, *12*(4): 561–577.

Sowell, T. (1993). *Inside American Education: The Decline, the Deception, and the Dogmas* (pp. 1–18, 285–303). New York, NY: The Free Press.

Theobald, N.D., and Rolle, R.A. (1995). Financing Professional Development Schools: Indiana School Finance Issues, 1995. In C. Edlefson (Ed.), *School Finance Policy Issues in the States and Provinces, Annual Update 1995* (pp. 54–58). Washington, DC: American Education Finance Association.

Tullock, G. (1965). *The Politics of Bureaucracy*. Washington DC: Public Affairs Press.

Walberg, H.J., and Fowler, W.J., Jr. (1987). Expenditure and Size Efficiencies of Public School Districts. *Educational Researcher*, *16*(7): 5–13.

Walberg, H.J., and Walberg, H. (1994). Losing Local Control. *Educational Researcher*, *23*(5): 19–26.

Wildavsky, A. (1964). *The Politics of the Budgetary Process*. Boston, MA: Little, Brown, and Company.

Young, R.A. (1991). Budget Size and Bureaucratic Careers. In A. Blais and S. Dion (Eds.), *The Budget-Maximizing Bureaucrat: Appraisals and Evidence* (pp. 33–58). Pittsburgh, PA: University of Pittsburgh Press.

Occupational Choices and the Academic Proficiency of the Teacher Workforce

Dan D. Goldhaber

University of Washington and The Urban Institute

Albert Yung-Hsu Liu

The Urban Institute

About the Authors

Dan D. Goldhaber is a Research Associate Professor in the Evans School of Public Affairs at the University of Washington and an Associated Scholar of the Urban Institute. Dr. Goldhaber received his Ph.D. in Economics from Cornell University.

Albert Yung-Hsu Liu is a Research Assistant at the Urban Institute.

Occupational Choices and the Academic Proficiency of the Teacher Workforce

Dan D. Goldhaber

University of Washington and The Urban Institute

Albert Yung-Hsu Liu

The Urban Institute

Acknowledgments

This project is supported by a Field Initiated Study grant from the Office of Educational Research and Improvement (now the Institute of Education Sciences), U.S. Department of Education. The views expressed in this paper do not necessarily reflect those of the University of Washington, the Urban Institute, or the U.S. Department of Education. We wish to thank Alissa Anderson, Emily Anthony, Meghan McElroy, and Dan Player for helpful comments.

Introduction

Recent research continues to support assertions by policymakers and professional educators that teacher quality is of paramount importance in promoting higher levels of student achievement. Among schooling characteristics, teacher effectiveness has been shown to explain the largest share of the variation in student achievement (Rivkin, Hanushek, and Kain 1998; Goldhaber, Brewer, and Anderson 1999). Differences in teacher quality have been found to explain more than one grade level equivalent of performance on standardized tests by their students (Hanushek 1992). Moreover, the impact of having particular teachers appears to explain students' achievement growth for several years (Sanders and Rivers 1996; Wright, Horn, and Sanders 1997).

There is broad agreement on the critical importance of teachers. However, there are also longstanding concerns about the quality of the current K–12 teacher workforce.[1] Dating back to the early and mid-1980s, commissions such as the National Commission on Excellence in Education (1983), the Carnegie Forum on Education and the Economy (1986), and more recently, the National Commission on Teaching and America's Future (1996) have all stressed the importance of teachers and the need to upgrade the skills of the teacher workforce. Today, teachers are better qualified, by some measures, than other college graduates. We find that teachers are more likely to hold advanced degrees and tend to have higher undergraduate grade point averages. But teachers also tend to be less academically proficient as measured by college entrance exam scores, the number of remedial courses

[1] Unless otherwise noted, we limit our discussion of teachers to those employed by public sector local education agencies.

they take in college, and the selectivity of the undergraduate colleges from which they graduate (Henke, Geis, and Giambattista 1996).

Some have suggested large across-the-board salary increases as a means of addressing concerns about the academic proficiency of the teacher workforce (as well as perceived teacher shortages). Others, however, point out that across-the-board increases may not be a particularly effective means of drawing more skilled personnel into teaching (Ballou and Podgursky 1995, 1997). In this paper, we investigate the hypothesis that observed differences in demonstrated academic proficiency may be due to the dissimilarity between teaching and other occupations in the *structure* of compensation. That is, we explore how compensation structures influence the career path decisions of prospective teachers.

We use data from the Baccalaureate and Beyond Longitudinal Study (B&B) to estimate the probability of progress through the teacher pipeline of a cohort of academic year (AY) 1992–93 college graduates. We find that, among other factors, college selectivity and college entrance exam scores predict progress through the teacher pipeline. We then estimate the returns to various attributes in the teacher and non-teacher labor markets and find that, while the public sector teacher labor market primarily rewards experience and advanced degree, the non-teacher labor market rewards these two attributes as well as college selectivity and technical major. These differential returns imply opportunity costs to enter the teaching profession that vary systematically based on an individual's college and undergraduate major.

This paper is arranged as follows: first, we provide background literature examining the relationship between teacher quality and opportunity costs. Second, we describe the data used in our analyses. Third, we present results from models describing the probability of progress through several stages of the teacher pipeline as well as salaries in chosen occupations. Fourth, we simulate opportunity costs of entering the teaching profession for men and women with different academic backgrounds. Finally, we offer some concluding thoughts on policy implications and our research agenda.

Researchers and professional educators have been unable to reach a consensus on a concise set of teacher characteristics that correlate with student achievement.

Background

Teacher Quality and Effectiveness

Using holistic measures of effectiveness that include observable and unobservable attributes, educational research has shown the overall impact of teacher quality to be the most important predictor of student achievement among school-related variables. Rivkin, Hanushek, and Kain (1998) estimate that, at a minimum, teacher quality accounts for 7.5 percent of the total variation in student achievement, a much larger share than any other educational input such as class size. This estimate is similar to that of an analysis by Goldhaber, Brewer, and Anderson (1999), who found that just over 8 percent of the variation in student achievement is due to differences between teachers.

These findings strongly suggest that raising the quality of the teacher workforce may be an effective lever for policymakers to raise student achievement levels. However, researchers and professional educators have been unable to reach a consensus on a concise set of teacher characteristics that correlate with student achievement. Goldhaber, Brewer, and Anderson (1999), for instance, estimated that only 3 percent of the contribution teachers make toward explaining student achievement is correlated with teacher experience, degree level, and other teacher characteristics included in their statistical model (e.g., race and gender). The remaining 97 percent is associated with teacher qualities or behaviors that could not be isolated and identified, such as understanding how children learn, being able to convey academic content, and connecting with the community. These traits and actions are certainly components of teacher quality but are difficult to include in statistical analyses.

Among studies that focus on observable inputs, relatively few studies that relate teacher characteristics to student outcomes include variables designed to measure the academic skills of teachers. Research that does include these attributes, however, tends to show correlations with teacher effectiveness. For example, studies all the way back to and including Coleman and Campbell (1966) have found teacher performance on a variety of standardized tests to be a good predictor of student achievement.[2]

While not all studies show a positive relationship between measures of teacher academic skills and student outcomes,[3] a meta-analysis by Greenwald, Hedges, and Laine (1996) suggests that teacher academic skills are correlated with student outcomes more often than characteristics such as graduate education and experience levels. Of the 24 studies of teacher ability reviewed, 50 percent reveal a positive and statistically significant effect, 4 percent a negative and statistically significant effect, and 46 percent no statistically significant effect.[4] This contrasts with studies of teacher experience and teacher education in which 72 percent of 46 studies and 68 percent of 68 studies, respectively, fail to show statistically significant effects.

A meta-analysis suggests that teacher academic skills are correlated with student outcomes more often than characteristics such as graduate education and experience levels.

Unfortunately, the teacher workforce tends to consist of college graduates of lesser academic proficiency: teachers are more likely to be drawn from the lower end of the distribution of standardized test scores and are more likely to have taken remedial coursework in college relative to their college graduate counterparts (Henke, Gies, and Giambattista 1996). Moreover, several studies suggest that this trend has become more pronounced in recent decades (Ballou 1996; Murnane and Singer 1991; Turner 1998; Vance and Schlechty 1982).

Compensation Structures

The differences between the structures of compensation in the teacher and non-teacher labor market may in part explain why individuals with stronger demonstrated academic skills tend to choose professions other than teaching. Although diverse, research on the structure of compensation in the non-teacher labor market suggests there are rewards for productivity on the job. Bretz and Milkovich (1989), for instance, estimate that 93 to 99 percent of private sector firms use some type of pay-for-performance plan for salaried individuals. Often, it takes an indirect form where individuals are rewarded for characteristics correlated with productivity. Studies have shown that the private sector labor market provides financial rewards for individuals who graduate from more selective colleges (Brewer, Eide, and Ehrenberg 1999) and who have higher standardized test scores (Murnane and Willet 1995).

In contrast, the explicit compensation structure used in over 95 percent of public local education agencies (LEAs) is known as the single salary schedule. This compensation structure differs significantly from that of most other occupations because it typically rewards only two characteristics within a given LEA: teacher experience and degree level (Odden and Kelley 1997). It is important to note that despite its rigidity, the teacher labor market may still reward individual characteristics such as college selectivity and undergraduate major. For example, there may be informal sorting of teachers

[2] Ferguson (1998), for instance, found measures of teachers' literacy skills, as measured by the Texas Examination of Current Administrators and Teachers, were associated with student achievement gains on mathematics tests, and Ferguson (1996) found a relationship between teachers' American College Testing Program (ACT) scores and students' fourth-grade reading test scores. Ehrenberg and Brewer (1994) found that college selectivity predicted students' test performance, and Strauss and Sawyer (1986) found a relationship between teachers' performance on licensure exams and students' test scores.

[3] For example, see Summers and Wolfe (1977) and Pugach and Raths (1983).

[4] Greenwald, Hedges, and Laine (1996) classify all teacher background characteristics as teacher ability except for teacher education and teacher experience. Examples include verbal and quantitative test scores.

on the salary schedule *within* an LEA that leads to differential rewards for characteristics *other than degree and experience levels*.[5] Additionally, we may observe sorting of teachers *between* LEAs if, for example, teachers from more selective colleges tend to teach at higher paying schools. Empirical evidence suggests that teachers with strong academic skills tend to migrate to LEAs and schools with high socioeconomic and high-achieving students (Lankford and Wyckoff 2000).

Such structural differences between the teaching profession and all other occupations suggest opportunity costs that vary markedly depending on individual background characteristics. For instance, if the non-teacher labor market rewards college selectivity more than the teacher labor market, we may predict that individuals who attend more selective colleges make greater financial sacrifices to enter the teaching profession relative to those who attend less selective colleges. Similarly, if the premium for majors such as math, science, or engineering is greater in the non-teacher labor market relative to the teacher labor market, we may predict similar differences in the opportunity cost of entering the teaching profession.

Structural differences between the teaching profession and all other occupations suggest opportunity costs that vary markedly depending on individual background characteristics.

The Teacher Pipeline

The purpose of this paper is to better understand career decisions related to the teaching profession made by recent college graduates. The teacher pipeline is a useful theoretical construct that provides a framework for a discussion of the supply of K–12 teachers. Our theoretical pipeline is presented in figure 1. Each node in the pipeline is based on a question from B&B (discussed below). The first node is based on the question, "*Have you ever trained or worked as a teacher at the preschool, grade school, or high school level, or are you currently considering teaching at these levels?"* Related questions pertain to completion of a student teaching class or a teacher certification program, or both, but fall outside the scope of this paper. Individuals who answer yes to the first node are then asked, *"Beginning around your graduation, how many applications for teaching positions have you submitted?"*

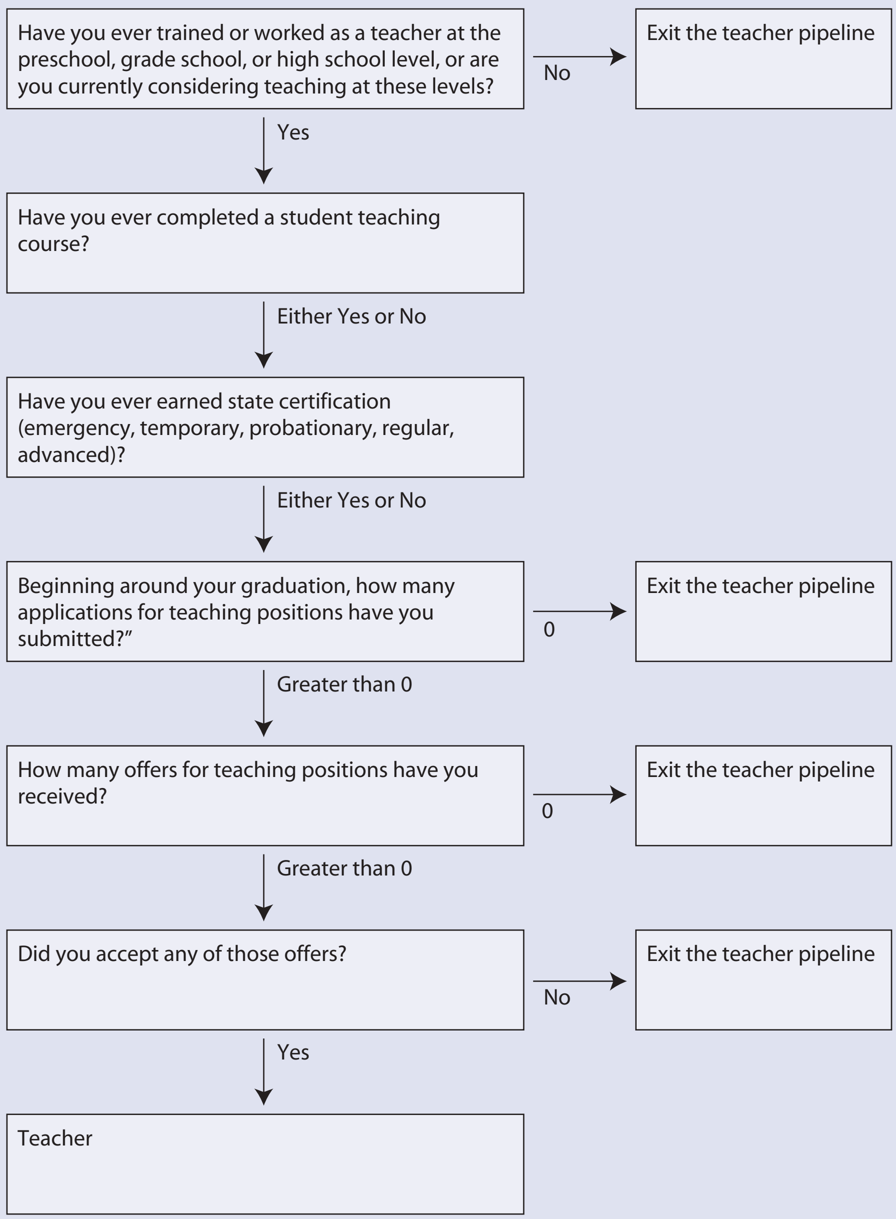
The subsequent nodes in the teacher pipeline are based on the questions *"How many offers for teaching positions have you received?"* and *"Did you accept any of those offers?"* These stages are fundamentally different from the previous two in that they reflect demand side forces in the teacher labor market. LEAs play the primary role in providing job offers, and accepting an offer is conditional on receiving at least one job offer.

If there are differentials between the teacher labor force and other recent college graduates on measures of academic skills, how are they reflected, if at all, in the various stages of the pipeline leading up to employment? Research on pathways to the classroom that compares teacher candidates to other college graduates suggests that attrition is spread throughout the pipeline. High school seniors who intend to major in education score lower on college entrance exams than their college-bound peers (Gitomer, Latham, and Ziomek 1999). Henke, Chen, and Gies (2000) and Henke, Gies, and Giambattista (1996) found that the college entrance examination scores for those who prepared to teach, were teaching, or were considering teaching were lower than those of their undergraduate counterparts.

Empirical research on decisions made by teachers is consistent with these trends as well. For instance, in a study on the attrition rate out of the teaching profession, Henke, Chen, and Gies (2000) found that graduates in the top quartile of college entrance examination scores are twice as likely as those in the bottom quartile to leave the profession in less than 4 years (32 percent vs.

[5] For example, it may be the case that teachers with attributes that are more in demand (e.g., math and science teachers) tend to be credited with more years of experience than those with backgrounds in other subjects. Such an example shows how districts might reward attributes other than degree and experience levels while ostensibly staying within the framework of the single salary schedule.

Figure 1. The teacher pipeline



SOURCE: Created by the authors based on questions from the U.S. Department of Education, National Center for Education Statistics (NCES), Baccalaureate and Beyond Longitudinal Study (B&B).

16 percent). Existing research on the teacher pipeline is succinctly summarized by Murnane and Singer (1991), who write that "college graduates with high test scores are less likely to take jobs, employed teachers with high test scores are less likely to stay, and former teachers with high test scores are less likely to return."

Data and Methodological Approach

This paper analyzes data derived from the Baccalaureate and Beyond Longitudinal Study (B&B), a nationally representative survey of more than 10,000 individuals who completed an undergraduate degree in AY 1992–93. Participants were initially selected from the 1992–93 National Postsecondary Student Aid Study (NPSAS:93) and were interviewed for B&B in 1994 (B&B:93/94) and again in 1997 (B&B:93/97). Students provided comprehensive information on themselves, including demographic characteristics, family background experiences, undergraduate and graduate level educational achievement, and labor market experiences through 1997. B&B devotes special attention to career decisions related to the teaching profession.

This paper analyzes data from the Baccalaureate and Beyond Longitudinal Study (B&B), a nationally representative survey of more than 10,000 individuals who completed college in 1992–93.

In the pipeline analysis portion of our paper, we confine our attention to the 10,080 individuals who responded to the first follow-up (B&B:93/94). For each node of the teacher pipeline construct that we analyze, we use a logistic probability model to estimate the probability of progress, where the dependent variable is a discrete choice variable that takes on binary values. The first node estimates affirmative answers to "*Have you ever trained or worked as a teacher at the preschool, grade school, or high school level, or are you currently considering teaching at these levels?"* We refer to this node as "*Have you ever considered teaching as a profession?*" The second node we estimate is conditional on progress through the previous node and is derived from the question, *"Beginning around your graduation, how many applications for teaching positions have you submitted?"* For simplicity, we truncate all numbers of applications greater than zero to 1, transforming the node to the question *"Beginning around your graduation, have you submitted at least one teaching application?"*

Omitting nonrespondents gives us a sample size of 9,845 observations, of which 3,235 have taught, trained, or are considering teaching as a profession. Of these 3,235 individuals, 1,702 submitted at least one teaching application; 1,533 submitted zero applications.

We group our explanatory variables into three vectors: demographic characteristics (gender and race/ethnicity), family background variables (parents' income, and parent occupation), and demonstrated academic proficiency (college selectivity index, undergraduate major, undergraduate grade point average, and college entrance exam score). Math, science, and engineering majors are aggregated into a composite technical major dummy variable. College selectivity data comes from *Barron's Profiles of American Colleges* and the College Board. We define the college selectivity as the average Scholastic Aptitude Test (SAT) score of the incoming class of AY 1989–90. College entrance exam scores are re-centered scores on the SAT or equivalent ACT scores. Details on variable construction are provided in appendix A.

Columns (1) and (2) in table 1 present the average characteristics of respondents who have not and have considered teaching, respectively. (Unless otherwise noted, differences are significant at the $p = .05$ level.) In the sample, men comprise 50 percent of the respondents who have not considered teaching, in contrast to 31 percent of the respondents who have considered teaching as a profession. And, while the proportions of White students in the two groups are statistically equal, we see differences between the groups for minority students. Native Americans and African Americans make up greater shares of those considering teaching relative to those who do not. Conversely, Asian and Pacific Islander Americans make up a lesser share of those considering teaching relative to those who do not.

Among family background characteristics, students who have considered teaching report having a mother employed as a teacher more frequently and tend to come from families with lower family incomes. The mean college entrance exam score and college selectiv-

Table 1. Descriptive statistics of the teacher pipeline

Characteristic	Have not considered teaching (1)	Considered teaching[1] (2)	Considered, have not applied (3)	Considered, applied[2] (4)
Male (in percent)	49.59	31.07	38.94	23.97
Native American (in percent)	0.53	0.87	0.78	0.94
Asian and Pacific Islander American (in percent)	5.42	2.01	3.00	1.12
African American (in percent)	5.17	7.23	8.94	5.70
Non-White Hispanic (in percent)	4.95	5.53	4.96	6.05
White (in percent)	83.24	83.77	81.47	85.84
Mother is a teacher[3] (in percent)	7.56	10.64	8.47	12.52
Father is a teacher[3] (in percent)	2.79	2.92	3.16	2.73
Family income (in dollars)	47,473	42,651	44,797	40,715
College selectivity index[4]	1,009	987	996	980
Technical major (in percent)	22.31	11.90	15.59	8.58
Undergraduate grade point average	3.06	3.12	3.04	3.18
College entrance exam score	1,015	977	995	961
Total respondents	6,610	3,235	1,533	1,702

[1]Considered teaching is defined as yes to *"Have you ever considered teaching as a profession?"*

[2]Applied is defined as yes to *"Beginning around your graduation, have you applied to at least one teaching job?"*

[3]Parent occupation comes from the parent survey of the 1992–94 National Postsecondary Student Aid Study (NPSAS:93), which was given only to a sample of the NPSAS:93 respondents.

[4]Average Scholastic Aptitude Test (SAT) score of the incoming class of AY 1989–90.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 1993/94 Baccalaureate and Beyond Longitudinal Study (B&B:93/94).

ity index for those who consider teaching are 38 and 22 points less, respectively, than the mean score and selectivity index of their counterparts. These results are consistent with previous research, discussed above, showing a negative relationship between some measures of academic skills and the propensity to enter and remain in the teaching profession. Respondents who have considered teaching also report majoring in technical fields at lower rates than those who have not considered teaching.

Sample characteristics for those who applied to zero teaching jobs and those who applied to at least one teaching job, conditional on having considered teaching as a profession, are presented in columns (3) and (4), respectively. The differences between these two groups are similar to the previous node in the teacher pipeline. Males comprise 39 percent of those who do not apply, 15 percentage points more than the percent of men among applicants. White students make up a higher share of nonapplicants; Asian and Pacific Islander Americans and African Americans make up lesser shares of applicants than of nonapplicants; and Native Americans and non-White Hispanic Americans make up greater shares of applicants than of nonapplicants.

The differences in family background characteristics are consistent with those of the previous node. The mean college entrance exam score and college selectivity index of those who apply are 34 and 16 points less, respectively, than the mean score and index of those who choose not to apply. And, among those who considered teaching, those with technical majors make up a lesser share of applicants than of nonapplicants.

An important point to note in the construction of the pipeline is that college choice may not be independent of the decision to become a teacher. It may be the case that individuals who intend to teach select a different set of colleges than their peers (Reback in press). If less selective colleges are more likely to offer teacher-training programs, then college selectivity is endogenous to our model. A similar line of reasoning is appropriate for undergraduate major because an education major is often a requirement for successful completion of a teacher training and certification program. Furthermore, if there is a correlation between undergraduate grade point average and major, grade point average may also be endogenous to our pipeline model. Thus, we estimate two variants of our pipeline model. The first includes the college-specific variables college selectivity, technical major, and undergraduate grade point average that we treat as exogenous. In the second variant, we substitute college entrance exam scores for all college-specific variables.

Our wage regression model analyzes a different subset of B&B data. To obtain sufficient variation in our sample, we stack observations from B&B:93/94 and B&B:93/97 and estimate the returns to various characteristics separately for teachers and non-teachers.[6] Because individuals potentially appear in our sample multiple times, we estimate our wage regressions using a random effects model.

We limit our analysis to individuals who provide information on their occupation and salary for their job in April of the survey year (B&B:93/97 asks for most recent job if the respondent is not employed in April). We further restrict our sample to individuals who worked full time, defined as working at least 30 hours per week. Because the single salary schedule is less prevalent among private sector teachers, we exclude teachers who are employed by private schools or whose sector of occupation cannot be determined. As a result, our final sample size is 13,636 observations, of which 1,421 are public sector teachers and 12,215 are non-teachers.

Our wage regression model analyzes a different subset of B&B by estimating the rates of return to various characteristics separately for teachers and non-teachers.

The dependent variable for our wage model is the natural log of annual salary. In each survey, respondents report dollar figures and the unit of time in which they report their salary, which can be any of per hour, per day, per week, per month, or per year. For purposes of comparison, we convert all wages to annual salary, using the conversion suggested by the documentation in the B&B, and adjust to 1997 dollars.[7]

The explanatory variables in the wage model are vectors of demographic characteristics (gender, race/ethnicity, marital status, and number of dependents), demonstrated academic proficiency (college selectivity index, undergraduate major, undergraduate grade point average, college entrance exam score, and advanced degree status), and labor market characteristics (years of full-time work experience and state of residence). Advanced degree is a dummy variable that indicates any of master's degree, first professional (e.g., JD), or doctoral degree.

It is important to note that by estimating the rates of return in the teacher and non-teacher labor market separately, we do not consider the role of wages in the selection of occupation. This potentially biases our findings if choice of occupation is correlated with both included explanatory variables and wages.

Table 2 presents the basic descriptive statistics of teachers in column (1) and of non-teachers in column (2). A number of differences between the two groups resonate with previous research and our findings from the teacher pipeline. The higher percentage of females among teachers is consistent with the historical gender composition of this profession (Bacolod 2001). In addition, relative to all other occupations, the teacher labor force is made up of higher shares of Whites and non-White Hispanic Americans and a lower share of Asian and Pacific Islander Americans. There are some striking differences between teachers and non-teachers in terms of their academic attributes and skills. Perhaps not surprisingly, teachers report majoring in technical fields less often than non-teachers. The mean college entrance exam score of teachers is 54 points less than that of non-teachers; for the college selectivity index, this difference is 41 points. However, teachers have higher rates of advanced degree attainment and a higher mean grade point average than non-teachers.

6 Such an approach is warranted by Chow tests that indicate structural differences between these subsamples.

7 Because teachers typically have 2 months of leave from work during their students' summer vacation, we also estimate our models using an annualized salary that assumes a 180-day year, a 36-week year, and a 10-month year. This is in contrast to the 260-day year, 30-week year, 12-month year for non-teachers. We experiment with these conversions, but these do not affect our wage regression results in significant ways.

Table 2. Descriptive statistics for the wage regression model

Characteristic	Public sector teachers (1)	Non-teachers (2)
Male (in percent)	22.17	47.09
Native American (in percent)	0.77	0.61
Asian and Pacific Islander American (in percent)	0.77	3.91
African American (in percent)	5.14	5.83
Non-White Hispanic (in percent)	6.26	4.70
White (in percent)	86.70	84.36
Married (in percent)	55.26	35.22
Dependents (including self)	2.15	1.72
College selectivity index	962	1,003
Technical major (in percent)	7.53	19.21
Undergraduate grade point average	3.24	3.03
College entrance exam score	950	1004
Advanced degree (in percent)	8.44	5.93
Years of experience	2.55	2.37
Salary (in dollars)	24,378	30,474
Total respondents	1,421	12,215

SOURCE: U.S. Department of Education, National Center for Education Statistics, 1993/94 Baccalaureate and Beyond Longitudinal Study (B&B:93/94).

Results

Progress Through the Teacher Pipeline

The first logistic probability model we estimate analyzes the decision to consider teaching as a profession. Table 3 presents the marginal probabilities for this model. (Marginal probabilities are calculated for a person with the mean characteristics.) Column (1) presents results that include college-specific characteristics (college selectivity index, undergraduate major, and undergraduate grade point average). Perhaps not surprisingly, males are predicted to be less likely (by about 14 percentage points) than females to consider teaching; in fact, based on the overall sample mean, they are almost 60 percent less likely than females to answer in the affirmative. African Americans are significantly more likely to consider teaching while Asian and Pacific Islander Americans are significantly less likely to consider it.

Family background characteristics also appear to influence the decision to consider teaching. Individuals whose parents have higher incomes are significantly less likely to consider teaching as a profession; for every additional $10,000 in income of an individual's parents, the probability that one considers teaching falls by 0.4 percentage points. Individuals who report a mother employed as a teacher are also 9 percentage points more likely to consider teaching.

Majoring in a technical field is associated with an 11 percentage point decline in the probability of answering yes. Also, having a 3.5 undergraduate grade point average rather than a 2.5 is predicted to increase the probability that individuals choose teaching by 3 percentage points. One explanation for the divergent findings between grade point average and college selectivity is that there may be grade inflation in education programs.

As we discussed above, it may be inappropriate to include measures of college selectivity, undergraduate major, and undergraduate grade point average in this stage of the pipeline analyses because individuals may choose their college and college courses based on their desire to teach. In column (2) of table 3, we present the marginal probabilities from a model that substitutes college entrance exam scores for these three measures of demonstrated academic proficiency. The coefficients of all nonacademic variables in the model change little in this specification of the model, and the results with regard to demonstrated academic skills also remain quite similar. An increase of 100 points in one's college entrance exam score is predicted to de-

Table 3. Marginal probabilities for the pipeline node I[1]

Characteristic	(1)		(2)	
Constant	0.0294	(0.0555)	0.0489**	(0.0297)
Male	-0.1415*	(0.0102)	-0.1622*	(0.0099)
Native American	0.0679	(0.0563)	0.075	(0.0565)
Asian and Pacific Islander American	-0.1833*	(0.0301)	-0.209*	(0.03)
African American	0.0568*	(0.0201)	0.0305	(0.0199)
Non-White Hispanic	0.0213	(0.0216)	0.0013	(0.0214)
Parents' income (each additional $10,000)	-0.0038*	(0.0011)	-0.0037*	(0.0011)
Mother is a teacher	0.0922*	(0.0260)	0.0939*	(0.026)
Father is a teacher	0.0087	(0.0454)	0.0012	(0.0454)
College selectivity index[2]	-0.0202*	(0.0046)		†
Technical major	-0.1129*	(0.0139)		†
Undergraduate grade point average	0.0336*	(0.0100)		†
College entrance exam score[2]		†	-0.0133*	(0.0028)
Total respondents	9,845		9,845	
-2 log likelihood	-5,965		-6,010	

*Indicates p value < .05.

**Indicates p value < .10.

†Not applicable.

[1]Affirmative to *"Have you ever considered teaching as a profession?"*

[2]Each 100-point increase.

NOTE: Standard errors are in parentheses.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 1993/94 Baccalaureate and Beyond Longitudinal Study (B&B:93/94).

crease the probability of considering teaching by roughly 1 percentage point.

The second node in the teacher pipeline that we estimate models responses conditional on having considered teaching as a career. Column (1) of table 4 presents the marginal probabilities for the model that includes college-specific variables. The results for this decision node are strikingly similar to those for considering teaching as a profession. Again, gender is a significant predictor of application among those who considered teaching as a profession; males are almost 14 percentage points less likely to apply for a teaching job than females.

Family background variables continue to predict progress through the teacher pipeline as well. An increase of $10,000 in parent income is associated with a 0.6 percent loss of the probability of progress, and having a mother who is a teacher increases the probability of progress by 13 percentage points.

The marginal probability for college selectivity is 2 percentage points for every 100 point change in college selectivity. For example, an individual who has considered teaching and graduates from Dartmouth College (college selectivity = 1130) is about 3 percentage points less likely to apply for at least one teaching job than someone at the University of Kentucky (college selectivity = 990), all else equal. Technical majors are 11 percentage points less likely to apply for at least one teaching job. These results should be treated with caution, however, because of the potential correlation between unobservable career desires and our included explanatory variables.

In column (2), we report the results when we substitute college entrance exam scores for college selectivity, technical major, and undergraduate grade point average. Again, the estimated coefficients change little and support the general finding that at this second node, individuals who attend more selective colleges (or have higher college entrance exam scores) are less likely to actually apply for a teaching job.

Table 4. Marginal probabilities for the pipeline node II[1]

Characteristic	(1)		(2)	
Constant	-0.0402	(0.1054)	0.3196*	(0.056)
Male	-0.1386*	(0.0202)	-0.1709*	(0.0197)
Native American	0.0343	(0.0973)	0.0364	(0.0974)
Asian and Pacific Islander American	-0.2122*	(0.0707)	-0.2325*	(0.0703)
African American	-0.1091*	(0.0366)	-0.1553*	(0.0358)
Non-White Hispanic	0.0497	(0.0408)	0.0238	(0.0401)
Parents' income (each additional $10,000)	-0.0055*	(0.0022)	-0.0064*	(0.0022)
Mother is a teacher	0.1332*	(0.0489)	0.1301*	(0.0487)
Father is a teacher	-0.0657	(0.0871)	-0.0376	(0.0863)
College selectivity index[2]	-0.0214*	(0.0087)		†
Technical major	-0.1125*	(0.0292)		†
Undergraduate grade point average	0.1194*	(0.0198)		†
College entrance exam score[2]		†	-0.0199*	(0.0055)
Total respondents	3,235		3,235	
-2 log likelihood	-2,140		-2,163	

*Indicates *p* value < .05.

†Not applicable.

[1]Affirmative to *"Beginning around graduation, have you applied to at least one teaching job?"*

[2]Each 100-point increase.

NOTE: Standard errors are in parentheses.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 1993/94 Baccalaureate and Beyond Longitudinal Study (B&B:93/94).

Wage Structures in the Teacher and Non-Teacher Labor Markets

Table 5 presents the random effects coefficient estimates of wages in the teacher and non-teacher labor markets. It is generally accepted practice to estimate the wages of males and females separately, given that there are differential returns by gender to various individual characteristics (Deolaliker 1993; Lundberg and Rose 2002; Schultz 1993). Since there are relatively few male teachers, the coefficient estimates for the male teacher model tend to be insignificant. For this reason, we do not report these results by gender; however, the magnitudes of the coefficient estimates were generally consistent with the estimates from the pooled sample of men and women.

Column (1) of table 5 presents the coefficient estimates for all teachers. We see a number of results that are consistent with the determination of salaries based on the single salary schedule. There are significant and positive rates of return to experience and attainment of an advanced degree. Each additional year of experience is associated with a 5.7 percent higher salary, and the completion of an advanced degree is correlated with a 9.5 percent higher salary, all else equal. In contrast, there are no differences in salaries, all else equal, based on a teacher's race/ethnicity (Native Americans are the exception), marital status, or number of dependents.[8] Nor are there statistically significant differences based on college selectivity, major, or undergraduate grade point average, measures that proxy for demonstrated academic proficiency and training.

These findings are certainly consistent with the use of the single salary schedule as a compensation structure. It is useful to recall that even with the use of the single salary schedule it was possible to observe returns to demonstrated academic proficiency, subject matter training, or other attributes in the teacher labor

[8] Men are found to receive a small wage premium in the teacher labor market, but the magnitude of this effect is far less than that in the non-teacher labor market.

Table 5. Random effects model results

Characteristic	Public sector teachers (1)		Non-Teachers (2)		Public sector teachers (3)		Non-Teachers (4)	
Constant	9.8999*	(0.1083)	9.4727*	(0.0735)	10.0405*	(0.0606)	9.6663*	(0.0424)
Male	0.0501*	(0.0189)	0.1630*	(0.0123)	0.0469*	(0.0184)	0.1681*	(0.012)
Native American	-0.1735*	(0.0864)	0.1021	(0.077)	-0.1786*	(0.0862)	0.0720	(0.0776)
Asian and Pacific Islander American	-0.0220	(0.1002)	0.0558**	(0.0311)	-0.0069	(0.0998)	0.0708*	(0.0312)
African American	0.0593**	(0.035)	-0.0113	(0.0262)	0.0546	(0.0347)	-0.0201	(0.0262)
Non-White Hispanic	-0.0113	(0.0340)	-0.0285	(0.0296)	-0.0058	(0.0339)	-0.0399	(0.0298)
Married	0.0057	(0.0200)	0.0123	(0.0147)	0.0033	(0.0201)	0.0219	(0.0147)
Dependents (including self)	-0.0030	(0.0084)	0.0486*	(0.0069)	0.0014	(0.0087)	0.0435*	(0.0070)
College selectivity index[1]	0.0077	(0.0087)	0.0122*	(0.0058)		†		†
Technical major	0.0389	(0.0293)	0.1306*	(0.0154)		†		†
Grade point average	0.0356**	(0.0189)	0.0883*	(0.0124)		†		†
College entrance exam score[1]		†		†	0.0054	(0.0049)	0.0208*	(0.0035)
Advanced degree	0.0953*	(0.0271)	0.0666*	(0.0216)	0.0968*	(0.0269)	0.0827*	(0.0216)
Years of experience	0.0570*	(0.0050)	0.0943*	(0.0031)	0.0562*	(0.0050)	0.0934*	(0.0031)
Total respondents	1,421		12,215		1,421		12,215	
-2 res log likelihood	458.9		20,301.8		450.7		20,375.7	

*Indicates p value < .05.

**Indicates p value < .10.

†Not applicable.

[1]Each 100-point increase.

NOTE: Standard errors are in parentheses.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 1993/94 Baccalaureate and Beyond Longitudinal Study (B&B:93/94).

market as a result of sorting within or between districts. As we describe below, however, the fact that we do not observe returns to academic skills or subject matter training implies that the teacher labor market differs markedly from the non-teacher labor market.

Column (2) shows the random effects model estimates of the returns for the non-teacher labor market. Unlike the teacher labor market, we find strong evidence of returns to gender, family composition, measures of demonstrated academic skills and training, and labor market experience. (Unless otherwise noted, all of the differences between the teacher and non-teacher labor markets discussed below are statistically significant at the 95 percent confidence level). Males are predicted to earn a pay premium of about 16 percent compared to a premium in the teacher labor market of only 5 percent. There is also a premium for individuals with children; we observe a 5 percent higher wage for every additional dependent. These findings for the non-teacher labor market are broadly consistent with previous findings in the literature.

There are important differences between the two labor markets in the returns to demonstrated academic proficiency and specialization. While there is little evidence in the teacher labor market of returns to college selectivity, in the non-teacher labor market, individuals who attend more selective colleges are predicted to earn more. An individual from a college with a 100 point higher selectivity index than the average is predicted to make 1.2 percent more in salary. We also observe that the return to undergraduate grade point average is higher in the non-teacher labor market than in the teacher labor market. A 1-point increase in grade point average is associated with an 8.8 percent increase in salary, whereas in the teacher labor market the same increase is associated with only a 3.6 percent increase in salary (the difference between these two estimates is significant at the 90 percent confidence level). Finally, the non-teacher labor market appears to significantly reward individuals who major in technical subjects; we estimate a 13 percent pay premium for those who have either a math, science, or engineering major.

It is worth noting that the return to experience in the non-teacher labor market, where there is estimated to be a 9.4 percent pay premium for an additional year of experience, is significantly larger than the 5.7 percent return to an additional year of experience in the teacher labor market. This difference is potentially important if long-term earnings potential influences occupational choice. Given that we are estimating wages for a sample of recent college graduates (they graduated in AY 1992–93), most have relatively little labor market experience. Furthermore, we do not know how many years of *teaching* experience teachers are credited as having. Thus, our estimates of the returns to experience should be treated with caution.

As we discussed in the teacher pipeline section, the choice of college—and by extension, undergraduate major and grade point average—may be endogenous to the selection of occupation. We address this problem in the same manner as in our approach in the teacher pipeline section and substitute college entrance exam score for college-specific variables in this paper.[9] In columns (3) and (4) we present the results from our random effects model that substitutes college entrance exam score for college selectivity, undergraduate major, and undergraduate grade point average. Our overall findings with regard to academic skills change little with this model specification. In the non-teacher labor market, wages are predicted to be higher for individuals with higher SAT scores (by about 2 percent for every 100 SAT points); however, there is no corresponding premium for SAT scores in the teacher labor market.

In the non-teacher labor market, wages are predicted to be higher for individuals with higher SAT scores; however, there is no corresponding premium for SAT scores in the teacher labor market.

In our final model specification (which is not included in the table), we estimate models that include both college-specific variables and college entrance exam scores. Consistent with our prior findings, neither individual SAT score or college selectivity are rewarded in the teacher labor market. In contrast, we observe marginally significant (at the 10 percent level) rates of return for these two characteristics in the non-teacher labor market.

The differences in compensation structure between the teacher and non-teacher labor markets imply that the financial opportunity costs associated with teaching vary systematically based on individual background characteristics. The implications of these differences are discussed below.

Opportunity Cost Simulations

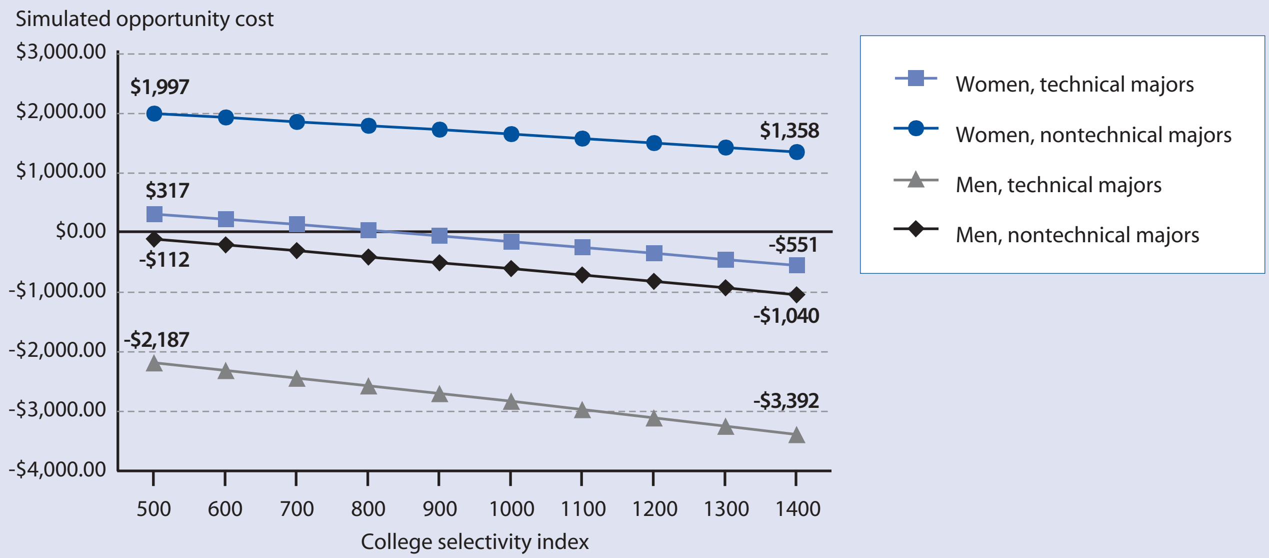
In this section, we discuss our simulations of the financial opportunity costs associated with teaching as opposed to entering the non-teacher labor market. We simulate the costs for men and women with different academic backgrounds (training, degree level, and college selectivity) as well as for individuals with different experience levels. In general, we calculate opportunity costs as the average of the difference of the predicted salary as a teacher less the predicted salary as a non-teacher. A more detailed discussion of the simulation is provided in appendix B.

Based on the actual characteristics in our sample, we estimate what men and women would have earned had they chosen the alternate occupation (e.g., teacher entered the non-teacher labor market and vice versa). The average predicted wage for female teachers is $23,692, about $600 less than the predicted wage for female college graduates outside of teaching. The average predicted wage for male teachers is $24,975, which is about $5,000 less than the predicted wage for males outside of teaching.

Figures 2 through 4 show the results of various simulations. The horizontal axis illustrates the opportunity cost depending on the selectivity of the college attended.

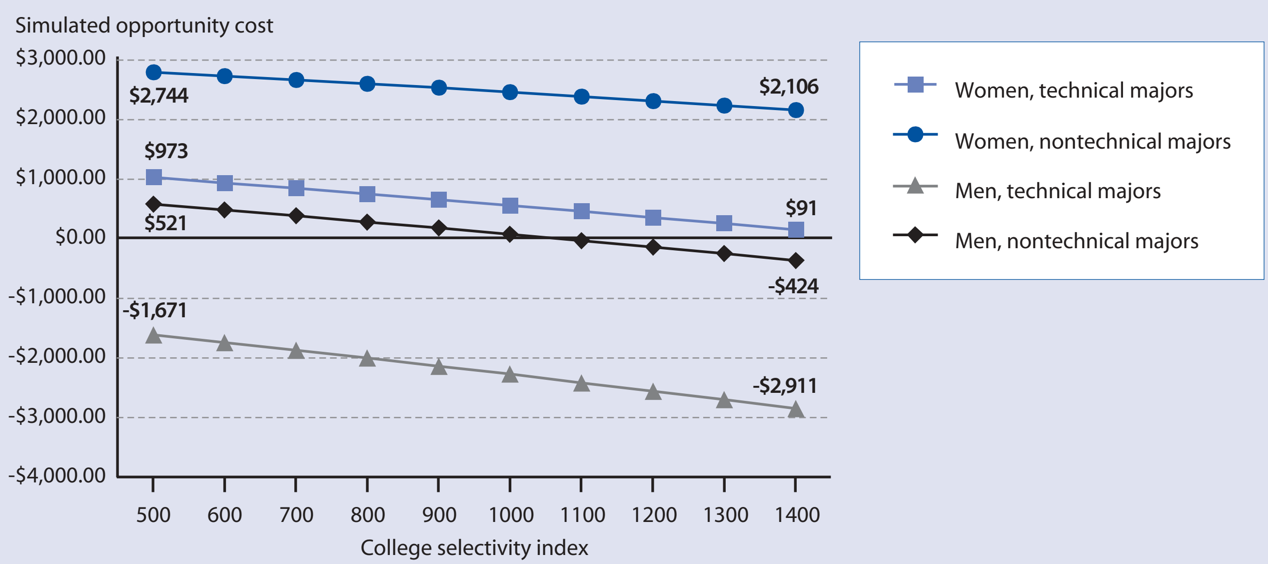
[9] One might argue that scores on college entrance exams are themselves endogenous; for instance, individuals who wish to enter a particular occupation may study more than those wishing to enter a different occupation. This argument, however, seems less plausible than the argument for the endogeneity of college selectivity and major.

Figure 2. Simulated opportunity costs of entering the teaching profession: No work experience, no advanced degree



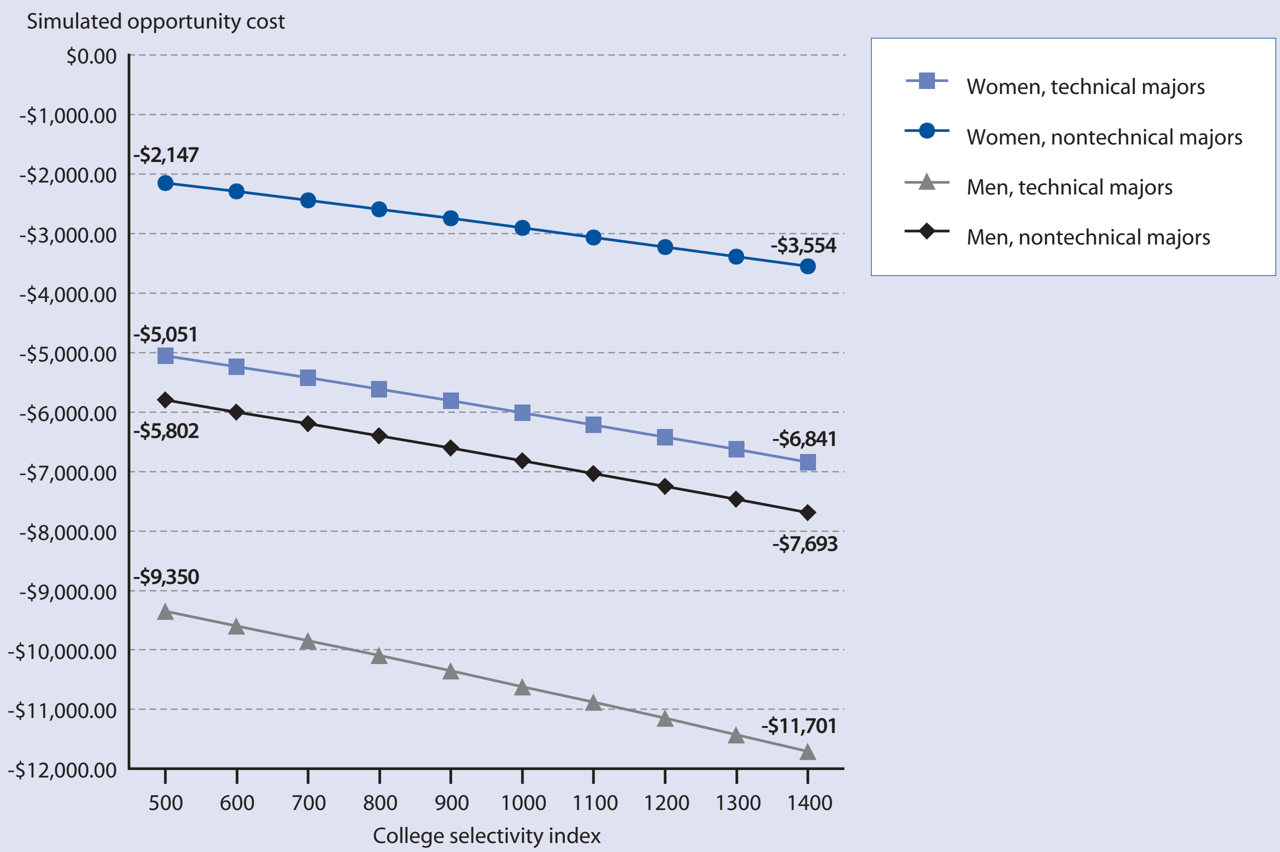
SOURCE: Authors' simulations based on the U.S. Department of Education, National Center for Education Statistics, Baccalaureate and Beyond Longitudinal Study (B&B).

Figure 3. Simulated opportunity costs of entering the teaching profession: No experience, with advanced degree



SOURCE: Authors' simulations based on the U.S. Department of Education, National Center for Education Statistics, Baccalaureate and Beyond Longitudinal Study (B&B).

Figure 4. Simulated opportunity costs of entering the teaching profession: Five years' experience, no advanced degree



SOURCE: Authors' simulations based on the U.S. Department of Education, National Center for Education Statistics, Baccalaureate and Beyond Longitudinal Study (B&B).

The plotted line represents the predicted salary for an individual who enters the teaching profession less the predicted salary if that individual were instead to enter the non-teacher labor market. Negative numbers imply that one is predicted to have a higher salary as a non-teacher whereas positive numbers imply that one is predicted to make more as a teacher.

Figure 2 presents the simulated opportunity cost for the average individual with zero years of work experience and no advanced degrees. The negative slope of each line reflects the lower estimated return to college selectivity in the teacher labor market relative to the non-teacher labor market. The top line, representing the opportunity costs for women who have a nontechnical major, lies above zero regardless of the selectivity of the college attended (though it is closer to zero for more selective colleges), implying that it is more financially lucrative for them to become teachers (the exact opportunity cost, of course, is contingent on the calculation of annual salaries). This is not necessarily the case for women who completed a technical major in college. Although those who attend less selective institutions earn slightly more as teachers than they otherwise are predicted to earn, women with technical majors who attend a college with an average SAT score of about 800 or greater are predicted to earn more outside of the teaching profession. In contrast to women, men are predicted to earn more outside of teaching regardless of their choice of college or major.

However, the opportunity costs rise for men who attend more selective institutions and have technical majors.

Because the teacher labor market provides explicit returns to both experience and level of education, we present simulations in which we vary these dimensions. These results reflect the patterns discussed above: men face larger opportunity costs than women, those with technical majors face larger opportunity costs than nontechnical majors, and opportunity costs rise for individuals who attend more selective institutions.

Figure 3 shows the simulated opportunity costs for individuals who hold an advanced degree but have no work experience. Women who major in nontechnical subjects are still predicted to earn more as teachers regardless of college selectivity, significantly more than is the case for women who do not hold an advanced degree. For example, a woman who attends a college with an average SAT of 1000 and holds a master's degree is predicted to earn $2,400 more as a teacher, while a similar woman who does not hold a master's degree is predicted to earn only $1,700 more as a teacher. The opportunity costs for women with technical majors, with and without an advanced degree, are $503 and –$146, respectively. The opportunity costs for men without technical majors, with and without an advanced degree, are $17 and –$607, respectively. These figures imply that an advanced degree makes it more financially worthwhile to be a teacher for women with or without technical majors and for men with nontechnical majors. The opportunity costs for men with technical majors still imply that they are predicted to be financially better off as non-teachers.

For most individuals, earnings are predicted to be higher outside of teaching than in the teaching profession.

In our final simulation, we simulate the opportunity costs for individuals with 5 years of labor market experience and no advanced degree. Figure 4 illustrates the importance of the differential returns to experience between labor markets. With 5 years of experience and no advanced degree, women with nontechnical majors (college selectivity = 1000) must now *sacrifice* $2,900 to teach, in contrast to the same individuals with no experience who *receive* $1,657 to teach. The effects of 5 years of experience for all groups, regardless of college selectivity, major, or gender, show that for most individuals, earnings are predicted to be higher outside of teaching than in the teaching profession. At the extreme, the estimated opportunity cost to enter the teaching profession for males with technical majors who graduate from more selective colleges can reach $10,000!

Conclusions

The results presented in this study suggest that measures of demonstrated academic proficiency predict the likelihood of potential teachers advancing through the teacher pipeline. Individuals with stronger demonstrated academic proficiency (e.g., higher college entrance exam scores or college selectivity) are less likely to consider teaching and less likely to apply for a job as a teacher. This may be explained, in part, by the compensation structure in teaching, since our salary structure results reveal important differences between the teacher and nonteacher labor markets in terms of the rewards associated with academic skills and training. These differences, which are consistent with the use of the single salary schedule, suggest that individuals with stronger academic backgrounds or technical training face greater opportunity costs to being teachers, all else equal.

Unless individuals systematically differ in terms of the value they place on nonpecuniary job characteristics, we would expect those with higher SAT scores, technical majors, or graduates from more selective colleges to be less likely to teach. This is exactly what we find in our analyses of various points on the teacher pipeline: those with higher college entrance exam scores, those who go to more selective colleges, and those who graduate with a technical major are less likely to have taught, trained as a teacher, or considered teaching as a profession.

Though the results of the study are suggestive of a causal connection between compensation structure and the decisions made by individuals in the teacher pipeline, we are cautious about drawing strong conclusions since the current study is limited in several respects. In particular, in the analyses of teacher compensation structure, we focus exclusively on salaries, omitting nonpecuniary rewards as well as bonuses and rewards for nonclassroom work in the school. Individuals certainly consider other characteristics of jobs that are part of a compensation package (e.g., health and retirement benefits) as well as other nonpecuniary job characteristics (e.g., pressure at work and collegiality). There is in fact evidence that teachers are particularly sensitive to nonpecuniary job characteristics when making decisions about the schools and districts in which they teach (Loeb 2001). Furthermore, the underlying assumption of career choice models is that individuals choose careers and jobs that maximize utility, so a limitation of the current study is that we do not explicitly treat the selection of occupation as endogenous, despite the fact that individuals self-select into occupations.[10] Future work on the impact of the compensation structure in teaching on the decisions made by individuals to enter or remain in the profession should explore these issues more fully.

Appendix A: Variable Construction

This appendix details how we constructed the variables used in our analysis. Questions regarding the survey should be directed to the National Center for Education Statistics (NCES) in the U.S. Department of Education. Unless otherwise noted, constructed variables with missing data are coded to 0 and flagged.

Demographic Characteristics

We identify the gender of the respondent using B2RSEX from the B&B:93/97 and supplement it with the gender variables RSEX, GENDER, SEX, and M_STGEN from B&B:93/94 and NPSAS:93.

We use B2ETHNIC to create a set of separate dummy variables for race for the categories Native American, Asian and Pacific Islander American, Black, Non-White Hispanic, Other, and White.

Demonstrated Academic Skills and Educational Attainment

Respondents' college selectivity is defined as the average SAT score of the incoming class of AY 1989–1990 at the undergraduate institution from which they graduated. Data are imputed from *Barron's Profiles of American Colleges*. If only ACT scores are available, we convert these scores to SAT using the conversion table found at the web site *http://www.collegeboard.com/sat/cbsenior/yr2001/pdf/ten.pdf*. Missing values are supplemented with data supplied by the College Board.

We use B2BAMAJR to construct a set of separate dummy variables for undergraduate major for the categories business and management, education, engineering, health professions, public affairs/social services, biological sciences, mathematics and other sciences, social science, history, humanities, psychology, and other. We then construct a composite technical major by combining engineering, biological sciences, and mathematics and other sciences majors. All other majors are defined as nontechnical majors.

For undergraduate grade point average, we use NORMGPA from the transcript survey if it is available. Otherwise, we use data from CUMULGPA in B&B:93/94, which asked respondents about their cumulative grade point average.

Advanced degree is calculated from questions on higher educational achievement. We acknowledge an advanced degree if the respondent has earned a master's, first professional, dual degree in which one degree was master's or first professional, or doctoral degree before or on the month of relevant employment in each of the surveys (see Salary and Employment). For B&B:93/94 respondents, we use PB01DGDT-PB03DGDT for the date one received the degree, and PB01PROG-PB03PROG for the program type. In the event that it is a dual degree, program types can be found in P01PRG01 and P01PRG02. Construction for B&B:93/97 recipients are analogous to B2P01PRG-B2P05PRG as program

[10] For instance, it is possible that individuals who know they are not likely to excel or earn a high salary in teaching are likely instead to choose other professions (and vice versa). To the degree that variables omitted from our wage equations and important to the determination of salaries are correlated with both measures of academic skills and the choice of occupation, our coefficient estimates will be biased.

types and B2P01DGD-B2P05DGD as the dates respondents received the degree.

Family Status

Marriage status in April 1994 is derived from B2MAR494, and marriage status in April 1997 is derived from B2MAR497. Single, divorced, separated, widowed, and living in a marriage-like relationship are considered not married.

Total number of dependents (including the respondent) is derived from TOTNUMDP and B2TOTDEP for B&B:93/94 and B&B:93/97 respondents, respectively.

Labor Market Experience

We calculate labor market experience using EMPL9207-EMPL9212 for B&B:93/94 respondents and B2EM9207-B2EM9212 for B&B:93/97 respondents. We consider a month employed if the month is after the graduation date and before or equal to the relevant month of employment (see Salary and Employment).

Parent Background Variables

Parents' income is derived from CINCOME in the NPSAS:93 survey, which is defined as total family income from the 1991 calendar year. If the student is not a dependent, then CINCOME contains the student's income.

Parents of a subset of the NPSAS:93 respondents were interviewed about the financing of their child's undergraduate education. MOMOC and DADOC contain the occupation of the mother and father, respectively. We code a flag if parent occupation is "school teacher."

Pipeline Variables

The first pipeline node, "*Have you ever trained or worked as a teacher at the preschool, grade school, or high school level, or are you currently considering teaching at these levels?*" is derived from TEACH in B&B:93/94. For those who replied yes to TEACH, responses to "*Beginning around your graduation, how many applications for teaching positions have you submitted?"* are found in APPLICAT. We recode this continuous variable to a dichotomous one, where values greater than 0 are coded to 1 and responses of 0 are coded to 0.

Salary and Employment

For B&B:93/94, respondents provided earnings data on their primary job in April 1994. Data are reported as real dollar figures in APRANSAL, which we convert to 1997 dollars. For B&B:93/97, respondents provided data for their job in April 1997, or the last month of their most recent job. Earnings are reported in B2AJBSAL with a corresponding wage rate (per hour, per day, per week, per month, per year) in B2AJRATE. We use the strategy employed by NCES in converting all figures to annual wages. If wages were reported per hour, we calculate the weekly wage with the hours worked per week, B2AJBHRS. We assume a 260-day work year, a 52-week year, and a 12-month year. Wages are converted to 1997 dollars.

For each salary figure, respondents were asked about their occupation. Because the occupation codes for the two follow-ups differ, MPR Associates reconstructed the variables to match the coding scheme of B&B:93/97. These two new variables, AJOBOCCR and B2AJOBR, are available from MPR Associates.

State of Employment

B2STATE and B1STATE provide information for respondents to B&B:93/94 and B&B:93/97, respectively. Individuals in Puerto Rico, the Virgin Islands, or other countries were coded as other, and those with missing information were coded as missing. We construct a set of separate dummy variables for each state.

Teacher Sector

For respondents to B&B:93/94 we use TCHSCHL, which provides the NCES code for the school taught at most recently. We assume that the questions for April 1994 salaries correspond to this school. Schools without NCES codes but for which there are sufficient data on name, city, and state of school are imputed with codes from the Common Core of Data (CCD) and the Private School Universe Survey (PSS).

Teachers in B&B:93/97 are assigned sectors using BSCL01-B2SCL05. We identify the relevant school in which the respondent was teaching in April 1997 or most recent job. We also impute codes for schools for those for whom we can.

Appendix B: Simulations

This appendix describes our approach in estimating opportunity costs of entering the teaching profession. For each profession, we estimate the predicted natural log of salary using each individual's characteristics in the wage regression model. We then convert these estimates to salary in dollars and take the average. To estimate the opportunity cost, we take the mean of the difference between salary as a teacher and salary as a non-teacher. When we estimate the opportunity cost for, say, women without technical majors, we estimate the mean opportunity cost for all individuals in our sample and use each person's individual characteristics, except that we force each person to be a woman and to have a technical major. For figures 2 through 4, we estimate the opportunity cost at different levels of the index of college selectivity.

References

Bacolod, M. (2001). *The Role of Alternative Opportunities in the Female Labor Market in Teacher Supply and Quality: 1940–1990*. Unpublished Paper from the University of California, Los Angeles.

Ballou, D. (1996). Do Public Schools Hire the Best Applicants? *The Quarterly Journal of Economics, 111*(1): 97–133.

Ballou, D., and Podgursky, M. (1995). Recruiting Smarter Teachers. *Journal of Human Resources, 30*(2): 327–338.

Ballou, D., and Podgursky, M. (1997). *Teacher Pay and Teacher Quality.* Kalamazoo, MI: Upjohn.

Bretz, R., and Milkovich, G. (1989). *Performance Appraisal in Large Organizations: Practice and Research Implications*. Ithaca, NY: Cornell University.

Brewer, D., Eide, E.A., and Ehrenberg, R.G. (1999). Does it Pay to Attend an Elite Private College: Cross-Cohort Evidence on the Effects of College Type on Earnings. *Journal of Human Resources, 34*(1): 104–123.

Carnegie Forum on Education and the Economy. (1986). *A Nation Prepared: Teachers for the 21st Century.* New York: The Forum.

Coleman, J.S., and Campbell, E.Q. (1966). *Equality of Educational Opportunity*. Washington, DC: U.S. Government Printing Office.

Deolaliker, A. (1993). Gender Differences in the Returns to Schooling and in School Enrollment Rates in Indonesia. *Journal of Human Resources, 28*(4): 899–932.

Ehrenberg, R.G., and Brewer, D.J. (1994). Do School and Teacher Characteristics Matter? Evidence from High School and Beyond. *Economics of Education Review, 13*(1).

Ferguson, R. (1996). How and Why Money Matters: An Analysis of Alabama Schools. In H. Ladd (Ed.), *Holding Schools Accountable*. Washington, DC: Brookings Institution Press.

Ferguson, R. (1998). Can Schools Narrow the Black-White Test Score Gap. In C. Jencks and M. Phillips (Eds.), *The Black-White Test Score Gap*. Washington, DC: Brookings Institution Press.

Gitomer, D., Latham, A., and Ziomek, R. (1999). *The Academic Quality of Prospective Teachers: The Impact of Admissions and Licensure Testing.* Princeton, NJ: Educational Testing Service.

Goldhaber, D., Brewer, D., and Anderson, D.J. (1999). A Three Way Error Components Analysis of Educational Productivity. *Education Economics, 7*(3): 199–208.

Greenwald, R., Hedges, L., and Laine, R. (1996). The Effect of School Resources on Student Achievement. *Review of Education Research, 66*(3): 361–396.

Hanushek, E. (1992). The Trade-Off Between Child Quantity and Quality. *Journal of Political Economy, 100*(1): 84–117.

Henke, R., Chen, X., and Geis, S. (2000). *Progress Through the Teacher Pipeline: 1992–93 College Graduates and Elementary/Secondary School Teaching as of 1997* (NCES 2000–152). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Henke, R., Geis, S., and Giambattista, J. (1996). *Out of the Lecture Hall and Into the Classroom: 1992–93 College Graduates and Elementary/Secondary School Teaching* (NCES 96–899). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Lankford, H., and Wyckoff, J. (2000). *The Labor Market for Public School Teachers: A Descriptive Analysis of New York State's Teacher Workforce.* Paper presented at the New York State Educational Finance Research Consortium, New York, NY.

Loeb, S. (2001). Teacher Quality: Its Enrichment and Potential for Improving Pupil Achievement. In D. Monk and H. Walberg (Eds.), *Improving Education Productivity*. Greenwich, CT: Information Age Publishing.

Lundberg, S., and Rose, E. (2002). The Effects of Sons and Daughters on Men's Labor Supply and Wages. *The Review of Economics and Statistics, 84*(2): 251–268.

Murnane, R., and Singer, J. (1991). *Who Will Teach?* Cambridge, MA: Harvard University Press.

Murnane, R.J., and Willett, J.B. (1995). The Growing Importance of Cognitive Skills in Wage Determination. *The Review of Economics and Statistics, 55*(2): 251–266.

National Commission on Excellence in Education. (1983). *A Nation At Risk: The Imperative for Educational Reform*. Washington, DC: U.S. Department of Education.

National Commission on Teaching and America's Future. (1996). *What Matters Most: Teaching for America's Future*. New York: Columbia University, National Commission on Teaching and America's Future.

Odden, A., and Kelley, C. (1997). *Paying Teachers for What They Know and Do: New and Smarter Compensation Strategies to Improve Schools*. Thousand Oaks, CA: Corwin Press, Inc.

Pugach, M., and Raths, J.D. (1983). Testing Teachers: Analysis and Recommendations. *Journal of Teacher Education, 34*(1): 37–43.

Reback, R. (in press). The Impact of College Course Offerings on the Supply of Academically Talented Public School Teachers. *Journal of Econometrics*.

Rivkin, S., Hanushek, E., and Kain, J. (1998). *Teachers, Schools, and Academic Achievement* (Working Paper No. 6691). Cambridge, MA: National Bureau of Economic Research.

Sanders, W., and Rivers, J. (1996). *Research Project Report: Cumulative and Residual Effects of Teachers on Future Student Academic Achievement*. Knoxville, TN: University of Tennessee, Value-Added Research and Assessment Center.

Schultz, P.T. (1993). Investments in the Schooling and Health of Women and Men: Quantities and Returns. *Journal of Human Resources, 84*(4): 694–732.

Strauss, R., and Sawyer, E. (1986). Some New Evidence on Teacher and Student Competencies. *Economics of Education Review, 5*(1): 41–48.

Summers, A., and Wolfe, R. (1977). Do Schools Make a Difference? *American Economic Review, 67*(4): 639–652.

Turner, S.E. (1998). *The Training of Teachers: The Changing Degree Output in the Area of Education.* Paper presented at the 1998 meeting of the Association of Public Policy and Management.

Vance, V.S., and Schlechty, P.C. (1982). The Distribution of Academic Ability in the Teaching Force: Policy Implications. *Phi Delta Kappan, 64,* 22–27.

Wright, P., Horn, S., and Sanders, W. (1997). Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation. *Journal of Personnel Evaluation in Education, 11,* 57–67.

Variation in the Rewards for a Teacher's Performance: An Application of Quantile Regressions

Sherrilyn M. Billger
Illinois State University

About the Author

Sherrilyn M. Billger is Assistant Professor of Economics at Illinois State University. Her research focuses on incentives for teachers and principals, school performance, and evaluating single-sex education. Previously, she taught at Union College, and was also a guest researcher at the Max Planck Institute for Human Development in Berlin, Germany.

Variation in the Rewards for a Teacher's Performance: An Application of Quantile Regressions

Sherrilyn M. Billger

Illinois State University

Introduction

A major difficulty in the implementation of incentives for educators lies in the measurement of performance. In addition, given salary schedules and union contracts at public schools, it is difficult to examine the impact of teaching performance on pay. On the other hand, private schools tend to have more flexible pay structures, offer larger merit awards, and have broader teacher support for incentives.[1] In addition, flexible pay could imply that variation in pay is correlated with teacher performance, as measured by skills, principals' assessments, training, mentor status, and student achievement. Heterogeneity (as estimated with quantile regressions) in the effects of such measures on pay also reflects how the returns to measured performance correlate with unobservables. Specifically, this study explores how teacher quality and pay structure impact salaries using quantile regressions. This technique enables the examination of a few important phenomena: the factors that contribute to, or detract from, salary dispersion; differential impacts of qualifications and performance throughout the conditional salary distribution; and the interaction between unobservable determinants of salary (perhaps correlated with ability) and individual covariates.

Pay for Performance?

Incentive pay in education is indeed regaining popularity, and a number of case studies highlight the results of such programs. For instance, Ladd (1999) investigates the Dallas school accountability and incentive program using panel data for urban schools in Texas. She finds that the program positively impacted seventh-grade test scores, though only for Whites and Latinos. In addition, Jacobson (1989) examines a New York State school district that implemented an incentive pay plan to reduce absenteeism. Though teachers did not know the exact amount of the bonus they would receive, absenteeism was substantially reduced and perfect attendance increased as a result of the plan.

Existing and proposed programs reward school and/or teacher performance, but teacher-level incentives create certain challenges, as outlined in Murnane and Cohen

[1] Ballou and Podgursky (1997).

(1986). They state that the very nature of a teacher's work makes incentive compensation difficult to implement. For instance, teachers could be rewarded for student test score gains that are the result of the cumulative education received, rather than the impact of one strong teacher. Furthermore, much of the education produced in schools results from collaborative work among all staff. Nonetheless, the authors state that successful (i.e., long-lived) merit pay plans have common characteristics. Specifically, schools that offer these plans offer very small merit bonuses, serve rather homogenous student bodies, and have very good working conditions and high salaries. Murnane and Cohen cite additional benefits that arise from these plans: more meaningful dialogue among staff and greater community support for the schools. Both may be valuable, even apart from effects on student achievement.

Successful (i.e., long-lived) merit pay plans have common characteristics.

Solmon and Podgursky (2000) provide a more recent discussion of the relevant issues concerning merit pay for teachers. Using feedback from practitioners, the authors cite the remaining problems involving performance measurement, fair implementation, and teacher morale. Nonetheless, increasing numbers of teachers favor performance-based compensation. Furthermore, Solmon and Podgursky suggest that such compensation should depend upon factors including the number of tasks/functions, quality of work, awards received, degrees, evaluated performance, and student achievement in terms of test score gains and attendance. In fact, Ballou (2001) reveals that teachers at approximately 10 percent of public districts and private schools are affected by incentive compensation. He finds that bonuses are substantial, particularly in some private schools. It remains to be seen whether this compensation is properly linked to performance. Also important is determining which attributes are rewarded in teachers' compensation.

The existing literature documents the role of personal, school, and community characteristics in determining teacher salaries. As Hanushek (2002) states, education and experience explain much of the existing variation, specifically in schools with salary schedules. Also significant are urban and regional factors, as Lankford, Loeb, and Wyckoff (2002) find using New York State data. Chambers (1996) reveals that gender, race, school level, class size, college major, and additional time spent on school-related activities are also significant determinants of teacher salaries.

Of particular interest are factors that are, at least potentially, related to teacher quality. Such factors may include education, experience, training, principal assessments, and student performance. For instance, Figlio (1997) finds that public school teacher salaries in local labor markets are positively related to two measures of quality: the selectivity of the college where a teacher earned his or her bachelor's degree and subject matter expertise.

Many of the existing analyses focus on public schools, for obvious reasons. However, private schools provide an appropriate focus for examining the relationship between teacher pay and performance.[2] For instance, Ballou and Podgursky (1998) state that, in addition to retaining high quality teachers, private schools have more flexible pay, have more supervision and mentoring, employ more non-certified teachers, have more staff development with training and mentoring, and have greater freedom to dismiss bad teachers. While public school salaries are almost entirely determined by education and experience, "the fact that schedule variables are less informative about the compensation of private school faculty suggests that unobserved factors (for example, individual merit) play a greater role in determining salaries" (p. 411).

In addition, controlling for region, education, and experience, Ballou and Podgursky (1997) find that private school teachers earn lower salaries than their counterparts in the public sector. On the other hand, pri-

[2] Also important is variation across private religious schools, a phenomenon only briefly mentioned in much of the literature on private schools. See Hanushek (2002) and Chambers (1996).

vate school teachers appear as good or better, according to indicators such as college selectivity and academic major. Furthermore, while principals rate beginning teachers similarly across sectors, experienced teachers at private schools are rated more highly than those at public schools. Are these teachers indeed of better quality, and do salaries reflect this? Moreover, is the unobserved component of salary (correlated with quality) associated with covariates in intuitive ways?

Data

Data for individual teachers, their schools, and their principals are compiled from the 1990–1991 Schools and Staffing Survey (SASS). This analysis focuses on secondary private schools, as they have more flexible pay structures than public schools. The SASS data are merged with county-level community characteristics including poverty level and median house value from the 1990 School District Data Book (SDDB). While the SDDB reports key variables such as expenditures for public schools, this information is unfortunately unavailable for private schools. After dropping observations that are missing data, the sample includes 2,372 teachers from 1,104 private secondary schools.

The SASS data are merged with county-level community characteristics including poverty level and median house value from the 1990 School District Data Book (SDDB).

Summary statistics are listed in table 1. Column 3 lists salary differentials that correspond to the indicator variables *(x)*. Specifically, *differential* = $salary_{x=1} - salary_{x=0}$ and *t*-statistics for testing the equality of means appear in column 4. In this study, teacher salary includes annual base salary and additional compensation for evening classes, coaching, and other similar school-related work. This variable is very highly correlated with base salary, but is a superior measure of compensation for teaching activities. Mean teacher salary during the 1990–1991 school year is $20,471. Salary differences arising from qualifications are mostly as expected. First, average teaching experience in the sample is 12 years, and 9 percent of the teachers have less than 1 year of experience. There appear to be substantial penalties for very little experience, as well as distinct returns to schooling. Teacher training is positively correlated with salary, as is subject matter expertise. Nearly half of these teachers are not state certified, revealing the hiring flexibility that has been previously documented. Furthermore, it appears that certification does not significantly impact pay, as evidenced by *t*-test results.

Nine percent of teachers are classified as "contributed-service," meaning that they accept a lower salary, often as a member of a religious order. This is consistent with a nonprofit motive and employees donating labor in order to benefit the mission of the school. Indeed, contributed-service teachers earn 18 percent lower salaries, on average. This is likely tied to salary differences across school affiliation. As seen in the second panel of table 1, there appears to be substantial variation across Catholic, other religious, and nonsectarian schools. For instance, teachers at Catholic parochial schools and those at conservative Christian schools earn quite low salaries. On the other hand, teachers at private order Catholic schools appear to earn the highest salaries (28 percent higher on average). In addition, only 12 percent of private school teachers in the sample work at nonsectarian schools, and they earn on average $5,231 more than teachers at other private schools.

Descriptive statistics also highlight the presence of evaluation and incentive programs. The 10 percent of teachers who are designated as masters earn 26 percent higher salaries.[3] Merit pay programs also coincide with higher average salaries for teachers within the school. Furthermore, teachers who receive a merit bonus earn 22 percent higher salaries, on average. Nonetheless, these differentials could be the result of school or community characteristics that are necessarily excluded from such simple tests. Salary differences that control for these characteristics are estimated in regression analyses in the next section, "The Determinants of Individual Teacher Salaries."

[3] The "master" teacher designation is an NCES data convention. These teachers are so determined by their individual schools or districts, and generally are mentors to younger teachers, aiding in their development in the crucial first years of teaching.

Table 1. Descriptive statistics for private secondary school teachers: 1990–91

Variable	Mean	Standard deviation	Salary Differential[1]	Absolute *t*-statistic[2]
Teacher characteristics				
Salary (in dollars)	20470.38	8307.89	†	†
Years of experience	12.410	9.936	†	†
Less than 1 year of experience	0.088	0.283	–5699.07	9.631
Bachelor's degree	0.582	0.493	–4020.41	11.968
Master's degree	0.330	0.470	4143.33	11.747
Professional/doctoral degree	0.059	0.237	4817.39	6.740
Male	0.414	0.493	3537.25	10.444
Part-time teacher	0.194	0.395	–7641.75	18.993
Contributed service[3]	0.093	0.290	–3722.90	6.384
State certified in field	0.588	0.492	490.35	1.415
Master teacher[4]	0.097	0.297	5328.62	9.432
Career ladder	0.238	0.427	2203.62	5.537
Receive merit pay	0.052	0.223	4492.99	5.904
Subject same as major	0.549	0.498	1322.00	3.868
Education training in college	0.899	0.301	1409.07	2.489
Training seminar of more than 30 hours	0.459	0.498	2235.09	6.588
Hours required to be at school	32.509	9.450	†	†
Teachers distributed across schools, school characteristics				
Catholic—parochial	0.165	0.371	–4338.17	9.627
Catholic—diocese	0.160	0.367	–1441.20	3.105
Catholic—private order	0.118	0.322	5821.61	11.284
Conservative Christian	0.055	0.228	–5238.38	7.087
Other religious affiliated	0.309	0.462	–175.64	0.476
Religious unaffiliated	0.069	0.254	–1046.72	1.557
Nonsectarian school	0.124	0.329	5230.96	10.312
Urban area	0.476	0.500	1338.27	3.930
Suburban area	0.327	0.469	130.43	0.359
Rural area	0.197	0.398	–2288.80	5.371
Merit pay program	0.126	0.332	3979.25	7.852
Number of observations	2,372	2,372	2,372	2,372
Number of schools	1,104	1,104	1,104	1,104

†Not applicable.

[1]Average salary differential for indicator variables, salary(*variable*=1)–salary(*variable*=0).

[2]Equality of means *t*-statistic.

[3]Indicator variable if teacher works on contributed-service basis, as with a religious order.

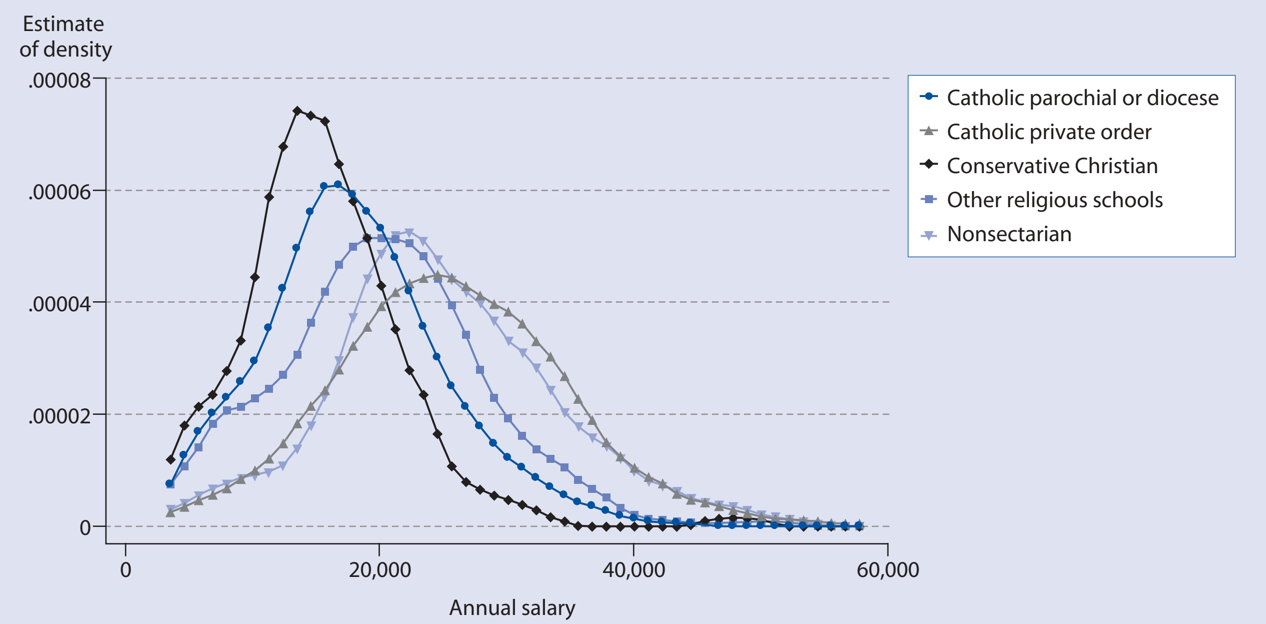
[4]The "master" teacher designation is an NCES data convention. These teachers are so determined by their individual schools or districts, and generally are mentors to younger teachers, aiding in their development in the crucial first years of teaching.

SOURCE: U.S. Department of Education, National Center for Education Statistics: Schools and Staffing Survey (SASS), 1990–1991 and School District Data Book (SDDB), 1990.

Results in table 1 imply that substantial variation in teacher salaries arises from school affiliation. To further explore this variation, figures 1 and 2 display Epanechnikov kernel density estimates and box-and-whisker plots of teacher salary across school affiliation. First, figure 1 provides an estimate of the probability density function for salary, and reveals that teachers at conservative Christian schools earn the lowest (and least flexible) salaries.[4] Salaries at nonsectarian and Catholic private order schools are highest, and exhibit the most variance. This variation could correspond to differences in teacher quality, revealing flexibility in

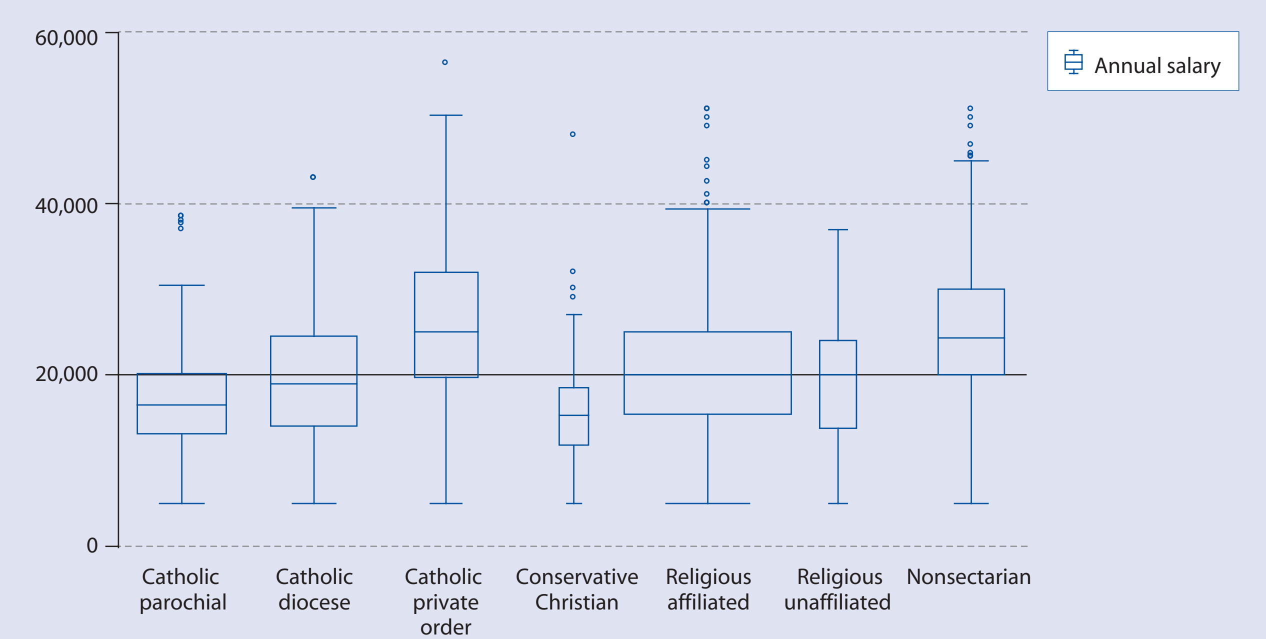
[4] The bandwidth varies across different density estimates, and is determined using the formula $h = 0.9m/n^{1/5}$, where $m = \min(\sqrt{\text{var}}, \text{range}/1.349)$. For instance, the bandwidth for the density estimate of nonsectarian school teacher salaries is 2,142, corresponding to a moving interval of \$4,284.

Figure 1. Teacher salary (in dollars) by school affiliation: 1990–91



SOURCE: U.S. Department of Education, National Center for Education Statistics: Schools and Staffing Survey (SASS), 1990–1991.

Figure 2. Teacher salary (in dollars) by school affiliation: 1990–91



NOTE: The box-and-whisker plots in the figure represent the medians and interquartile ranges, and the circles represent outliers. Outliers are either (1) greater than $x_{\theta=0.75} + 1.5(x_{\theta=0.75} - x_{\theta=0.25})$ or (2) less than $x_{\theta=0.25} - 1.5(x_{\theta=0.75} - x_{\theta=0.25})$ where θ represents the quartile. The horizontal line at 20,000 represents the median salary (in dollars) in the sample, and the width of each box corresponds to the relative size of each affiliation subsample.

SOURCE: U.S. Department of Education, National Center for Education Statistics: Schools and Staffing Survey (SASS), 1990–1991.

rewarding teachers. In contrast, the left tails for many religiously affiliated schools likely reveal the salaries for contributed-service teachers.

Another view of differences in salary distributions is revealed in box-and-whisker plots in figure 2. The boxes represent the medians and interquartile ranges, and the circles represent outliers.[5] The horizontal line at 20,000 represents the median salary (in dollars) in the sample, and the width of each box corresponds to the relative size of each affiliation subsample. Clearly, a large portion of private school teachers are at Catholic schools, but salaries are higher and more varied at Catholic private order schools than at Catholic parochial/diocese schools, suggesting that not all Catholic schools are equal in terms of teacher salaries. In fact, confirming kernel density estimate results, the distribution of teacher salaries at Catholic private order schools seems most like that for nonsectarian schools. Given such considerable differences across school affiliation, the popular use of three categories (Catholic, other religious, and nonsectarian) appears inappropriate, and categories used in this study include Catholic parochial/diocese, Catholic private order, conservative Christian, other religious, and nonsectarian.

> *Given such considerable differences across school affiliation… categories used in this study include Catholic parochial/diocese, Catholic private order, conservative Christian, other religious, and nonsectarian.*

The Determinants of Individual Teacher Salaries

Results in table 1 suggest significant salary differences based upon teacher and school characteristics, and regression analysis generates estimates of the salary determinants that are conditional on observed teacher and school characteristics. To investigate the determinants of private school teacher salaries, the following linear specification is estimated:

(1) $$lsalary = X'\beta + \varepsilon$$

The dependent variable is log annual base salary and X denotes a matrix of teacher, school, and community characteristics. Teacher covariates include experience, degree attainment, hours required to be at school per week, hours spent on after-school activities, number of students in class, training, subject taught, state certification, and receipt of incentive compensation, as well as controls for gender, contributed-service, part-time, and additional responsibilities. School characteristics include affiliation, location, salary schedule indicator, presence of merit pay, and principal's rating of teaching staff (relative to "very good"). The community characteristics of median house value and the percent of the population above the poverty level are merged into the data by county.

Equation (1) is estimated using two distinct techniques: Ordinary Least Squares (OLS) and quantile regressions.[6] OLS provides an adequate baseline for mean effects. Specifically, OLS involves the estimation of the conditional mean $E(y|x) = x'\beta$, yielding the response parameters for the "average" observation, generated by minimizing the sum of squared residuals (i.e., deviation from the predicted salary to the actual salary). In contrast to OLS, quantile regressions involve the estimation of quantiles of the conditional distribution of teacher salary, specifically, $Quant_\theta(y|x) = x'\beta_\theta$ where $\theta \in (0,1)$.[7] $\theta = 0.5$ corresponds to the conditional median, found by minimizing the sum of absolute deviations from the regression line. The median reveals a measure of location that improves upon the mean as it is not skewed by outliers in the data. This is particularly useful whenever the conditional distribution of the dependent variable is fat-tailed, as appears to be the case with private school teacher salaries. An additional benefit of this technique is that it provides estimates of effects that may vary, providing a more thor-

5 Outliers are either (1) greater than $x_{\theta=0.75}$ + 1.5 $(x_{\theta=0.75} - x_{\theta=0.25})$ or (2) less than $x_{\theta=0.25}$ + 1.5 $(x_{\theta=0.75} - x_{\theta=0.25})$, where θ represents the quartile.

6 Quantile regressions has been applied to student achievement studies such as Eide and Showalter (1998) and Levin (2001).

7 Koenker and Hallock (2001) provide an introduction to the quantile regressions technique that is developed in Koenker and Bassett (1978).

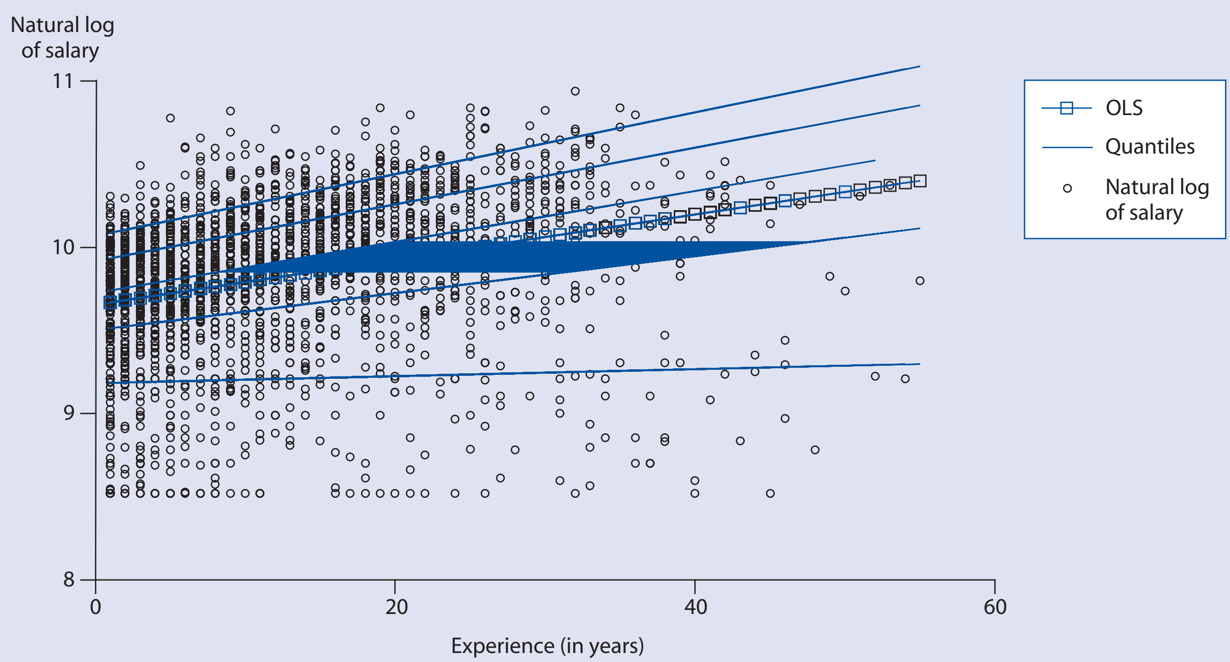
ough depiction of the determinants of variation in the dependent variable. Quantiles other than the median are estimated by differential weighting of positive and negative absolute deviations. For instance, the first decile, $\theta = 0.10$, is predicted where 90 percent of the deviation from the regression line is above the line and 10 percent is below. Note that this does not correspond to partitioning the data and performing a regression on the observations in the lowest 10 percent of the *un*conditional salary distribution.[8]

A simplified version of equation (1) is used to graphically illustrate the value of this technique. Regressing ln(salary) on years of experience results in fitted lines represented in figure 3. Each circle represents an actual experience–ln(salary) pair within the data. The OLS (mean) regression line is denoted by squares, and represents the average effect of an additional year of experience on compensation. Additional lines represent the 90th, 75th, 50th, 25th, and 10th quantile regression coefficients. A common shortcoming of OLS is that estimates can be skewed by outliers. It appears that these mean estimates are so impacted by the observations with high levels of experience and low salaries. Specifically, the median (50th percentile) line lies fully above the mean line, and has a steeper slope. This suggests that the outlier observations generate a downward-biased OLS estimate of the effect of experience on salary.

Additional quantile estimates provide a fuller understanding of the relationship between salary and experience. For instance, increasing quantile lines have greater slopes, revealing that the effect of additional experience

[8] For instance, in examining figure 3, truncating the unconditional distribution would correspond to creating horizontal segments through the scatterplot and then fitting a line through the points in each of the separate segments.

Figure 3. Comparing the impact of experience on salary: Ordinary Least Squares (OLS) and regression quantiles: 1990–91



NOTE: This figure uses a simplified version of $lsalary = X'\beta + \varepsilon$. Regressing ln(salary) on years of experience results in fitted lines represented in the figure. Each circle represents an actual experience–ln(salary) pair within the data. The OLS (mean) regression line is denoted by squares, and represents the average effect of an additional year of experience on compensation. Additional lines represent the 90th, 75th, 50th, 25th, and 10th quantile regression coefficients.

SOURCE: U.S. Department of Education, National Center for Education Statistics: Schools and Staffing Survey (SASS), 1990–1991.

increases throughout the conditional distribution of salary. This suggests that additional experience "explains" many high salaries, indeed contributing to greater salary dispersion, and perhaps has little impact on teachers at low levels in the conditional distribution.

In order to interpret results, it is useful to discuss the conditional distribution of the dependent variable, here teacher salary. A substantial number of controls are included in equation (1), and remaining variance in salary is due to variation in the error term. It is somewhat common in the labor economics literature to interpret the residual as representing unmeasured ability.[9] Within this context, if salary truly reflects performance, the residual can encompass performance that is visible to administrators but not to researchers. Thus, lower conditional salary quantiles would represent relatively "bad" teachers and higher conditional quantiles reflect relatively "good" teachers. Estimated coefficients then reveal how performance interacts with the covariates to affect salary. On the other hand, if salaries do not reflect performance, or if additional important regressors are omitted from the specification, the residual does not directly indicate performance. For instance, the error could simply represent luck, perhaps correlated with characteristics such as parental involvement that could affect teacher salary. Nonetheless, the residual does likely contain unmeasured ability/performance, and flexibility in private school teacher pay suggests that ability and pay are, to a substantial degree, correlated.

An additional interpretation of conditional quantiles is nevertheless possible. Perhaps one of the goals of incentive pay in education is to increase variation in salaries, reflecting rewards and penalties based upon performance. Estimated coefficients that increase monotonically through quantiles of the salary distribution reveal a factor that increases salary dispersion, while monotonically decreasing coefficients signal a source of greater equity in salaries.

Perhaps one of the goals of incentive pay in education is to increase variation in salaries, reflecting rewards and penalties based upon performance.

Estimation results for equation (1) appear in table 2. OLS results incorporate SASS teacher weights, robust standard errors, and allow for correlation across teachers within a school. Quantile results are listed for five separate quantiles: θ = 0.10, 0.25, 0.50, 0.75, 0.90. To provide quantile regression results that are robust to heteroskedasticity, reported standard errors are generated from 1000 bootstrapping repetitions.[10]

First, we see substantial differences across school affiliation. Not all Catholic schools are equal, as seen in the positive and significant returns to Catholic private order schools relative to the baseline Catholic parochial/diocese schools. However, OLS results appear to overstate this somewhat, as quantile regression coefficients decline through the conditional salary distribution, suggesting that employment at a Catholic private order school mostly alleviates a low salary. Stated differently, the premium to such employment is not uniform, and is greatest for teachers earning the lowest (conditional) salaries. Interestingly, this premium seems quite similar to that for teachers at nonsectarian schools. Not surprisingly, employment at a conservative Christian school coincides with significantly lower salaries throughout the distribution.

The community in which the school is located is also a significant salary predictor, and the effect of poverty appears quite diverse. Though the mean impact is significant (0.4 percent higher salaries with 1 percent more population above the poverty line), it is not a consistent effect. High-paid and low-paid teachers appear unaffected by this measure, suggesting there may be very little variation in poverty level at these schools.

As expected, experience increases salary by approximately 1 percent per year, and there are substantial wage penalties in the first year. In addition, the effect of experience on salary increases monotonically throughout the conditional distribution of salary, suggesting that the returns to experience are greatest for

---

9 See for instance, Schultz and Mwabu (1998) and Arias, Hallock, and Sosa-Escudero (2001).

10 To obtain estimated standard errors that improve upon those in Koenker and Bassett (1982), bootstrap replications are used, as per Efron and Tibshirani (1993).

Table 2. Ordinary Least Squares (OLS) and quantile regressions results for natural log of teacher salary: 1990–91

		Quantile				
Variable	**OLS**	**0.10**	**0.25**	**0.50**	**0.75**	**0.90**
Catholic school, private order	**0.263** (0.032)	**0.218** (0.033)	**0.219** (0.028)	**0.206** (0.022)	**0.204** (0.026)	**0.187** (0.025)
Conservative Christian school	**–0.126** (0.040)	**–0.086** (0.038)	**–0.138** (0.033)	**–0.120** (0.034)	**–0.101** (0.032)	**–0.079** (0.044)
Other religious school	**0.106** (0.029)	**0.057** (0.026)	**0.100** (0.024)	**0.115** (0.021)	**0.119** (0.019)	**0.138** (0.022)
Nonsectarian school	**0.242** (0.032)	**0.229** (0.040)	**0.206** (0.032)	**0.270** (0.025)	**0.238** (0.025)	**0.214** (0.028)
Median house value (in thousands of dollars)	**0.001** (0.0002)	**0.001** (0.0001)	**0.001** (0.0001)	**0.001** (0.0001)	**0.001** (0.0001)	**0.001**(0.0001)
Percent above poverty level	**0.393** (0.166)	0.185 (0.201)	**0.307** (0.147)	**0.303** (0.132)	**0.220** (0.128)	–0.023 (0.173)
Years of experience	**0.009** (0.001)	**0.006** (0.001)	**0.010** (0.001)	**0.012** (0.001)	**0.013** (0.001)	**0.013** (0.001)
Less than 1 year of experience	**–0.082** (0.038)	**–0.100** (0.030)	**–0.066** (0.027)	**–0.099** (0.024)	**–0.095** (0.026)	**–0.144** (0.027)
Bachelor's degree	**0.204** (0.071)	**0.187** (0.059)	**0.192** (0.054)	**0.241** (0.066)	**0.163** (0.095)	0.129 (0.066)
Master's degree	**0.293** (0.072)	**0.219** (0.063)	**0.248** (0.055)	**0.329** (0.066)	**0.245** (0.096)	**0.232** (0.067)
Professional/doctoral degree	**0.327** (0.083)	**0.276** (0.075)	**0.296** (0.066)	**0.358** (0.075)	**0.308** (0.100)	**0.273** (0.073)
Hours required per week	**0.013** (0.002)	**0.020** (0.003)	**0.017** (0.002)	**0.013** (0.002)	**0.007** (0.002)	**0.006** (0.002)
Hours per week after school	**0.003** (0.001)	**0.003** (0.001)	**0.002** (0.001)	**0.002** (0.001)	**0.002** (0.001)	0.001 (0.001)
Number of students in class divided by 100	0.008 (0.015)	0.012 (0.021)	**0.026** (0.014)	**0.029** (0.014)	**0.029** (0.014)	**0.051** (0.016)
College courses in teaching	–0.012 (0.032)	0.021 (0.039)	–0.031 (0.028)	–0.001 (0.025)	0.007 (0.023)	0.005 (0.025)
Workshop in teaching methods	0.023 (0.017)	0.012 (0.021)	0.018 (0.017)	0.020 (0.015)	**0.031** (0.015)	0.019 (0.015)
Teach mathematics	**0.041** (0.025)	0.033 (0.036)	**0.044** (0.022)	0.022 (0.022)	–0.008 (0.020)	0.018 (0.025)
Teach science	**0.056** (0.031)	0.055 (0.037)	**0.050** (0.026)	**0.017** (0.024)	0.014 (0.021)	**0.050** (0.029)
Teach English	**0.043** (0.026)	–0.011 (0.040)	**0.060** (0.026)	**0.053** (0.022)	**0.041** (0.022)	**0.043** (0.019)
Teach same as major	**–0.029** (0.017)	0.015 (0.024)	0.007 (0.017)	0.004 (0.015)	–0.012 (0.015)	–0.012 (0.015)
School requires private certification	0.011 (0.023)	–0.009 (0.027)	0.009 (0.020)	0.025 (0.016)	0.001 (0.017)	–0.008 (0.017)
State certified	0.000 (0.019)	–0.016 (0.021)	0.002 (0.018)	0.003 (0.018)	0.003 (0.018)	–0.006 (0.016)
School has merit pay	**0.057** (0.025)	0.056 (0.037)	**0.099** (0.023)	**0.048** (0.019)	**0.082** (0.022)	**0.081** (0.026)
Master teacher[1]	**0.082** (0.029)	**0.089** (0.047)	**0.070** (0.023)	**0.044** (0.023)	0.031 (0.021)	–0.001 (0.025)
Receive merit bonus	**0.106** (0.037)	0.033 (0.045)	0.044 (0.042)	0.030 (0.029)	0.037 (0.032)	0.0004 (0.035)
Receive step on career ladder	**0.036** (0.011)	**0.043** (0.022)	0.014 (0.020)	**0.030** (0.016)	**0.029** (0.017)	**0.034** (0.018)
Principal's rating of teachers overall is "good"	**-0.093** (0.033)	0.023 (0.058)	0.031 (0.037)	–0.024 (0.034)	–0.018 (0.032)	–0.043 (0.028)
Principal's rating of teachers overall is "excellent"	0.003 (0.020)	0.000 (0.022)	**0.033** (0.017)	**0.032** (0.015)	0.020 (0.015)	**0.030** (0.018)
R^2 or pseudo R^2	0.538	0.455	0.395	0.327	0.300	0.302
Number of observations	2,372	2,372	2,372	2,372	2,372	2,372

[1]The "master" teacher designation is an NCES data convention. These teachers are so determined by their individual schools or districts, and generally are mentors to younger teachers, aiding in their development in the crucial first years of teaching.

NOTE: Additional covariates: contributed-service, part-time, region and urban/suburban, male, additional responsibilities, salary schedule, and intercept. Baselines include Catholic parochial or diocese school principal's rating of teaching staff as "very good." Coefficients in bold are significant at the 10 percent level. Standard errors are in parentheses. OLS results incorporate SASS teacher weights and robust standard errors, allowing for correlation across teachers within schools. Quantile regressions results are from 1000 bootstrapping repetitions.

SOURCE: U.S. Department of Education, National Center for Education Statistics: Schools and Staffing Survey (SASS), 1990–1991 and School District Data Book (SDDB), 1990.

the highest performing teachers. As in public schools, experience increases dispersion in private school teacher salaries. The other most commonly cited salary determinant, education, also demonstrates the expected impact on salary, but OLS estimates appear upward-biased for the majority of teachers.

There is also substantial variation in the effect of hours required per week. The OLS coefficient suggests that an additional hour improves salary by 1.3 percent, but this effect declines as salary rises. Thus, longer school days appear to alleviate low teacher salaries, and decrease salary dispersion. Hours spent after school on activities like coaching also improve salary, particularly for what may be lower quality teachers. Also interesting is that class size has no apparent mean impact on salary (coincident with some previous studies), but it does explain some high salaries. In addition, higher quality teachers (with larger positive residuals) seem to benefit from teaching larger classes.

Training in teaching methods appears to have little impact on salary, and is significant in only one specification. On the other hand, subject taught does impact private school salaries. However, mean effects are not confirmed for all teachers. Contrary to expectations, subject matter expertise, measured as teaching the same subject as college major, appears to have no significant impact on salary, according to quantile regression estimates. Perhaps this mirrors the result that neither state nor private certification appears to impact salaries, suggesting that these factors are not rewarded in private schools.

Teachers at schools with a merit pay program earn 6 percent higher salaries.

On the other hand, incentive programs do impact salary. For instance, ceteris paribus, teachers at schools with a merit pay program earn 6 percent higher salaries, on average. Certainly, these results cannot reveal causation (i.e., whether salaries rise when merit pay is introduced, or whether high-paying schools tend to introduce merit pay), but a positive correlation does emerge. In addition, the returns to having a merit pay program are higher (8 percent) for highly paid teachers, suggesting that teachers who are more able do benefit from merit programs. The effect of receiving a higher salary step on a career ladder is similarly significant. On the other hand, a teacher who is designated as a master or mentor teacher receives an 8 percent higher salary on average, but the benefit accrues only to low-paid (potentially poor-performing) teachers. Intuitively, a correlation between experience and master teacher status may explain this, but the inclusion of an interaction term suggests this is not the case. Receiving a merit bonus also greatly improves salary (by 10 percent), but the rewards are not confirmed with regression quantiles. Finally, one might surmise that principal ratings reveal teacher quality. While principals' ratings of individual teachers are unavailable, principals' schoolwide ratings of teachers are provided in the SASS. As expected, the lowest rating corresponds to lower salaries, on average, and the highest rating improves some teachers' salaries.

Conclusion

The main objective of this work is a new measurement of the relationship between teacher performance and pay. As private school salaries exhibit substantial variation and greater use of incentives, such relationships can be better estimated. This is also particularly useful from a policy perspective, as new reforms may look to the private sector for potentially successful accountability methods. Quantile regression estimation provides additional benefits, including an investigation of the factors that increase (or decrease) salary dispersion, as well as the correlation between unobservable salary determinants (e.g., ability and luck) and often-cited teacher qualifications. This study finds that performance and incentives impact pay in many expected ways. Specifically, unobservable performance appears positively correlated with experience and some incentives, thus resulting in higher compensation for high-quality teachers. Employing quantile regressions and studying private schools therefore provides an additional method for examining the relationship between teacher pay and performance.

References

Arias, O., Hallock, K.F., and Sosa-Escudero, W. (2001). Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression Using Twins Data. *Empirical Economics, 26*(1): 7–40.

Ballou, D. (2001). Pay for Performance in Public and Private Schools. *Economics of Education Review, 20*(1): 51–61.

Ballou, D., and Podgursky, M. (1997). *Teacher Pay and Teacher Quality.* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Ballou, D., and Podgursky, M. (1998). Teacher Recruitment and Retention in Public and Private Schools. *Journal of Policy Analysis and Management, 17*(3): 393–417.

Chambers, J. (1996). *The Patterns of Teacher Compensation* (NCES 95–829). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Efron, B., and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap.* New York: Chapman and Hall.

Eide, E., and Showalter, M.H. (1998). The Effect of School Quality on Student Performance: A Quantile Regression Approach. *Economics Letters, 58*(3): 345–350.

Figlio, D.N. (1997). Teacher Salaries and Teacher Quality. *Economics Letters, 55*(2): 267–271.

Hanushek, E.A. (2002, February). *Publicly Provided Education* (Working Paper No. 8799). Cambridge, MA: National Bureau of Economic Research.

Jacobson, S.L. (1989). The Effects of Pay Incentives on Teacher Absenteeism. *The Journal of Human Resources, 24*(2): 280–286.

Koenker, R., and Bassett, G., Jr. (1978). Regression Quantiles. *Econometrica, 46*(1): 33–50.

Koenker, R., and Bassett, G., Jr. (1982). Robust Tests for Heteroscedasticity Based on Regression Quantiles. *Econometrica, 50*(1): 43–62.

Koenker, R., and Hallock, K.F. (2001). Quantile Regression. *Journal of Economic Perspectives, 15*(4): 143–156.

Ladd, H.F. (1999). The Dallas School Accountability and Incentive Program: An Evaluation of Its Impacts on Student Outcomes. *Economics of Education Review, 18*(1): 1–16.

Lankford, H., Loeb, S., and Wyckoff, J. (2002). Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis. *Educational Evaluation and Policy Analysis, 24*(1): 37–62.

Levin, J. (2001). For Whom the Reductions Count: A Quantile Regression Analysis of Class Size and Peer Effects on Scholastic Achievement. *Empirical Economics, 26*(1): 221–246.

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

Schultz, T.P., and Mwabu, G. (1998). Labor Unions and the Distribution of Wages and Employment in South Africa. *Industrial and Labor Relations Review, 51*(4): 680–703.

Solmon, L.C., and Podgursky, M. (2000). *The Pros and Cons of Performance-Based Compensation.* Milken Family Foundation. Available: *http://www.mff.org/publications/publications.taf*

National Evidence on Racial Disparities in School Finance Adequacy

Ross Rubenstein
Syracuse University

About the Author

Ross Rubenstein is an Associate Professor of Public Administration at the Maxwell School of Citizenship and Public Affairs, Syracuse University. He was previously on the faculty of the Andrew Young School of Policy Studies at Georgia State University. His research focuses on education policy, public budgeting, and public finance, specifically funding equity and adequacy in education, performance measurement, and merit-based financial aid. He has conducted training programs in public finance and budgeting for government officials from numerous countries and served as a staff member on Georgia's Education Reform Study Commission. He received his Ph.D. from New York University in 1997.

National Evidence on Racial Disparities in School Finance Adequacy

Ross Rubenstein

Syracuse University

Acknowledgments

I wish to thank William Fowler, Jr. of the National Center for Education Statistics and Catherine Freeman and David Sjoquist of the Fiscal Research Program at Georgia State University for insightful comments on an earlier draft. All responsibility for omissions and errors is mine.

Introduction

The past decade has witnessed a subtle yet fundamental expansion in the focus of school finance policy and research. State school finance structures, often arising in response to legal challenges, have traditionally focused on the provision of equitable educational opportunities for all students. Since Kentucky's 1989 *Rose v. Council for Better Education* (1989) suit, though, interest has increasingly focused on the adequacy of state school finance systems, with courts ruling in favor of plaintiffs challenging state education finance systems in Alabama, Tennessee, North Carolina, South Carolina, Wyoming, and New Hampshire. While equity concerns generally focus on disparities in resources across school districts (or individual schools), adequacy-based legal challenges are more likely to focus on whether educational resources are sufficient to provide students the opportunity to meet state standards or more general educational goals.

The level and adequacy of resources in districts with high proportions of minority students have also figured prominently in a number of school finance lawsuits. For example, in the *Campaign for Fiscal Equity v. New York State* case, a New York State Supreme Court justice found funding in New York City "so deficient that it falls below the constitutional floor set by the education article of the New York State Constitution" (Goodnough 2001). The court went on to state that the system disproportionately harmed minority students, who make up the majority of New York City's public school students.[1]

As adequacy claims have increased in state courts, school finance research on adequacy issues has grown over the past decade. This paper contributes to that body of research by examining school finance adequacy across the United States. Specifically, it quantifies dif-

[1] The Supreme Court Appellate Division overturned the decision in June 2002. As of this writing, the case has been appealed to the New York State Court of Appeals.

ferences in adequacy across states and across racial groups within states, estimates the cost to bring all students to selected adequacy levels, and analyzes adequacy in relation to district racial composition and location. The next section provides conceptual and historical background on school finance adequacy and its relationship to equity concerns, followed by discussion of the data, methods, and empirical results. A final section draws conclusions for policy and future research.

Conceptual Basis of School Finance Adequacy

A large body of research has explored school finance equity within states (see, e.g., Goertz 1992; Hertert, Busch, and Odden 1994; Johnston and Duncombe 1998) and across states (see Berne and Stiefel 1984; Evans, Murray, and Schwab 1997; General Accounting Office 1997; Moser and Rubenstein 2002; Parrish, Hikido, and Fowler 1998; Parrish, Matsumoto, and Fowler 1995; Wyckoff 1992). While equity concerns have been well documented, much less research has examined adequacy, particularly from a cross-state perspective. Equity analyses typically compare school districts to each other, while adequacy analyses measure education funding relative to an absolute standard. At its most basic, an adequate funding level is one that provides all students the opportunity to achieve specified benchmarks and goals. Determining these goals, and understanding the ways in which the inputs to education help students reach these goals, are among the difficult challenges facing policymakers and analysts working to determine adequate funding levels.

An adequate funding level is one that provides all students the opportunity to achieve specified benchmarks and goals.

While the details of state funding systems are typically left to state policymakers, courts are increasingly responding to litigation by defining the broad goals of states' education systems. For example, the Kentucky Supreme Court specified seven "capacities" that an adequate education should provide for children, including "oral and written communication skills to enable students to function in a complex and rapidly changing civilization" and "sufficient understanding of governmental processes to enable the student to understand the issues that affect his or her community, state and nation" (*Rose v. Council for Better Education* 1989). Odden and Clune (1998) take a broader and more ambitious approach to adequacy, defining the goal as "high achievement for all students." They note that because certain students and school systems may require higher levels of resources to achieve desired performance goals, an important component of an adequate system would include additional resources for students with special needs. Therefore, the adequate funding level will likely vary according to student and district characteristics.

The measurement of adequacy is more difficult and less well developed than the measurement of equity. While analysts have used numerous dispersion and relationship measures to examine equity (Berne and Stiefel 1984), no generally accepted methods are available to determine adequate funding levels for different types of students. Since the nature of the relationship between educational inputs and outputs is not fully understood, identifying the level of resources that is necessary and sufficient to produce a given level of achievement is particularly challenging. Despite these difficulties, a number of researchers have addressed the issue head-on and attempted to determine adequate funding levels for districts within individual states. Three methods have primarily been used:[2]

1. **A "professional expert" approach.** In this approach, experienced educators and researchers convene to identify preferred instructional strategies for achieving educational goals (Guthrie and Rothstein 1999). The expert groups then estimate the price of the necessary components. Variations on this approach have been used by Chambers and Parrish (1994) to develop their Resource Cost Model, and by Guthrie and

[2] See Rubenstein and Picus (2000) for further discussion of methods to assess adequacy.

Rothstein (1999) to develop estimates of adequate funding in Wyoming.

2. **An empirical "exemplary district" approach.** In this approach, researchers identify districts and/or schools that are representative of the state as a whole and of subgroups within the state, such as high poverty and rural districts (Augenblick 1997). Districts with higher performance and lower spending levels are then identified within each group. The researchers investigate the instructional strategies and expenditure patterns used in the exemplary districts (or schools) to identify the adequate per pupil funding level for each type of district. This approach has been used to develop estimates of adequate funding levels in Ohio, Illinois, and Mississippi.

3. **An econometric approach**. This approach is built on the development of cost functions (Duncombe and Yinger 1997; Reschovsky and Imazeki 1998). Cost functions relating expenditures to various measures of student performance and need are used to construct a "cost index" that measures differences across districts in the resource levels required to produce a given level of student performance. The estimates control for factors that are assumed to be outside the control of the district, such as the mix of students and the cost of hiring teachers, as well as inefficiencies found in some districts.

National research quantifying school finance adequacy (or inadequacy) has been relatively limited to date. Odden and Busch (1998), using the 1991–92 Common Core of Data (CCD) from the National Center for Education Statistics (NCES), estimate the cost of raising all districts in the United States to the median level of per pupil state and local revenues in each state, as well as to the national median. They find that approximately one-third of all districts would require additional revenues to raise spending to the national median, at a total cost of $16.56 billion. Inflating that figure to 1996–97 dollars, they estimate a total cost of $22.3 billion. *Education Week* newspaper, in its yearly *Quality Counts* report, has also attempted to measure adequacy and to grade states on their efforts (Orlofsky and Olson 2001). Using cost-adjusted NCES data, they divide each state's average expenditures by a national benchmark of $7,652[3] to derive a score out of 100. Using this methodology, only West Virginia achieves a score of 100, while Arizona has the lowest score (44) of all states.

Data and Methods

The analyses in this paper examine inter- and intrastate[4] differences in funding adequacy across the United States. All expenditure data come from the CCD for the 1996–97 school year. To exclude atypical districts and those not providing primarily general education services, I exclude very small districts (those with fewer than five students), those not reporting current expenditures, those with over 50 percent of students in special education as indicated by the presence of an Individualized Education Program (IEP), and any districts classified as college-grade, vocational/special education, nonoperating, or educational service agencies. These exclusions result in a total of 14,145 districts in the database.

The analyses in this paper examine inter- and intrastate differences in funding adequacy across the United States.

To account for differences in exogenous costs facing each district, the data were adjusted using the cost of education index created by Chambers (1998). Chambers' Geographical Cost of Education Index (GCEI) uses a hedonic wage model to control for factors outside local districts' control that affect their costs, including amenities that make teaching and other staff positions relatively more or less attractive.

In addition to the cost-of-education adjustments, I weight the enrollment data (fall membership) to account for student needs that may require the spend-

[3] This figure was derived by inflating their 1997 benchmark of $7,000 per pupil. Each state's rating was calculated as its cost-adjusted per-pupil expenditures divided by the benchmark.

[4] The District of Columbia is treated as a state in all comparisons presented in this article.

ing of additional resources. As described earlier, accounting for differences in student needs is a critical component in developing valid estimates of adequate funding levels. Individual student-level data do not currently exist at a national level to facilitate study of each student's resource needs, but it is possible to group students into broad categories that suggest differential resource needs. The most common of these categories are students requiring special education services, students from low-income families, and students with limited English proficiency (LEP). Students with these special needs typically require more intensive resources, such as smaller classes, special adaptive tools, or teachers with special training, to enable them to achieve at desired levels. The amount of additional resources is likely to vary across students, but estimates are available to give a general sense of the additional weights that should be applied to such students. Following Parrish, Matsumoto, and Fowler (1995), I use weights of 1.2 for students from low-income families and for LEP students, and a weight of 2.3 for students in special education. Thus, for example, a student in special education is assumed to require 2.3 times the funding of a student in general education. While the weights are simply an estimate of the additional funding these students require, they provide a more accurate assessment of resource needs than would unweighted data. Weighted per pupil expenditures are then created by dividing total current expenditures by the weighted student count. Because the weighted student count is, by construction, larger than the unweighted count, weighted per pupil expenditures will be lower. Therefore, districts with relatively high proportions of students with special needs but not the associated higher levels of expenditures will have low weighted expenditures per pupil relative to nominal expenditures.

Accounting for differences in student needs is a critical component in developing valid estimates of adequate funding levels.

While no consensus exists about the level of spending required to achieve adequacy for all students, Odden and Picus (2000) have developed a measure—the Odden-Picus Adequacy Index (OPAI)—that quantifies how far a given finance system is from achieving adequacy, assuming an adequate spending level is determined. The index is similar to the McLoone index in that it concentrates on students in districts below a given funding level. While the McLoone index uses a state or district median as the benchmark, the OPAI can be set at any level deemed to be "adequate." Specifically, it is calculated as

$$(OPAI = PCTABOVE_s + [PCTBELOW_s{}^* (EXPBELOW_s/EXPADEQ_s)]$$

where $PCTABOVE_s$ is the percentage of students in state *s* enrolled in districts spending above the adequate level, $PCTBELOW_s$ is the percentage of students in state *s* enrolled in districts spending below the adequate level, $EXPBELOW_s$ is total expenditures in districts spending below the median in state *s*, and $EXPADEQ_s$ is estimated expenditures in state *s* if all districts below the adequate level spent at the adequate level. Note that schools could be substituted for districts. School-level data, in fact, might provide a more accurate assessment of the resources that actually reach students, though such data are rarely available on a large scale (Berne and Stiefel 1994; Rubenstein 1998).

As the object of analysis for the OPAI calculations, I use current expenditures per pupil for elementary and secondary education.[5] The data are weighted to account for student needs and adjusted to reflect cost-of-education differences across districts.

One of the most difficult assumptions inherent in such analyses is the choice of an adequate funding level. As described above, researchers have used a variety of methods to assess adequacy. Odden and Clune (1998) review a number of strategies and suggest that the estimates are often very close to the national spending median. Odden and Busch (1998) examine the per pupil costs of several popular school reform models and conclude that raising spending in all districts to the national median would provide adequate funding to fi-

[5] This variable includes current operating expenditures for instruction, student support services, and "other" current expenditures such as food service. The variable excludes capital expenditures and expenditures for adult education and community services.

nance these reforms. Therefore, the analyses presented below use the national per pupil current expenditure median for 1996–97 (unweighted and unadjusted, as well as weighted and adjusted) as the adequacy benchmark for the calculations. The analyses also compare the percentage of students above and below the adequate level, additional total and per pupil spending required to bring all students up to the adequate level, and the relationship between the adequacy measures, district racial composition, and district location.

Analysis of Adequacy Across States

Table 1 displays mean spending per pupil per state for four current expenditure variables: nominal expenditures (unweighted and unadjusted), expenditures adjusted for cost differentials, expenditures using weighted pupil counts, and expenditures adjusted for student needs (weighted student counts) and cost differentials.[6] Note that in states with above average-costs, such as Alaska, cost-adjusted expenditures are well below nominal expenditures, while the opposite is true in lower cost states such as Alabama and Arkansas. Because the weighted student counts inflate the denominator in the per pupil expenditure calculation, weighted per pupil expenditures are, in all cases, lower than nominal expenditures.

The analyses compare the percentage of students below the adequate level and additional spending required to bring all students up to the adequate level.

Adequacy Using Nominal Expenditures

Table 2 contains adequacy statistics for each state using nominal current expenditures per pupil as the object of analysis. An OPAI of 1.0 indicates that all districts have current expenditures above the national median, which is $5,333 per pupil using the nominal data. Nationally, 6,141 districts spend below the benchmark while 8,004 districts spend above this level, though equal numbers of students attend districts above and below the benchmark.

Eight states have an OPAI of 1.0, while Utah has the lowest value at 0.714. The majority of states have an OPAI of 0.90 or above. Not surprisingly, Southeastern states (Mississippi, Arkansas, Louisiana, and Tennessee) are disproportionately represented in the bottom quintile of states. The remaining low-adequacy states (Utah, Arizona, Idaho, North Dakota, Oklahoma, and New Mexico) are in the western part of the United States. All of the states with an OPAI of 1.0, with the exception of Alaska and Hawaii, are in the Northeast. Thus, the rankings appear to reflect, in large part, traditional regional differences in spending levels.

Table 2 also lists the proportion of students and of districts in each state below the national median. If districts spending below the benchmark tend to be large (often urban) districts, then the proportion of *students* below the benchmark may be much larger than the proportion of *districts* below the benchmark. Most states have similar proportions of students and districts below the adequacy benchmark, but there are several notable exceptions. For example, in Nevada only 23.5 percent of districts spend below the national median, but these districts serve almost 85 percent of the state's students.[7] Conversely, in Ohio, 73.5 percent of the state's districts spend below the benchmark, but these districts serve only 53.7 percent of the state's students, suggesting that the larger districts tend to have higher per pupil spending.

Table 2 also includes estimates of the total and per pupil cost to bring all students up to the adequacy benchmark. The total estimated cost is just below $14 billion. The gaps are concentrated in the largest states, with California and Texas together accounting for over one-quarter of the required additional spending. On a per pupil basis, though, the additional expenditures required in these states amount to $400–$600 for every pupil below the benchmark, as compared to over $1,000 per pupil in the states with the lowest OPAI.

6 All means and medians used in this paper use a pupil level of analysis; that is, the calculations are weighted by the number of pupils per district.

7 Over half of the state's students are in Clark County.

Table 1. Current per pupil expenditure means, by state: 1996–97

State	Number of students	Number of districts	Nominal mean (in dollars)	Cost-adjusted mean (in dollars)	Weighted mean (in dollars)	Cost-adjusted and weighted mean (in dollars)
Alabama	737,386	127	4,642	5,202	3,848	4,311
Alaska	128,143	53	8,276	6,512	6,868	5,401
Arizona	783,543	213	4,410	4,458	3,772	3,810
Arkansas	457,349	311	4,533	5,201	3,886	4,459
California	5,540,189	985	4,964	4,462	4,265	3,833
Colorado	672,634	176	5,194	5,285	4,515	4,596
Connecticut	507,838	166	8,302	7,213	6,846	5,948
Delaware	104,673	16	6,913	6,747	5,871	5,727
District of Columbia	78,648	1	8,048	7,494	6,900	6,425
Florida	2,241,298	67	5,220	5,453	4,301	4,490
Georgia	1,346,761	180	5,317	5,707	4,609	4,946
Hawaii	187,653	1	5,774	5,790	4,976	4,990
Idaho	245,252	112	4,415	4,806	3,798	4,133
Illinois	1,948,372	899	5,707	5,506	4,756	4,583
Indiana	981,546	292	5,946	6,361	4,921	5,263
Iowa	502,941	378	5,312	6,035	4,457	5,063
Kansas	466,368	304	5,556	6,259	4,716	5,311
Kentucky	631,592	176	5,480	6,135	5,310	5,946
Louisiana	808,798	66	4,526	5,071	3,793	4,245
Maine	212,818	223	6,284	6,420	5,210	5,318
Maryland	818,583	24	6,747	6,605	5,699	5,579
Massachusetts	896,555	295	7,126	6,078	5,725	4,882
Michigan	1,671,574	554	6,453	6,338	5,945	5,841
Minnesota	843,812	341	6,134	6,268	5,238	5,352
Mississippi	502,326	149	4,033	4,630	3,337	3,831
Missouri	892,358	522	5,087	5,364	4,350	4,566
Montana	164,337	450	5,398	5,997	4,566	5,073
Nebraska	290,497	609	5,519	6,286	4,587	5,224
Nevada	282,131	17	5,076	5,333	4,344	4,563
New Hampshire	193,524	162	5,999	5,751	5,051	4,842
New Jersey	1,192,039	551	9,265	8,042	8,637	7,498
New Mexico	326,326	88	4,643	5,014	3,805	4,110
New York	2,805,678	691	8,531	7,597	7,159	6,377
North Carolina	1,208,695	117	4,935	5,380	4,136	4,506
North Dakota	118,170	232	4,667	5,506	4,001	4,718
Ohio	1,844,245	611	5,528	5,572	5,116	5,158
Oklahoma	620,179	548	4,618	5,160	3,936	4,400
Oregon	518,164	214	5,858	6,077	4,997	5,183
Pennsylvania	1,781,383	500	6,490	6,311	5,571	5,415
Rhode Island	150,433	36	7,425	6,746	5,936	5,396
South Carolina	641,925	91	5,066	5,596	4,256	4,699
South Dakota	135,601	173	4,641	5,468	3,978	4,687
Tennessee	886,517	138	4,612	5,048	3,780	4,134
Texas	3,826,366	1043	5,073	5,418	4,215	4,496
Utah	479,812	40	3,826	4,018	3,271	3,435
Vermont	100,277	246	6,385	6,463	5,548	5,614
Virginia	1,096,279	132	5,663	5,821	4,731	4,862
Washington	974,504	296	5,651	5,468	4,828	4,668
West Virginia	303,441	55	6,031	6,736	4,865	5,431
Wisconsin	878,283	425	6,721	7,029	5,651	5,910
Wyoming	98,777	49	5,982	6,553	5,068	5,550

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD); and author's calculations.

Table 2. Adequacy estimates, by state: Nominal 1997 expenditures

Rank	State	Odden-Picus Adequacy Index Index (OPAI)	Number of districts	Percent of districts below adequacy benchmark	Percent of students above adequacy benchmark	Percent of students below adequacy benchmark	Additional funds for adequacy (in dollars)	Additional funds per pupil for adequacy (in dollars)
Total additional adequacy funds							**13,984,553,164**	
1	Alaska	1.000	53	0.0	100.0	0.0	—	—
1	Connecticut	1.000	166	0.0	100.0	0.0	—	—
1	District of Columbia	1.000	1	0.0	100.0	0.0	—	—
1	Delaware	1.000	16	0.0	100.0	0.0	—	—
1	Hawaii	1.000	1	0.0	100.0	0.0	—	—
1	Maryland	1.000	24	0.0	100.0	0.0	—	—
1	New York	1.000	691	0.0	100.0	0.0	—	—
1	Rhode Island	1.000	36	0.0	100.0	0.0	—	—
9	New Jersey	1.000	551	0.4	99.9	0.1	138,534	99
10	West Virginia	1.000	55	1.8	98.8	1.2	619,159	176
11	Massachusetts	0.999	295	5.1	98.2	1.8	4,593,273	284
12	Wisconsin	0.999	425	4.5	97.8	2.2	5,490,954	287
13	Pennsylvania	0.998	500	7.4	94.8	5.2	22,720,045	247
14	Maine	0.996	223	7.6	90.1	9.9	4,433,314	211
15	Michigan	0.994	554	23.1	88.0	12.0	49,285,729	245
16	Washington	0.993	296	22.0	79.6	20.4	36,780,670	185
17	Minnesota	0.992	341	24.3	78.9	21.1	35,616,025	200
18	Wyoming	0.992	49	12.2	73.5	26.5	4,182,601	160
19	Oregon	0.990	214	12.1	82.3	17.7	26,308,790	287
20	Vermont	0.988	246	21.1	79.5	20.5	6,393,490	311
21	Indiana	0.987	292	38.0	73.4	26.6	70,076,221	268
22	New Hampshire	0.980	162	16.7	79.2	20.8	20,468,578	508
23	Kentucky	0.966	176	51.7	52.3	47.7	115,908,140	384
24	Iowa	0.965	378	59.3	44.6	55.4	92,824,804	333
25	Virginia	0.965	132	51.5	51.0	49.0	203,739,657	379
26	Kansas	0.964	304	23.7	67.2	32.8	88,975,417	582
27	Nebraska	0.963	609	37.8	59.0	41.0	57,110,297	479
28	Florida	0.957	67	65.7	38.1	61.9	515,900,579	372
29	Georgia	0.955	180	70.0	36.7	63.3	321,392,637	377
30	Colorado	0.949	176	46.0	25.4	74.6	184,628,800	368
31	Ohio	0.942	611	73.5	46.3	53.7	566,708,597	572
32	Illinois	0.940	899	65.9	55.7	44.3	624,254,518	723
33	Nevada	0.936	17	23.5	15.4	84.6	96,509,122	404
34	Texas	0.930	1043	42.9	18.8	81.2	1,427,761,391	460
35	South Carolina	0.927	91	69.2	30.0	70.0	250,269,465	557
36	California	0.917	985	73.6	23.7	76.3	2,447,360,067	579
37	North Carolina	0.917	117	70.9	17.5	82.5	537,318,910	539
38	Montana	0.913	450	42.7	34.7	65.3	76,114,587	710
39	Missouri	0.895	522	77.0	26.0	74.0	498,186,441	754
40	Alabama	0.866	127	89.8	10.4	89.6	527,544,348	798
41	South Dakota	0.856	173	74.6	8.3	91.7	104,228,229	838
42	Tennessee	0.855	138	91.3	15.0	85.0	683,384,043	906
43	New Mexico	0.855	88	46.6	8.8	91.2	251,609,125	845
44	Oklahoma	0.854	548	67.9	7.5	92.5	481,450,831	839
45	Louisiana	0.846	66	92.4	3.1	96.9	665,440,720	849
46	North Dakota	0.845	232	53.4	14.0	86.0	97,529,110	959
47	Arkansas	0.840	311	91.6	13.8	86.2	389,154,550	987
48	Idaho	0.819	112	66.1	8.4	91.6	236,881,447	1,054
49	Arizona	0.815	213	71.4	7.8	92.2	771,831,622	1,069
50	Mississippi	0.756	149	99.3	0.1	99.9	652,861,661	1,300
51	Utah	0.714	40	82.5	1.4	98.6	730,566,666	1,545

—Not available.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD); and author's calculations.

The $14 billion estimate is somewhat lower than Odden and Busch's (1998) estimate of $16.56 billion in additional required state and local revenues, using 1991–92 data. The amount of additional expenditures required is very sensitive to the choice of adequacy level, however. For example, modestly increasing the adequate expenditure level to $6,000 per pupil more than doubles the amount of additional expenditures required to over $32 billion (table 3).

Table 3. Additional cost to bring all districts to selected per pupil expenditure levels

Per pupil expenditure level (in dollars)	Additional cost (in billions of dollars)
5,000	7.496
5,333 (national median)	13.985
6,000	32.494
7,000	67.651

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD); and author's calculations.

Adequacy Using Cost- and Need-Adjusted Data

Table 4 presents the same information using need-weighted, cost-adjusted expenditures as the object of analysis. The median national expenditure level is $4,657. This lower expenditure level is the result of using a student count inflated by the student weightings. This figure implies that while $4,657 is adequate for a student without special needs, a student from a low-income family or with LEP would require $5,588, and a student in special education would require $10,711. The bottom row shows that when student needs and differential costs are taken into account, the total additional expenditures needed to raise all students to the adequacy benchmark rise to $15.6 billion. For comparability, the required additional expenditures are listed in nominal rather than cost-adjusted dollars.

The number of states with all students above the national benchmark falls from eight to six, with Alaska, Connecticut, and Rhode Island falling below 1.0, and Wyoming joining the list. The OPAI for traditionally high-spending states such as Connecticut and New Jersey falls just below 1.0 once student needs and the higher costs in these states are taken into account, with one or two districts falling below the benchmark. While nominal spending shows all students in Alaska above the benchmark, the cost-adjusted dollars suggest that, with substantially above-average costs, over one-third of students in Alaska receive average real resources below the national median. Similarly, some states with relatively lower nominal spending but below-average costs, such as South Carolina and Wyoming, have substantially higher OPAI values after factoring in these cost and need differences.

Table 4 also presents each state's proportion of students from low-income families, with LEP, and in special education. California, which has higher than average costs and serves large numbers of students with LEP, falls to near the bottom of the pack once need and cost differences are taken into account. Of the over $15 billion in additional required expenditures nationally, almost 40 percent ($6.18 billion) would be in California, with Texas accounting for the next largest share at $1.42 billion. Only Utah, though, would require additional expenditures over $1,000 for each pupil below the national benchmark.

Using the weighted, adjusted data, most states have a higher proportion of *students* below the adequacy benchmark than *districts* below the benchmark. In Alaska, for example, only one district has average expenditures below the benchmark, but that district (Anchorage) serves over one-third of the state's students. In California, 74 percent of the districts have average expenditures below the benchmark, but these districts serve almost all the students in the state (97.3 percent). This pattern (using the cost-adjusted and need-weighted data) is not surprising since large urban districts may have higher costs and serve disproportionately high proportions of students with special needs.

Adequacy and Race

Table 5 displays the percentage of African American and minority students by state, along with each state's OPAI value and rank. While African American students constitute the largest minority group in most states, several states have large proportions of Hispanic, Asian and Pacific Islander students. For example, Texas, New Mexico, and California have large

Table 4. Adequacy estimates, by state: Cost- and need-adjusted 1997 expenditures (median = $4,657)

Rank	State	Odden-Picus Adequacy Index (OPAI)	Percentage of districts below adequacy benchmark	Percentage of students below adequacy benchmark	Additional funds for adequacy (in dollars)	Additional funds per pupil for adequacy (in dollars)	Percent of low-income students	Percent of limited English proficient (LEP) students	Percent of special education students
Total additional adequacy funds					**15,608,516,021**				
1	District of Columbia	1.000	0.0	0.0	—	—	25.4	2.5	8.5
1	Delaware	1.000	0.0	0.0	—	—	11.4	0.9	11.9
1	Hawaii	1.000	0.0	0.0	—	—	19.1	5.7	8.5
1	Maryland	1.000	0.0	0.0	—	—	9.8	1.1	12.7
1	New York	1.000	0.0	0.0	—	—	18.0	0.9	12.2
1	Wyoming	1.000	0.0	0.0	—	—	12.5	0.4	11.8
7	Connecticut	1.000	0.6	0.3	102,284	55	9.7	1.6	14.6
8	New Jersey	1.000	0.4	0.1	206,827	162	6.4	1.4	4.5
9	Kentucky*	1.000	0.6	0.2	87,149	70	15.7	0.3	—
10	West Virginia	1.000	1.8	1.2	52,975	12	19.4	0.4	15.5
11	Michigan	1.000	2.7	0.9	2,442,020	153	16.3	0.6	4.0
12	Wisconsin	1.000	1.4	0.7	1,418,226	187	12.9	0.8	12.5
13	Rhode Island	0.997	8.3	8.6	2,940,390	182	11.7	2.1	17.2
14	Pennsylvania	0.997	9.4	19.5	33,239,796	82	13.7	0.8	10.6
15	Indiana	0.994	14.7	13.1	29,033,977	187	12.3	0.6	14.0
16	Iowa	0.992	6.3	11.6	20,068,648	287	12.1	0.6	12.9
17	Maine	0.991	10.3	17.0	10,622,805	244	11.5	0.4	14.0
18	Minnesota	0.990	8.8	14.3	45,938,243	326	7.4	0.8	12.3
19	Oregon	0.990	9.8	22.9	28,587,527	206	13.7	1.2	11.0
20	Kansas	0.988	9.5	16.7	28,408,603	310	11.9	0.7	11.7
21	Vermont	0.986	14.2	18.1	7,378,302	352	10.2	0.2	10.2
22	Georgia	0.986	16.7	29.0	96,823,811	215	10.4	0.5	10.3
23	Ohio	0.985	34.4	27.9	138,980,411	250	15.4	0.5	3.7
24	Nebraska	0.979	13.6	42.8	32,335,524	216	11.5	0.5	13.9
25	Virginia	0.971	34.1	42.0	172,830,681	314	12.3	0.9	13.1
26	South Carolina	0.967	44.0	55.0	107,988,867	257	18.2	0.4	11.7
27	Washington	0.964	32.1	56.9	209,239,518	323	13.2	1.4	10.9
28	Nevada	0.961	17.6	83.8	57,044,860	207	13.3	1.7	10.6
29	Massachusetts	0.960	47.8	45.1	241,480,916	480	10.9	1.8	16.7
30	Alaska	0.958	1.9	37.7	38,748,558	666	10.5	1.1	13.8
31	Colorado	0.951	27.3	68.4	176,488,929	333	10.8	0.8	9.9
32	North Carolina	0.951	49.6	70.2	306,389,105	303	14.8	0.8	12.5
33	Florida	0.950	52.2	74.0	605,525,236	301	18.4	1.9	13.4
34	South Dakota	0.949	30.6	68.4	33,452,166	310	12.0	0.3	10.9
35	New Hampshire	0.945	32.7	50.0	64,062,938	557	6.7	0.5	13.4
36	Montana	0.935	29.6	55.0	54,241,456	507	17.5	0.3	11.4
37	Texas	0.932	28.3	75.2	1,418,289,236	410	21.2	3.6	11.8
38	Arkansas	0.931	57.2	70.7	149,103,855	395	14.8	0.3	10.5
39	North Dakota	0.926	25.0	60.9	41,388,528	494	14.3	0.2	10.5
40	Missouri	0.924	55.9	68.3	345,823,182	482	15.8	0.5	11.1
41	Alabama	0.915	73.2	81.4	314,831,049	435	18.0	0.4	13.1
42	Illinois	0.909	62.1	72.0	1,029,321,166	610	14.7	1.9	11.5
43	Oklahoma	0.904	35.4	74.9	305,847,267	562	9.1	0.4	11.9
44	Louisiana	0.903	84.8	93.8	387,394,892	427	25.4	0.6	11.1
45	Tennessee	0.878	87.0	85.1	555,860,598	604	19.0	0.5	14.0
46	Idaho	0.867	51.8	86.2	164,365,222	669	14.2	1.0	10.2
47	New Mexico	0.860	38.6	89.5	242,857,000	682	17.9	3.0	13.8
48	Mississippi	0.821	94.6	98.2	442,496,574	742	18.7	0.2	13.2
49	California	0.819	73.8	97.3	6,181,773,959	985	13.9	6.0	9.7
50	Arizona	0.808	68.1	93.2	821,758,833	962	18.7	3.1	9.7
51	Utah	0.735	82.5	98.5	661,243,912	1,199	11.2	0.7	11.1

—Not available.

*Special education data are not available for Kentucky.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD); and author's calculations.

Table 5. State adequacy rankings and racial composition

Rank	State	Odden-Picus Adequacy Index (OPAI)	Percent African American students	Percent African American students in districts spending below adequacy benchmark	Percent African American students in districts spending above adequacy benchmark	Percent minority students	Percent minority students in districts spending below adequacy benchmark	Percent minority students in districts spending above adequacy benchmark
1	District of Columbia	1.000	87.0	—	87.0	96.0	—	96.0
1	Delaware	1.000	30.2	—	30.2	37.0	—	37.0
1	Hawaii	1.000	2.6	—	2.6	78.4	—	78.4
1	Maryland	1.000	36.2	—	36.2	44.2	—	44.2
1	New York	1.000	20.5	—	20.5	44.4	—	44.4
1	Wyoming	1.000	1.1	—	1.1	11.2	—	11.2
7	Connecticut	1.000	13.3	9.7	13.3	28.0	20.3	28.1
8	New Jersey	1.000	18.4	12.7	18.5	38.1	38.1	38.1
9	Kentucky	1.000	10.1	7.5	10.1	11.3	8.6	11.3
10	West Virginia	1.000	4.1	1.2	4.1	4.9	2.1	5.0
11	Michigan	1.000	18.7	0.6	18.9	24.2	5.0	24.4
12	Wisconsin	1.000	9.8	0.3	9.8	17.8	2.4	17.9
13	Rhode Island	0.997	7.3	6.9	7.4	22.4	32.7	21.5
14	Pennsylvania	0.997	14.5	42.6	7.8	20.3	53.7	12.4
15	Indiana	0.994	11.5	1.8	12.9	15.1	4.0	16.8
16	Iowa	0.992	3.5	6.6	3.1	8.2	9.8	8.0
17	Maine	0.991	0.9	0.6	1.0	2.7	1.6	3.0
18	Minnesota	0.990	5.4	1.5	6.1	14.1	5.7	15.5
19	Oregon	0.990	2.8	1.3	3.2	16.2	13.2	17.1
20	Kansas	0.988	8.6	3.5	9.7	18.7	16.6	19.2
21	Vermont	0.986	0.9	0.6	1.0	2.1	0.0	2.6
22	Georgia	0.986	38.4	23.0	44.5	43.3	28.1	49.3
23	Ohio	0.985	15.6	4.4	20.0	18.2	6.0	23.0
24	Nebraska	0.979	6.2	11.8	2.0	13.7	23.1	6.7
25	Virginia	0.971	27.1	32.7	23.1	34.5	36.7	32.9
26	South Carolina	0.967	41.9	37.5	47.4	43.9	39.7	49.2
27	Washington	0.964	5.0	3.9	6.4	23.3	20.6	26.8
28	Nevada	0.961	9.6	11.2	1.0	36.7	39.8	20.3
29	Massachusetts	0.960	8.6	4.2	12.2	22.6	10.9	32.4
30	Alaska	0.958	4.7	8.7	2.4	37.4	33.0	40.0
31	Colorado	0.951	5.6	7.4	2.0	28.8	33.4	19.1
32	North Carolina	0.951	31.0	29.6	34.6	36.8	36.0	38.9
33	Florida	0.950	25.4	24.7	27.6	43.9	36.9	63.8
34	South Dakota	0.949	1.0	1.3	0.3	11.9	9.5	17.1
35	New Hampshire	0.945	1.0	1.2	0.7	3.7	5.1	2.3
36	Montana	0.935	0.6	0.7	0.4	12.8	9.1	17.4
37	Texas	0.932	14.4	14.9	12.8	55.2	56.6	50.9
38	Arkansas	0.931	23.8	17.3	39.5	27.1	21.2	41.4
39	North Dakota	0.926	0.9	1.1	0.5	9.8	8.1	12.4
40	Missouri	0.924	16.6	7.9	34.8	19.2	10.1	38.3
41	Alabama	0.915	36.4	36.2	37.1	38.6	38.3	40.2
42	Illinois	0.909	21.1	24.2	13.2	37.6	42.6	25.0
43	Oklahoma	0.904	10.7	12.3	5.7	31.9	30.5	36.4
44	Louisiana	0.903	46.7	46.9	42.5	49.7	50.1	44.7
45	Tennessee	0.878	23.2	22.0	29.9	22.4	20.4	33.5
46	Idaho*	0.867	—	—	—	—	—	—
47	New Mexico	0.860	2.4	2.6	0.5	62.7	62.5	64.3
48	Mississippi	0.821	51.4	50.8	79.2	52.5	52.0	79.5
49	California	0.819	8.7	8.8	5.7	60.8	61.5	37.0
50	Arizona	0.808	4.3	4.2	5.8	44.4	42.2	75.5
51	Utah	0.735	0.8	0.8	0.1	11.4	11.2	31.1

— Not available.

*Racial composition data are not available for Idaho.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD); and author's calculations.

Hispanic populations, while Hawaii has a large Asian and Pacific Islander population.[8]

Looking at the table, no clear relationship between adequacy and the proportion of African American students is apparent. For example, two states with OPAI values of 1.0 (Hawaii and Wyoming) have relatively low proportions of African American students, while two others (the District of Columbia and Maryland) serve student populations that are over one-third African American. At the other end of the scale, most of the states with the lowest OPAI values (Utah, Arizona, California, and New Mexico) serve a small percentage of African American pupils, though low-ranked Mississippi is over 50 percent African American. With the exception of Utah, though, each of these low-adequacy states has a high proportion of minority group students, primarily Hispanics.

A more systematic analysis also reveals a mixed picture. The Pearson correlation coefficient (pupil-weighted) between the percentage of African American students in a state and its OPAI is 0.164, reflecting a weak positive relationship between adequacy and a state's racial composition. Thus, as the percentage of African American students increases, the state's OPAI also tends to increase. Examining the relationship between the percentage of minority students and adequacy, however, yields a very different result. The correlation between percent minority and OPAI is –0.522, reflecting a strong negative relationship between adequacy and the percentage of a state's students from minority groups. The difference may be explained in large part by several large states (California, Illinois, Texas, and Arizona) with relatively low OPAI values and large numbers of Hispanic students.

Most states with higher proportions of African American students in districts below the national benchmark also have higher proportions of all minority students in these districts.

Statewide averages may mask important intrastate disparities, however. For example, if a state has a high proportion of minority students and a high OPAI, but the districts above the adequate level serve primarily White students, then the relationship between adequacy and student race may be stronger than appears by examining the statewide average. To assess this relationship, table 5 also includes the percentage of African American and minority students in the state as a whole, and in districts above and below the adequacy benchmark. Six states have no districts below the benchmark. Of the remaining 45 states, 8 have well below average proportions of African American students in districts above the adequacy benchmark (Pennsylvania, Nebraska, Virginia, Nevada, Colorado, Illinois, Oklahoma, and Louisiana). All but Louisiana also have above-average percentages of African American students in lower spending districts. In other words, African American students in these states are likely to be in districts spending below the adequacy benchmark.

Most states, though, have proportions of African American students in districts above and below the benchmark that reflect the statewide demographic composition of students. Several states, such as Georgia, South Carolina, Arkansas, and Missouri, have well above average proportions of African American students in higher spending districts, and below-average proportions of African American students in lower spending districts. In Michigan and Wisconsin, where the state proportions of African American students are 19 and 10 percent, respectively, districts above the benchmark have percentages of African American students that reflect state demographics, but the districts below the benchmark serve almost exclusively White student populations.

Examining the spending patterns in relation to the proportion of all minority students (African American, Hispanic, Asian, and Pacific Islander) produces similar results. Most states with higher proportions of African American students in districts below the national benchmark also have higher proportions of all minority students in these districts, though disparities become more pronounced in a limited number of states, such as Texas and Rhode Island. Likewise, most states

[8] The data on student race are aggregated from the school to the district level for the 1996–97 school year. I thank William Fowler of NCES for providing these data.

with a higher proportion of African American students in districts above the benchmark exhibit a similar pattern for all minority students. The differences are even larger in some states, such as Florida, which has a slightly above average proportion of African American students in districts above the benchmark, but a well above average proportion of minority students (64 percent in districts spending above the adequacy benchmark as compared to the state average of 44 percent).

Interestingly, the within-state differences are most pronounced in some of the states with the lowest overall adequacy rankings. In Arizona, for example, districts spending above the national benchmark serve over 75 percent minority children on average, while districts below the benchmark have 42 percent minority children. Similarly, in Utah the districts spending below the benchmark have primarily White student populations (89 percent) while those above the benchmark are only 69 percent White. Because the vast majority of students in these states are in districts spending below the national benchmark, though, the above-benchmark averages include relatively few students.

In most states, minority children are not systematically overrepresented in the lowest spending districts. Minority children, particularly Hispanics, are often heavily concentrated in lower spending states, however.

A small number of states exhibit the opposite pattern. For example, lower spending districts in California tend to have much higher proportions of minority students than do higher spending districts (62 percent in lower spending districts vs. 37 percent in higher spending districts). In Nebraska, almost 14 percent of the state's students are racial minorities, yet districts below the adequacy benchmark average 23 percent and districts above the benchmark average less than 7 percent. Pennsylvania has the most dramatic contrast, with districts spending below the national benchmark averaging 54 percent minority students as compared to 12 percent in districts above the benchmark and just over 20 percent in the state as a whole. Unlike California, most minority students in Pennsylvania are African American. Despite these exceptions, though, most states have similar or lower proportions of African American and minority students in districts below the adequacy benchmark as compared to the state average. These results suggest that African American, Hispanic, and Asian children are not systematically overrepresented in the lowest spending districts in most states. Minority children, particularly Hispanics, are often heavily concentrated in lower spending states, however.

Adequacy by District Location

Given that racial demographics may be closely related to location, examining the relationship between adequacy and district location may also shed some light on these patterns. The CCD contains location descriptors from the U.S. Bureau of the Census categorizing each district in one of seven categories: large central city, urban fringe of large city, mid-size central city, urban fringe of mid-size city, large town, small town, and rural. I combine large central city and mid-size central city into a category called "urban," urban fringe of large city and urban fringe of mid-size city into a category called "urban fringe" and large town, small town, and rural into a category called "rural."

Table 6 displays the percentage of districts above and below the adequacy level falling into each of these three categories. In most states, urban and urban fringe districts are more likely to spend below the benchmark, while rural districts are more likely to spend above the benchmark.[9] For example, California has 727 districts below the benchmark and 258 districts above. Of the districts below the national benchmark, 60 percent are urban fringe and 21 percent are urban. Of those above the benchmark, only 48 percent are urban or urban fringe. Similar patterns are apparent in a number of states (for example, Arizona, Colorado, Florida, Georgia, Nevada, Texas, and Washington). Only in six states is the proportion of rural districts below the benchmark higher than the proportion above the benchmark.

The higher spending in rural districts is somewhat surprising, but may be the result of several factors. Urban and urban fringe districts are likely to have higher costs and may have higher proportions of students with special needs. Therefore, even though nominal spending

9 This pattern ignores states in which only one or two districts fall below the benchmark (e.g., Kentucky and New Jersey).

Table 6. Distribution of districts by location and spending relative to national median of weighted adjusted current expenditures: 1997

State	Below or above national median	Percent: Urban	Percent: Urban fringe	Percent: Rural	Total number of districts
Alabama	Below	8	25	68	**93**
	Above	21	21	59	**34**
Alaska	Below	100	0	0	**1**
	Above	0	0	100	**52**
Arizona	Below	21	35	44	**145**
	Above	8	20	73	**66**
Arkansas	Below	6	10	85	**178**
	Above	3	2	95	**133**
California	Below	21	60	19	**727**
	Above	9	39	52	**258**
Colorado	Below	27	31	42	**48**
	Above	1	8	91	**128**
Connecticut	Below	0	100	0	**1**
	Above	7	54	39	**165**
Delaware	Below	0	0	0	**0**
	Above	19	38	44	**16**
District of Columbia	Below	0	0	0	**0**
	Above	100	0	0	**1**
Florida	Below	34	46	20	**35**
	Above	13	3	84	**32**
Georgia	Below	0	57	43	**30**
	Above	5	10	85	**150**
Hawaii	Below	0	0	0	**0**
	Above	0	100	0	**1**
Idaho	Below	2	7	91	**58**
	Above	0	0	100	**54**
Illinois	Below	5	42	53	**558**
	Above	5	40	55	**341**
Indiana	Below	7	47	47	**43**
	Above	12	23	65	**249**
Iowa	Below	13	25	63	**24**
	Above	3	3	94	**352**
Kansas	Below	10	34	55	**29**
	Above	3	3	94	**275**
Kentucky	Below	0	0	100	**1**
	Above	5	15	79	**175**
Louisiana	Below	16	21	63	**56**
	Above	0	10	90	**10**
Maine	Below	0	13	87	**23**
	Above	3	9	88	**200**
Maryland	Below	0	0	0	**0**
	Above	17	33	50	**24**
Massachusetts	Below	5	74	21	**141**
	Above	8	47	45	**150**
Michigan	Below	7	29	64	**14**
	Above	9	32	59	**539**
Minnesota	Below	0	47	53	**30**
	Above	3	18	79	**304**
Missouri	Below	2	19	78	**292**
	Above	3	8	89	**230**
Mississippi	Below	4	9	88	**141**
	Above	0	0	100	**8**
Montana	Below	8	2	90	**133**
	Above	1	3	96	**315**
Nebraska	Below	4	4	93	**83**
	Above	1	1	98	**524**
New Hampshire	Below	6	26	68	**53**
	Above	1	17	82	**109**
New Jersey	Below	0	100	0	**2**
	Above	3	88	10	**549**
New Mexico	Below	9	12	79	**34**
	Above	2	2	96	**54**
New York	Below	0	0	0	**0**
	Above	6	47	47	**680**
Nevada	Below	67	0	33	**3**
	Above	0	7	93	**14**
North Carolina	Below	19	31	50	**58**
	Above	8	3	88	**59**
North Dakota	Below	9	10	81	**58**
	Above	1	2	98	**172**
Ohio	Below	10	36	54	**210**
	Above	15	39	46	**401**
Oklahoma	Below	10	23	66	**194**
	Above	1	3	96	**350**
Oregon	Below	5	76	19	**21**
	Above	6	21	73	**191**
Pennsylvania	Below	4	53	43	**47**
	Above	10	49	41	**453**
Rhode Island	Below	33	67	0	**3**
	Above	9	64	27	**33**
South Carolina	Below	15	33	53	**40**
	Above	8	16	76	**51**
South Dakota	Below	4	6	91	**53**
	Above	0	0	100	**119**
Tennessee	Below	6	18	76	**120**
	Above	33	17	50	**18**
Texas	Below	25	39	36	**295**
	Above	4	8	87	**748**
Utah	Below	15	15	70	**33**
	Above	0	0	100	**7**
Vermont	Below	0	6	94	**35**
	Above	1	3	96	**211**
Virginia	Below	18	33	49	**45**
	Above	8	20	72	**87**
Washington	Below	16	41	43	**95**
	Above	8	8	84	**201**
West Virginia	Below	0	0	100	**1**
	Above	7	13	80	**54**
Wisconsin	Below	0	33	67	**6**
	Above	5	24	71	**418**
Wyoming	Below	0	0	0	**0**
	Above	4	0	96	**49**

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD); and author's calculations.

may be higher in urban districts, cost and need-adjusted spending could be lower in urban areas than in rural areas. In addition, rural districts tend to be small and unable to take advantage of economies of scale. When fixed district costs (such as administration) are divided by low numbers of pupils, per pupil averages are inflated. In Georgia, for example, rural districts average 3,589 students, as compared with 31,569 in urban districts and 20,222 in urban fringe districts.

Conclusions

This paper provides a starting point for estimating the cost of providing adequate educational resources nationwide and for examining disparities in adequate educational opportunities across racial groups. The analysis does not attempt to determine an adequate funding level for different types of students, but instead uses existing estimates of adequate funding and differential costs to cost out the additional funding needed to achieve adequacy. Several conclusions arise from the analyses:

While interstate differences are largely correlated with the proportion of minority children in the state, minority children within states do not appear to be concentrated in lower spending districts.

- Using the national median of per pupil spending as the estimate of an adequate funding level, additional spending of approximately $14–$16 billion is needed to raise all districts in the country to the national median, an increase of approximately 5 to 6 percent in total current expenditures. This figure is close to—though slightly below—previous estimates.
- The most consistent disparities across states are regional, with northeastern states generally having high levels of adequacy and southeastern states having low levels of adequacy. These differences largely remain even when differences in the cost of education and student needs are taken into account.
- Adequacy index values are only weakly (positively) correlated with the proportion of African American students in a state, but strongly negatively related to the percentage of minority students in a state. This result may be driven in large part by several large states, such as California, Texas, and Arizona, with low OPAI values and high proportions of Hispanic and other minority students.
- Interstate racial disparities in adequacy are generally greater than intrastate disparities. In most states, districts below the national median tend to have lower proportions of African American and minority students than do districts above the median. Only a small number have substantially higher than average proportions of African American and minority students in lower spending districts.
- Using cost- and need-adjusted expenditure data, rural areas tend to be disproportionately represented in districts spending above the median, while urban and urban fringe districts are more likely to be below the median. Lower costs and diseconomies of scale in rural districts may account for much of this pattern.

These results highlight several issues for future policy debates and research. For example, the estimates show that the additional cost of bringing average spending in all districts up to the national median is relatively low, though the resources would need to be heavily targeted to specific states and districts. Using other benchmarks for adequacy substantially changes the estimates, however. As table 3 shows, even raising the bar from $5,333 to $6,000 per pupil more than doubles the additional cost. Achieving a more ambitious goal, such as average spending of $7,000 per pupil, would cost an additional $67 billion, an increase of over 25 percent in national elementary and secondary education expenditures.

The analyses also produce somewhat surprising results regarding racial disparities in adequacy. While interstate differences are largely correlated with the proportion of minority children in the state, minority children within states do not appear to be concentrated in lower spending districts. Therefore, a national strategy to address these inequities may be more effective than state-level strategies. The results also highlight the importance of breaking out data on student race into specific racial categories. This is particularly important in states such as

California and Texas, which serve large (and increasing) numbers of Hispanic and Asian students. But the sensitivity of the estimates to the adequacy benchmark level suggests that more work needs to be done to accurately determine adequate resource levels for different students. In addition, it may not be sufficient to measure adequacy purely in terms of dollars spent. Rather, as a number of researchers have attempted to do, we may need to identify adequacy in terms of the resources (personnel and otherwise) that these dollars purchase. Only then can we hope to ensure that all students have the opportunity to achieve the educational goals set out for them.

References

Augenblick, J. (1997). *Recommendations for a Base Figure and Pupil-Weighted Adjustments to the Base Figure for Use in a New School Finance System in Ohio*. Columbus, OH: Ohio Department of Education.

Berne, R., and Stiefel, L. (1984). *The Measurement of Equity in School Finance*. Baltimore, MD: Johns Hopkins Press.

Berne, R., and Stiefel, L. (1994). Measuring Equity at the School Level: The Finance Perspective. *Education Evaluation and Policy Analysis*, *16*(4): 405–421.

Campaign for Fiscal Equity v. The State of New York, 86 N.Y.2d 307 (N.Y. 1995), 2001 N.Y. Misc. LEXIS 1(NY), *rev'd*, No. 05327, slip op. (App. Div., 1st Dept.), *appeal pending*.

Chambers, J.G., and Parrish, T. (1994). Modeling Resource Costs. In H.J. Walberg and W.S. Barnett (Eds.), *Advances in Educational Productivity, Volume 4* (pp. 7–22). Greenwich, CT: JAI Press.

Chambers, J.G. (1998). *Geographic Variations in Public Schools' Costs* (NCES 98–04). U.S. Department of Education. Washington, DC: National Center for Education Statistics Working Paper.

Duncombe, W., and Yinger, J. (1997). Why Is It So Hard to Help Central City Schools? *Journal of Policy Analysis and Management*, *16*(1): 85–113.

Evans, W.N., Murray, S., and Schwab, R.M. (1997). Schoolhouses, Courthouses, and Statehouses After Serrano. *Journal of Policy Analysis and Management, 16*(1): 10–31.

Goertz, M.E. (1992). School Finance Reform in New Jersey: The Saga Continues. *Journal of Education Finance, 18*(4): 346–365.

General Accounting Office. (1997). *State Efforts to Reduce Funding Gaps Between Poor and Wealthy Districts*. Washington, DC: U.S. Government Printing Office.

Goodnough, A. (2001, January 11). New York City Is Shortchanged in School Aid, State Judge Rules. *New York Times*, p. A1.

Guthrie, J.W., and Rothstein, R. (1999). Enabling Adequacy to Achieve Reality: Translating Adequacy Into State School Finance Distribution Arrangements. In H.F. Ladd, R. Chalk, and J.S. Hansen (Eds.), *Equity and Adequacy in Education Finance: Issues and Perspectives* (pp. 209–259). Washington, DC: National Academy Press.

Hertert, L., Busch, C., and Odden, A.R. (1994). School Financing Inequities Among the States: The Problem from a National Perspective. *Journal of Education Finance, 19*(3): 231–255.

Johnston, J.M., and Duncombe, W. (1998). Balancing Conflicting Policy Objectives: The Case of School Finance Reform. *Public Administration Review*, *58*(2): 145–158.

Moser, M., and Rubenstein, R. (2002). The Equality of Public School District Funding in the U.S., 1992–1995: A National Status Report. *Public Administration Review*, *62*(1): 47–56.

Odden, A.R., and Busch, C. (1998). *Financing Schools for High Performance: Strategies for Improving the Use of Educational Resources*. San Francisco: Jossey-Bass.

Odden, A.R., and Clune, W.H. (1998). School Finance Systems: Aging Structures in Need of Renovation. *Educational Evaluation and Policy Analysis*, *20*(3): 157–177.

Odden, A.R., and Picus, L.O. (2000). *School Finance: A Policy Perspective* (2nd ed.). New York, NY: McGraw Hill Publishers.

Orlofsky, G.F., and Olson, L. (2001, January 11). The State of the States. *Education Week, Quality Counts 2001: A Better Balance,* 86–88.

Parrish, T.B., Hikido, C.S., and Fowler, W.J., Jr. (1998). *Inequalities in Public School District Revenues* (NCES 98–210). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Parrish, T.B., Matsumoto, C.S., and Fowler, W.J., Jr. (1995). *Disparities in Public School District Spending 1989–90: A Multivariate, Student-Weighted Analysis, Adjusted for Differences in Geographic Cost of Living and Student Need*. U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Reschovsky, A., and Imazeki, J. (1998). The Development of School Finance Formulas to Guarantee the Provision of Adequate Education to Low-Income Students. In W.J. Fowler, Jr. (Ed.), *Developments in School Finance, 1997* (NCES 98–212) (pp. 121–148). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Rose v. Council for Better Education, 790 S.W.2d 186 (Ky. 1989).

Rubenstein, R. (1998). Resource Equity in the Chicago Public Schools: A School-Level Approach. *Journal of Education Finance*, *23*(4): 468–489.

Rubenstein, R., and Picus, L.O. (2000). Politics, the Courts and the Economy: Implications for the Future of School Financing. *Proceedings of the 92nd Annual Conference*. National Tax Association, 131–171.

Wyckoff, J.H. (1992). The Intrastate Equality of Public Primary and Secondary Education Resources in the U.S., 1980–1987. *Economics of Education Review*, *11*(1): 19–30.

Competing Perspectives on the Cost of Education

Lori L. Taylor

University of Texas at Austin and Federal Reserve Bank of Dallas

Harrison Keller

University of Texas at Austin

About the Authors

Lori L. Taylor is a Senior Economist and Policy Advisor with the Federal Reserve Bank of Dallas. During a recent leave of absence from the Bank, she served as Principal Researcher on the Texas Cost-of-Education Project. She has published several articles on the measurement of educational outcomes, school district efficiency, and the impact of competition in education markets. Dr. Taylor holds a B.A. in Economics and a B.S. in Business Administration from the University of Kansas and an M.A. and Ph.D. in Economics from the University of Rochester.

Harrison Keller is the Senior Policy Analyst for Education for Speaker Tom Craddick of the Texas House of Representatives. Prior to coming to the Speaker's office, he was Project Director for Education Policy at The University of Texas at Austin's Charles A. Dana Center, where his responsibilities included designing activities to support informed deliberation and decisionmaking about education policy. Dr. Keller has worked as an education policy analyst in both the House and Senate of the Texas Legislature and has taught at Georgetown University, St. Edward's University, and The University of Texas at Austin. Dr. Keller holds a B.A. in Philosophy from the University of Notre Dame and an M.A. and a Ph.D. in Philosophy from Georgetown University.

Competing Perspectives on the Cost of Education

Lori L. Taylor

University of Texas at Austin and Federal Reserve Bank of Dallas

Harrison Keller

University of Texas at Austin

Introduction

As discussions about education finance shift from considerations of fiscal equity to adequacy, researchers and policymakers are paying increasing attention to geographic variations in the costs of education. Unfortunately, there is no consensus about the best approach to measuring geographic cost variations. Each strategy for making cost adjustments to address these variations has certain conceptual strengths and limitations. Moreover, the picture of geographic cost variations can vary considerably across different strategies for making cost adjustments.

In 1999, the 76th Texas Legislature commissioned the Charles A. Dana Center at The University of Texas at Austin to study different approaches to adjusting school district funding to reflect geographic cost variations. The ensuing study was the most comprehensive of this issue previously attempted in any state, and included researchers from The University of Texas at Austin, Texas A&M University, and the Federal Reserve Bank of Dallas. In this article, after a brief discussion of current theory and practice regarding geographic cost adjustments, we compare and contrast the study's findings about the costs of public education in Texas as well as estimates generated by Jay Chambers and Jennifer Imazeki and Andrew Reschovsky. Notably, we find that different indexing strategies yield considerably different estimates of the costs of education in Texas. As such, we argue that there is a pressing need for greater theoretical guidance about appropriate strategies for cost adjustments. Neither the current strategies nor the estimates they generate should be applied lightly.

Geographic Cost Adjustment in Theory and Practice

The literature on strategies for adjusting school district funding to reflect geographic cost variations can be divided into two broad categories—cost-of-living and cost-of-education strategies.

The basic premise of cost-of-living strategies is familiar: areas with relatively higher costs of living have to pay higher salaries to attract school employees, thereby increasing the cost of operating schools and districts. The cost of living therefore acts as a proxy for the cost of education.

There are two basic strategies for estimating variations in the local cost of living. One strategy is to examine the cost of a specified "market basket" of goods and services used by consumers in each community. The total costs of the market basket of consumer goods and services in each community are then compared to illustrate differences in the costs of living. This sort of strategy is used, for example, to create the Consumer Price Index.

A second strategy for estimating geographic variations in the costs of living is the "comparable wage" strategy. Because all types of workers tend to demand higher wages in areas with a higher cost of living, economic theory suggests that systematic regional variations in wages will reflect variations in the cost of living. Therefore, one should be able to approximate the cost of living for educators by observing salaries of comparable workers who are not educators (Rothstein and Smith 1997; Guthrie and Rothstein 1999; Goldhaber 1999; Stoddard 2002).

Regardless of the strategy used, there are a number of advantages to using cost-of-living indexes to capture geographic variations in the costs of education. The principal advantage is that cost-of-living indexes measure costs that are clearly beyond the control of school administrators. In most areas, district officials are unable to manipulate the general labor market, which means that researchers do not have to draw controversial distinctions between controllable and uncontrollable costs. Furthermore, the calculation of a cost-of-living index can be quite straightforward and need not employ sophisticated statistical techniques. While there are still many complex measurement issues involved (Rothstein and Smith 1997; Wynne and Sigalla 1994), either cost-of-living approach produces cost measures that can be compared relatively easily and directly. Finally, a cost-of-living approach is easily understood by policymakers and easily communicated to the public. Variations on cost-of-living approaches have been used to adjust district funding to reflect geographic variations in Florida, Colorado, and Wyoming.

The principal advantage is that cost-of-living indexes measure costs that are clearly beyond the control of school administrators.

Cost-of-living strategies also have a number of limitations. First, high-quality consumer price data can be quite expensive to collect. For example, the state of Florida reports that it spends more than $100,000 per year collecting consumer price data for use in calculation of its cost index. Second, and more significantly, a cost-of-living strategy relies on comparability among market baskets and among workers. If either sort of comparability breaks down, a cost-of-living index then becomes a poor proxy for the cost of hiring educators. For example, if people choose radically different market baskets in one setting than in another, perhaps because in a rural community they grow more of their own food whereas in a city they eat more restaurant meals, then it would be inappropriate to use the same market basket of goods to measure the cost of living in both settings. Similarly, if tastes for goods and services or local amenities differ according to worker types, perhaps because professionals are more susceptible to the lure of city lights than other types of workers, then it would be inappropriate to include all types of workers in a comparable-wage index. Of course, a market-basket index or a comparable-wage index based on an overly small sample of workers or products would be susceptible to large measurement error.

A third limitation of cost-of-living strategies, which pertains only to market-basket indexes, is that they do not reflect local variations in community characteristics such as climate, crime rates, or cultural amenities.[1] Therefore, cost adjustments based on market baskets of consumer goods may overcompensate districts that face high costs of goods and services but that also have a number of amenities that make them desirable places to work (Rothstein and Smith 1997). Finally, on a related note, cost-of-living indexes measure the cost of living in broad labor markets. By design, they do not capture variations in the costs of education within labor markets.[2] Therefore, cost-of-living strategies may generate the same index

[1] To the extent that these factors influence the price of goods and services such as housing and haircuts, they would be partially reflected in a market basket. However, the weights are likely to be inappropriate.

[2] As McMahon (1996) argues, because teachers may live outside the district in which they teach, it would be misleading to construct cost-of-living index values for districts.

value for an advantaged school district as for its disadvantaged crosstown rival.

A different set of strategies for estimating geographic cost variations involves the construction of cost-of-education indexes (CEIs). This set of strategies uses data on district expenditures to estimate either the costs of providing comparable levels of educational services (Chambers 1998) or the costs of producing comparable educational outcomes (Duncombe, Ruggiero, and Yinger 1996; Imazeki and Rechovsky 1999). The former strategy generates an estimate of the additional amount each district would have to spend to operate a typical school—or at least, to hire a typical teacher. Chambers' Teacher Cost Index and Geographic Cost of Education Index are both examples of this approach. The latter strategy generates estimates of how much more or less each district would be predicted to spend to achieve a certain level of educational achievement—frequently, the average level of educational achievement.

Strategies for estimating geographic cost variations use expenditures to estimate either the costs of providing comparable levels of educational services or the costs of producing comparable educational outcomes.

Cost-of-education strategies have a number of attractive features. First, instead of using indirect proxies for education cost differences, as cost-of-living strategies do, they not only directly examine school district expenditures but also use statistical analyses to estimate the costs of providing equivalent levels of educational services or outcomes in particular districts. Cost-of-education strategies can therefore be used to take account of cost variations within labor markets—an option not available with cost-of-living adjustments. Second, for states that already maintain data on educator salaries and district expenditures, it can be much less expensive to construct a CEI than to apply a market-basket approach. Finally, CEIs that measure the costs of achieving educational outcomes can correct both for variations in the prices paid for resources and for the intensity with which those resources must be used. Cost-of-living indexes, on the other hand, only capture price variations.

Cost-of-education indexing strategies also have a number of potential disadvantages. For one, it is impossible to account completely for all relevant controllable and uncontrollable cost factors. For example, important differences in teacher quality or educational outcomes may not be observable in the data (Hanushek 1999; Goldhaber 1999; Alexander et al. 2000). Therefore, estimation bias is always a concern for researchers. In addition, there are good reasons to believe that existing patterns of district expenditure do not always reflect cost-minimizing behavior. For example, McMahon (1996) argues that district officials can manipulate expenditures, while Hanushek (1999) emphasizes the noncompetitive nature of most educational markets. As Rothstein and Smith (1997) rightly point out, CEIs can reward inefficiency by directing additional state aid to districts that spend the most.

The Texas Cost-of-Education Index Study

Texas is an ideal laboratory for examining geographic differences in the costs of public education. There are a large number of school districts and labor markets in the state, and the significant variation in demographics and economic conditions across those areas implies that the cost of education should vary substantially. Texas also maintains richer data on the financing and performance of its schools than any other state, which facilitates the construction of CEIs. Finally, the state has a decades-long history of adjusting its school finance formula to reflect geographic differences in the cost of education. Since 1984, Texas has incorporated some form of a CEI in its finance formula. The Current Texas CEI represents the systematic variation in teacher salaries arising from five uncontrollable factors—district size, district type, the percentage of low income students, the average beginning teacher salary in surrounding districts, and location in a county with a population less than 40,000—holding constant at the mean variations in property wealth per teacher, the total effective tax rate, the graduation rate, the percent minority teaching staff, nonsalary benefit expenditures per pupil and teacher characteristics (years of experience and indicators for whether the teacher has at least a B.A. or teaches at the secondary level).[3] The Current Texas CEI is somewhat dated, however, because it has not been updated since its adoption in 1990.

A number of researchers have estimated CEIs for Texas. Monk and Walker (1991) developed a Teacher Cost Index that was subsequently incorporated into the state's school finance formula as the Current Texas CEI. The Dana Center study (Alexander et al. 2000, 2002) faithfully updated the Texas CEI and then developed a new Teacher Cost Index (the Texas TCI).[4] Chambers used data from the National Center for Education Statistics Schools and Staffing Survey to estimate a nationwide Geographic Cost-of-Education Index (GCEI), which he also applied to Texas school districts (Chambers 1999). More recently, Imazeki and Reschovsky (2002) and Alexander et al. (2000) estimated cost functions from which they developed indexes (denoted as the I&R Cost Function Index and the A&A Cost Function Index, respectively) of the costs of producing average educational performance in Texas.[5] Finally, Alexander et al. (2000) followed a comparable-wage strategy to generate a cost-of-living index for each Texas school district.

Table 1 provides descriptive statistics on these seven Texas cost indexes. To facilitate comparisons, all the indexes have been rescaled so that the least cost Texas district is assigned an index value of one.

Although all these indexes point to substantial variations in the cost of education, they paint very different pictures of Texas. The Teacher Cost Indexes (the Current Texas CEI, the Updated Texas CEI, and the Texas TCI) range from 1 to 1.34, implying that the cost of education in the highest cost school district is no more than 34 percent greater than in the lowest cost school district. The GCEI ranges from 1 to 1.45, implying a somewhat greater range of educational costs. In contrast, both the cost-of-living index (COL Index) and the I&R Cost Function Index imply that the cost of education nearly doubles as one moves from the lowest cost district to the highest cost district. The A&A Cost Function Index shows the greatest range,

3 For districts with average daily attendance between 1,600 and 2,000 students, an adjusted CEI is used. The adjusted CEI = CEI × (1.0 + ((2000 – ADA) × .00014)).

4 Alexander et al. (2000, 2002) developed a series of Teacher Cost Indexes, and found that a comparatively parsimonious model generated index values that were highly correlated with those of a more complete specification. They demonstrated that their models were remarkably insensitive to the inclusion of health insurance benefits in the dependent variable.

They also demonstrated that index values from their models were reasonably stable across time. The discussion here focuses on the 3-year average salary and benefits index. Alexander et al. (2002) used data from the 1997–98, 1998–99 and 1999–2000 school years to calculate average values for the uncontrollable cost factors in each district. These 3-year average values were then multiplied by the estimated coefficients from the parsimonious salary and benefits model described in Alexander et al. (2000) to generate index values.

5 Our thanks to Jennifer Imazeki and Andrew Reschovsky for graciously making their index available.

Table 1. Descriptive statistics of Texas school districts measured by seven cost indexes

Variable	Number of school districts	Mean	Standard deviation	Minimum	Maximum
Current Texas CEI	1,041	1.06	0.03	1.00	1.18
Updated Texas CEI	1,042	1.07	0.04	1.00	1.20
Texas TCI	1,042	1.10	0.05	1.00	1.34
GCEI	1,042	1.20	0.10	1.00	1.45
A&A Cost Function Index	973	1.41	0.26	1.00	2.84
I&R Cost Function Index	879	1.35	0.15	1.00	1.94
COL Index	1,042	1.37	0.26	1.00	1.94

NOTE: All indexes have been rescaled so that the least cost Texas district is assigned an index value of 1.

SOURCE: Current Texas CEI: Monk and Walker (1991), and Texas Education Agency; Updated Texas CEI: Alexander et al. (2000, 2002); Texas TCI: Alexander et al. (2000, 2002); GCEI: Chambers (1999); A&A Cost Function Index: Alexander et al. (2000); I&R Cost Function Index: Imazeki and Reschovsky (2002); COL Index: Alexander et al. (2000).

with the index value for the highest cost district nearly triple the index value for the lowest cost district.[6]

As table 2 illustrates, there is little agreement *across the indexes* about the characteristics of high- and low-cost districts. The price indexes (the Teacher Cost Indexes, the GCEI, and the COL Index) indicate that the highest cost districts tend to be large and urban. Those are common characteristics of *low*-cost areas according to the cost function indexes, however. Expenditures per pupil are high relative to teacher salaries in the districts assigned high index values by the cost function indexes, and low relative to teacher salaries in the districts assigned high index values by the price indexes. For all the price indexes, average expenditures per pupil are higher for low-cost areas than for high-cost areas; both of the cost function indexes appear to suggest that teacher salaries are higher in low-cost areas than in high-cost areas.

There are also substantial differences *within index types*. High-cost areas are generally assumed to have a greater share of limited English proficient students than low-cost areas, but not according to the COL Index. According to the GCEI and the COL Index, *low*-cost districts have a much greater share of economically disadvantaged students than do high-cost districts. In contrast, according to the Teacher Cost Indexes, the share

[6] Both Imazeki and Reschovsky and Alexander et al. estimated their cost functions from data on districts that serve grades K–12. Imazeki and Reschovsky provided index values only for those districts included in their analysis, while Alexander et al. (2000) published cost function index values for all school districts. Given the obvious technological differences, however, Alexander et al. (2000) caution against relying on the cost function to impute index values for school districts that do not have a high school. In this analysis, we treat as missing the cost function index values for districts that do not serve all grades. If they were included, the A&A Cost Function Index would range from 1 to 5.93.

Table 2. Comparing the characteristics of high- and low-cost Texas school districts across seven cost indexes

	Current Texas CEI	Updated Texas CEI	Texas TCI	GCEI	A&A Cost Function Index	I&R Cost Function Index	COL Index
10 percent of districts with highest index values							
Expenditures per pupil (in dollars)	6,484	6,366	6,576	6,489	9,843	7,759	6,644
Average monthly salary for teachers with less than 5 years' experience (in dollars)	3,058	3,131	3,148	3,101	2,682	2,752	2,955
Average daily attendance	16,193	19,880	20,087	15,812	182	1,430	10,270
Economically disadvantaged (in percent)	57.18	45.47	46.54	42.31	56.86	70.94	34.86
Limited English proficient (in percent)	19.80	15.57	14.94	12.73	6.27	14.09	6.07
Miles to major urban area	82	33	42	26	179	137	0
Urban (in percent)	84.30	93.46	99.05	100.00	7.14	21.35	100.00
10 percent of districts with lowest index values							
Expenditures per pupil (in dollars)	6,839	7,575	7,640	9,488	6,191	6,350	8,045
Average monthly salary for teachers with less than 5 years' experience (in dollars)	2,651	2,665	2,641	2,665	2,964	2,954	2,694
Average daily attendance	635	487	305	175	9,078	8,606	749
Economically disadvantaged (in percent)	44.37	43.60	47.03	53.17	22.54	19.32	57.33
Limited English proficient (in percent)	3.10	2.78	3.72	4.79	3.30	2.71	9.43
Miles to major urban area	110	122	114	141	41	37	182
Urban (in percent)	27.67	21.26	37.38	0.88	81.63	92.13	0.00

NOTE: All district characteristics are as of the 1999–2000 school year. All indexes have been rescaled so that the least cost Texas district is assigned an index value of 1.

SOURCE: Current Texas CEI: Monk and Walker (1991), and Texas Education Agency; Updated Texas CEI: Alexander et al. (2000, 2002); Texas TCI: Alexander et al. (2000, 2002); GCEI: Chambers (1999); A&A Cost Function Index: Alexander et al. (2000); I&R Cost Function Index: Imazeki and Reschovsky (2002); COL Index: Alexander et al. (2000).

of economically disadvantaged students is either higher in high-cost districts or insignificantly different between high- and low-cost districts. Low-cost urban districts are virtually unheard of according to the GCEI and the COL Index, while the Teacher Cost Indexes imply that they are relatively common. The average high-cost district is larger than the state median according to the I&R Cost Function Index but much *smaller* than the median according to the A&A Cost Function Index.

Further confirmation of the dramatic differences across metrics can be found in table 3, which presents the Pearson correlations among the index values. The upper right-hand section of the table presents correlation coefficients for urban school districts; the lower left-hand section presents correlation coefficients for rural school districts.

As table 3 illustrates, the Teacher Cost Indexes and the GCEI are reasonably well correlated with one another in urban areas, but much less so in rural areas. The cost function indexes are well correlated with each other in rural areas and urban areas, but either uncorrelated or *negatively* correlated with the price indexes. None of the indexes are highly correlated with the COL Index, in part because the COL Index does not vary within labor markets as the other indexes do.

Table 4 provides another perspective on the differences within indexes between urban and rural areas. The Cur-

Table 3. Pearson correlation coefficients for urban and rural Texas school districts across seven cost indexes

	Urban school districts						
	Current Texas CEI	Updated Texas CEI	Texas TCI	GCEI	A&A Cost Function Index	I&R Cost Function Index	COL Index
Current Texas CEI		0.8148	0.7521	0.6646	−0.1716	0.1183	0.1869
		0.0001	0.0001	0.0001	0.0005	0.0194	0.0001
		429	429	429	404	390	429
Updated Texas CEI	0.6797		0.7967	0.7688	−0.3063	−0.0550	0.4079
	0.0001		0.0001	0.0001	0.0001	0.2783	0.0001
	612		429	429	404	390	429
Texas TCI	0.4500	0.4503		0.8290	−0.4020	−0.0331	0.3646
	0.0001	0.0001		0.0001	0.0001	0.5152	0.0001
	612	613		429	404	390	429
GCEI	0.1943	0.3733	0.3562		−0.4034	−0.0864	0.4930
	0.0001	0.0001	0.0001		0.0001	0.0882	0.0001
	612	613	613		404	390	429
A&A Cost Function Index	0.0378	−0.2664	−0.1358	−0.6693		0.7969	−0.2816
	0.3687	0.0001	0.0012	0.0001		0.0001	0.0001
	568	569	569	569		390	404
I&R Cost Function Index	0.0523	−0.0817	−0.1505	−0.4283	0.7328		−0.2224
	0.2480	0.0711	0.0008	0.0001	0.0001		0.0001
	489	489	489	489	489		390
COL Index	−0.1153	0.0067	−0.1583	0.1376	−0.2039	−0.2115	
	0.0043	0.8681	0.0001	0.0006	0.0001	0.0001	
	612	613	613	613	569	489	
	Rural school districts						

NOTE: Each cell presents Pearson correlation coefficients; Probability > |R| under H_o : Rho = 0; and number of observations. The upper right-hand section of the table presents correlation coefficients for urban school districts; the lower left-hand (shaded) section presents correlation coefficients for rural school districts.

SOURCE: Current Texas CEI: Monk and Walker (1991), and Texas Education Agency; Updated Texas CEI: Alexander et al. (2000, 2002); Texas TCI: Alexander et al. (2000, 2002); GCEI: Chambers (1999); A&A Cost Function Index: Alexander et al. (2000); I&R Cost Function Index: Imazeki and Reschovsky (2002); COL Index: Alexander et al. (2000).

Table 4. Geographic variations in Texas school districts across seven cost indexes

	Current Texas CEI	Updated Texas CEI	Texas TCI	GCEI	A&A Cost Function Index	I&R Cost Function Index	COL Index
Very sparse rural counties							
Mean	1.05	1.06	1.10	1.11	1.69	1.46	1.16
Standard deviation	0.03	0.03	0.02	0.04	0.33	0.14	0.13
Number of districts	185	186	186	186	177	133	186
Other rural counties							
Mean	1.05	1.06	1.07	1.15	1.41	1.38	1.23
Standard deviation	0.03	0.03	0.02	0.04	0.20	0.13	0.11
Number of districts	427	427	427	427	392	356	427
Small urban areas							
Mean	1.06	1.07	1.10	1.26	1.33	1.30	1.43
Standard deviation	0.04	0.04	0.05	0.06	0.20	0.15	0.11
Number of districts	228	228	228	228	211	200	228
Major urban areas							
Mean	1.09	1.12	1.16	1.34	1.26	1.29	1.79
Standard deviation	0.03	0.04	0.05	0.07	0.15	0.14	0.16
Number of districts	201	201	201	201	193	190	201
Mexican border							
Mean	1.09	1.10	1.13	1.23	1.47	1.46	1.27
Standard deviation	0.04	0.04	0.05	0.10	0.24	0.15	0.17
Number of districts	154	155	155	155	143	128	155

NOTE: All indexes have been rescaled so that the least cost Texas district is assigned an index value of 1.

SOURCE: Current Texas CEI: Monk and Walker (1991), and Texas Education Agency; Updated Texas CEI: Alexander et al. (2000, 2002); Texas TCI: Alexander et al. (2000, 2002); GCEI: Chambers (1999); A&A Cost Function Index: Alexander et al. (2000); I&R Cost Function Index: Imazeki and Reschovsky (2002); COL Index: Alexander et al. (2000).

rent Texas CEI and the Updated Texas CEI are much higher for major urban areas, but indicate little difference in cost between rural areas and small urban areas, such as Waco or Texarkana. In contrast, the Texas TCI suggests that costs are higher in sparsely populated rural areas than in some urban areas! Both the GCEI and the COL Index strictly increase with urban density. But the cost function indexes generally *decrease* with density.

The cost function indexes are highest in rural areas for a very simple reason—that's where the small schools are. And as figure 1 illustrates, both of the cost function indexes exhibit striking economies of scale.

Table 5 illustrates another perspective on this issue. According to the A&A Cost Function Index, the average school with less than 100 students has *twice* the index value of the average school with more than 10,000 students.[7] All but one rural school district has fewer than 10,000 students; only two urban K–12 districts have fewer than 100 students.

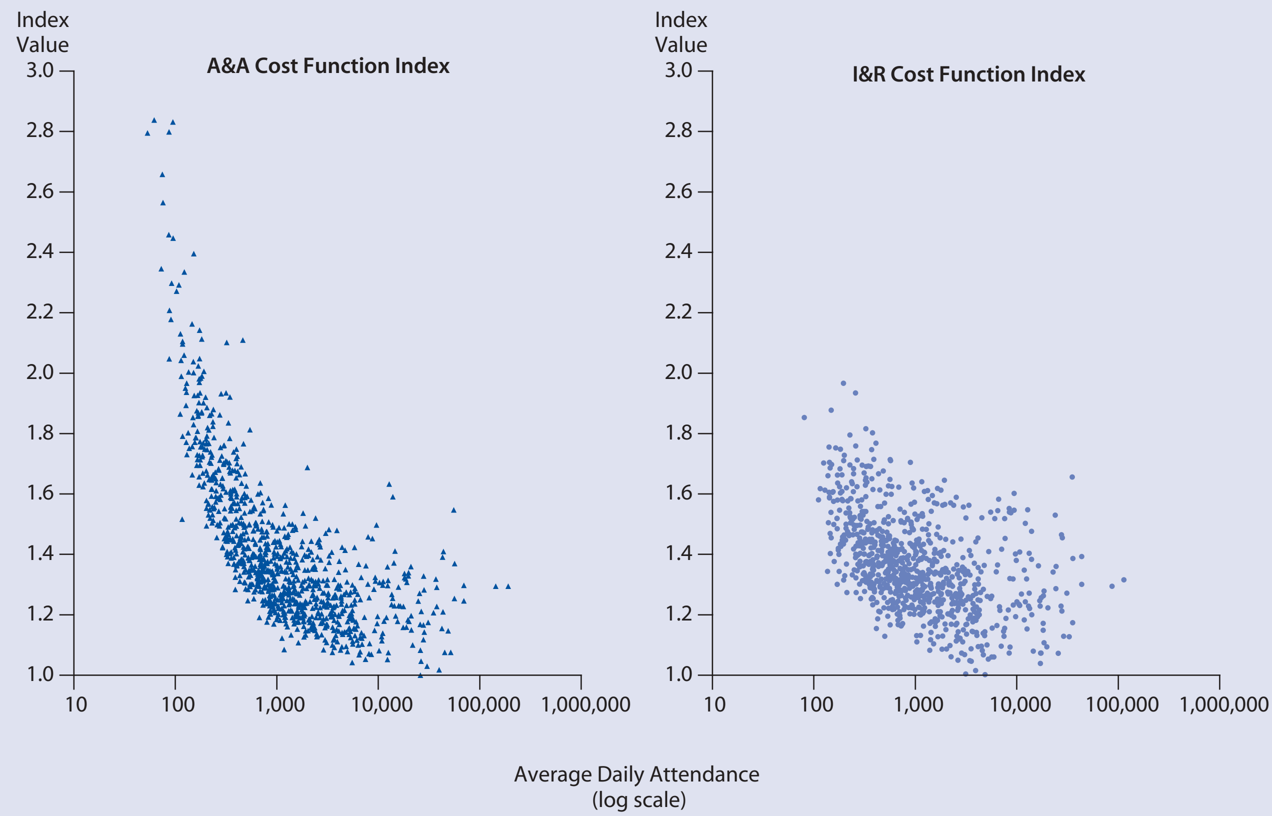
Not only do economies of scale explain most of the variation in the cost function indexes (78 percent for the A&A Cost Function Index, 82 percent for the I&R Cost Function Index), they also explain much of the difference in findings across the methodologies. More than half of the difference between any of the price indexes and the A&A Cost Function Index can be explained by school district size.[8] One-third of the difference between the I&R Cost Function Index and the other indexes can be explained by size.[9]

7 Imazeki and Reschovsky did not provide index values for school districts with fewer than 100 students.

8 This conclusion is based on the R-squares from a regression of the difference in the two indexes on the log of average daily attendance, its square, cube, and quartic.

9 Size has less power to explain the difference between the I&R Cost Function Index and the other indexes because the I&R Cost Function Index is not available for districts with less than 100 students in average daily attendance.

Figure 1. The cost function indexes suggest striking economies of scale



SOURCE: A&A Cost Function Index: Alexander et al. (2000); I&R Cost Function Index: Imazeki and Reschovsky (2002).

Table 5. Variations in Texas school districts according to average daily attendance across seven cost indexes

	Current Texas CEI	Updated Texas CEI	Texas TCI	GCEI	A&A Cost Function Index	I&R Cost Function Index	COL Index
Houston Independent School District	1.15	1.18	1.23	1.38	1.30	1.31	1.84
Dallas Independent School District	1.14	1.19	1.24	1.42	1.30	1.29	1.94
Average daily attendance is greater than 10,000 students and less than 100,000 students							
Mean	1.12	1.15	1.20	1.36	1.23	1.28	1.61
Standard deviation	0.03	0.03	0.04	0.06	0.12	0.15	0.23
Number of districts	73	73	73	73	73	73	73
Average daily attendance is greater than 1,000 students and less than 10,000 students							
Mean	1.07	1.09	1.12	1.25	1.26	1.30	1.43
Standard deviation	0.03	0.04	0.04	0.08	0.11	0.13	0.27
Number of districts	395	395	395	395	395	390	395
Average daily attendance is greater than 100 students and less than 1,000 students							
Mean	1.04	1.05	1.08	1.15	1.53	1.42	1.29
Standard deviation	0.02	0.03	0.03	0.06	0.22	0.14	0.22
Number of districts	525	526	526	526	490	413	526
Average daily attendance is less than 100 students							
Mean	1.05	1.06	1.06	1.07	2.50	—	1.26
Standard deviation	0.02	0.03	0.02	0.06	0.27	—	0.17
Number of districts	46	46	46	46	13	—	46

—Not available.

NOTE: All indexes have been rescaled so that the least cost Texas district is assigned an index value of 1.

SOURCE: Current Texas CEI: Monk and Walker (1991), and Texas Education Agency; Updated Texas CEI: Alexander et al. (2000, 2002); Texas TCI: Alexander et al. (2000, 2002); GCEI: Chambers (1999); A&A Cost Function Index: Alexander et al. (2000); I&R Cost Function Index: Imazeki and Reschovsky (2002); COL Index: Alexander et al. (2000).

Interestingly, these economies of scale tend to fade away at relatively low attendance levels. The correlation between average daily attendance (or its logarithm) and either of the cost function indexes is negligible for school districts with more than 2,000 students. Consequently, the indexing strategies generally indicate little difference in cost between the state's two largest districts—Houston and Dallas. With nearly 200,000 students, the Houston Independent School District has one-third more students than the Dallas Independent School District, yet the cost function indexes make little distinction between them. Only the COL Index identifies a substantial cost difference between the Houston Independent School District and the Dallas Independent School District, and it gives the nod to Dallas as being the higher cost area.

Another dimension about which the indexes yield very different perspectives involves the socioeconomic status of the students. As table 6 illustrates, the Teacher Cost Indexes and the GCEI exhibit a "U-shaped" or slightly "J-shaped" relationship. Apparent costs are high in districts with a high proportion of economically disadvantaged students (disadvantaged districts), and in districts with a low proportion of economically disadvantaged students (advantaged districts). On average, costs are lowest in districts in the middle of the range. For the Texas TCI and the Updated Texas CEI, there is no significant difference in index values between advantaged districts and disadvantaged districts. The Current Texas CEI is somewhat skewed, with the index values significantly higher in disadvantaged districts; the GCEI is skewed in the other direction, with significantly higher values in *advantaged* districts.

The other indexes yield linear, but contradictory relationships. The COL Index is lowest in disadvantaged districts and highest in advantaged districts. The cost function indexes are highest in disadvantaged districts, and lowest in advantaged districts. However, the I&R Cost Function Index is much more responsive than the A&A Cost Function Index to variations in the percent of disadvantaged students. Fully 61 percent of the variation in the I&R Cost Function Index can be explained by variations in the socioeconomic status of the students, while only 22 percent of the variation in the A&A Cost Function Index can be explained by students' socioeconomic status.

Table 6. Economically disadvantaged Texas school districts across seven cost indexes

	Current Texas CEI	Updated Texas CEI	Texas TCI	GCEI	A&A Cost Function Index	I&R Cost Function Index	COL Index
Greater than 75 percent economically disadvantaged							
Mean	1.10	1.10	1.13	1.22	1.61	1.59	1.28
Standard deviation	0.04	0.04	0.06	0.12	0.33	0.11	0.19
Number of districts	98	99	99	99	87	71	99
Economically disadvantaged greater than 25 percent and less than 75 percent							
Mean	1.05	1.07	1.10	1.19	1.42	1.36	1.34
Standard deviation	0.03	0.04	0.04	0.09	0.23	0.12	0.25
Number of districts	809	809	809	809	767	696	809
Less than 25 percent economically disadvantaged							
Mean	1.07	1.10	1.12	1.27	1.24	1.16	1.58
Standard deviation	0.03	0.05	0.05	0.11	0.26	0.09	0.27
Number of districts	134	134	134	134	119	112	134

NOTE: All indexes have been rescaled so that the least cost Texas district is assigned an index value of 1.

SOURCE: Current Texas CEI: Monk and Walker (1991), and Texas Education Agency; Updated Texas CEI: Alexander et al. (2000, 2002); Texas TCI: Alexander et al. (2000, 2002); GCEI: Chambers (1999); A&A Cost Function Index: Alexander et al. (2000); I&R Cost Function Index: Imazeki and Reschovsky (2002); COL Index: Alexander et al. (2000).

Conclusions and Implications

All of the estimates of the cost of education in Texas find substantial variations across the state. The most conservative estimate implies that costs in the highest cost districts are 18 percent higher than in the least cost districts. More liberal estimates imply a range *more than ten times* greater than the most conservative estimates. It is important to note, however, that these estimates are highly sensitive to the indexing strategy employed. No estimate can explain more than 69 percent of the variation in any other estimate. Estimates for rural Texas districts are even more inconsistent across models. To take an extreme example, index values for Allison Independent School District in rural Wheeler county range from 1.02 to 2.83.

So why the dramatic differences? Changes in the underlying characteristics of districts or shifts in the cost technology can explain some differences. However, they are clearly not the primary source of variation. Four of the seven indexes are drawn from data on the 1998–99 school year (Alexander et al. 2000, 2002), and the fifth was drawn from data on the 2000–2001 school year (Imazeki and Reschovsky 2002). Only the GCEI (1993–94) and the Current Texas CEI (1988–89) measure educational costs at markedly different points in time. Furthermore, despite a 10-year gap between estimates, the update to the Current Texas CEI is more highly correlated with its predecessor than with any of its contemporaries.

The primary differences across indexes are attributable to differences in methodology. Such sharp differences across estimation strategies support four important conclusions.

First, the lion's share of variations in input prices arises from variations across labor markets. Table 7 illustrates the extent of within-market variation in the indexes. As the table illustrates, between 66 and 82 percent of the variation in the Teacher Cost Indexes or the GCEI reflects variations across labor markets. Because within-market variations are relatively small compared to the between-market variations, the cost-of-living approach appears to be a viable indexing strategy.

Second, a somewhat crude estimate of comparable wages is only moderately successful at explaining these market-level variations. The modest correlation between the COL Index and the other price indexes implies that the COL Index is unduly noisy, that the population used to generate the COL Index is not comparable to educators, or that the hedonic salary models are all misspecified in some way. Given the imprecision with which the COL Index is measured, excessive noise is the most likely explanation. However, the fact that the COL Index is more than twice as correlated with the GCEI (which includes wage measures for classified personnel) as with the Teacher Cost Indexes (which reflect only teacher compensation) suggests that comparability might also be important. In either case, more refined analysis of a comparable-wage model could promise significant benefits.

Table 7. Within-market and between-market variations in Texas school districts across seven cost indexes

	Current Texas CEI	Updated Texas CEI	Texas TCI	GCEI	A&A Cost Function Index	I&R Cost Function Index	COL Index
Within-market variation	0.28	0.53	0.83	1.91	29.17	10.09	0.00
Between-market variation	0.97	1.32	1.62	8.67	35.46	9.91	69.87
Total variation	1.25	1.84	2.45	10.59	64.63	20.00	69.87
Share of variation that is within market (in percent)	22.5	28.7	33.9	18.1	45.1	50.4	0.0

NOTE: All indexes have been rescaled so that the least cost Texas district is assigned an index value of 1.

SOURCE: Current Texas CEI: Monk and Walker (1991), and Texas Education Agency; Updated Texas CEI: Alexander et al. (2000, 2002); Texas TCI: Alexander et al. (2000, 2002); GCEI: Chambers (1999); A&A Cost Function Index: Alexander et al. (2000); I&R Cost Function Index: Imazeki and Reschovsky (2002); COL Index: Alexander et al. (2000).

Third, there are significant variations across different specifications within each modeling strategy. Although the Teacher Cost Indexes are well correlated with one another in urban areas, the relationship is much weaker in rural parts of the state. Similarly, while the cost function indexes are highly correlated with one another in large school districts, they are much less so in small ones. The sensitivity of the index values to specification differences suggests that researchers should carefully examine the stability of their estimates and formally incorporate the imprecision of their estimates into their policy recommendations concerning finance formula adjustments.

Finally—and most importantly—the differences across these indexes strongly imply that the cost of educational inputs is a poor proxy for the cost of educational outcomes. There is at best no correlation and at worst an *inverse* correlation between cost estimates based on input prices and cost estimates based on educational outputs. Of course, serious measurement issues impede our ability to model the cost of producing educational outcomes, but the Texas estimates strongly imply that these problems must be addressed. As policy discussions about education finance shift from considerations of tax equity to considerations of educational adequacy, there will be an increasing need for accurate measures of the cost of producing educational outcomes. And the ability of researchers to address this need will in no small part depend on advancements in the area of geographic cost adjustments.

References

Alexander, C.D., Gronberg, T.J., Jansen, D.W., Keller, H., Taylor, L.L., and Treisman, P.U. (2000). *A Study of Uncontrollable Variations in the Costs of Texas Public Education* (Summary report prepared for the 77th Texas Legislature). Austin: Charles A. Dana Center, The University of Texas at Austin. Available: *http://www.utdanacenter.org*

Alexander, C.D., Gronberg, T.J., Jansen, D.W., Keller, H., Taylor, L.L., and Treisman, P.U. (2002). *A Study of Uncontrollable Variations in the Costs of Texas Public Education: Technical Supplement* (Supplement to summary report prepared for the 77th Texas Legislature). Austin: Charles A. Dana Center, The University of Texas at Austin. Available: *http://www.utdanacenter.org*

Chambers, J.G. (1998). *Geographic Variations in Public School Costs* (NCES 98–04). U.S. Department of Education. Washington, DC: National Center for Education Statistics Working Paper.

Chambers, J.G. (1999). Patterns of Variation in the Salaries of School Personnel: What Goes on Behind the Cost Index Numbers? *Journal of Education Finance, 25*(2): 255–280.

Duncombe, W., Ruggiero, J., and Yinger, J. (1996). Alternative Approaches to Measuring the Cost of Education. In H.F. Ladd (Ed.), *Holding Schools Accountable: Performance-Based Reform in Education* (pp. 327–356). Washington, DC: The Brookings Institution.

Goldhaber, D.D. (1999). An Alternative Measure of Inflation in Teacher Salaries. In W.J. Fowler, Jr. (Ed.), *Selected Papers in School Finance, 1997–99* (NCES 1999–334) (pp. 29–54). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Guthrie, J., and Rothstein, R. (1999). Enabling 'Adequacy' to Achieve Reality: Translating Adequacy Into State School Finance Distribution Arrangements. In H.F. Ladd, R. Chalk, and J.S. Hansen (Eds.), *Equity and Adequacy in Education Finance* (pp. 209–259). Washington, DC: National Academy Press.

Hanushek, E.A. (1999). Adjusting for Differences in the Costs of Educational Inputs. In W.J. Fowler, Jr. (Ed.), *Selected Papers in School Finance, 1997–99* (NCES 1999–334) (pp. 13–28). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Hanushek, E.A., and Taylor, L.L. (1990). Alternative Assessments of the Performance of Schools: Measurement of State Variations in Achievement. *Journal of Human Resources, 25*(2): 179–201.

Imazeki, J., and Reschovsky, A. (1999). Measuring the Costs of Providing an Adequate Public Education in Texas. In H. Chernick (Ed.), *Proceedings of the 91st Annual Conference on Taxation* (pp. 275–290). Washington, DC: National Tax Association.

Imazeki, J., and Reschovsky, A. (2002). *School Finance Reform in Texas: A Never-Ending Story?* Unpublished manuscript.

McMahon, W.W. (1996). Intrastate Cost Adjustments. In W.J. Fowler, Jr., (Ed.), *Selected Papers in School Finance, 1994* (NCES 96–068) (pp. 89–114). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Monk, D.H., and Walker, B.D. (1991). The Texas Cost of Education Index: A Broadened Approach. *Journal of Education Finance, 17*(2): 172–192.

Rothstein, R., and Smith, J.R. (1997, May). *Adjusting Oregon Education Expenditures for Regional Cost Differences: A Feasibility Study* (Submitted to Confederation of Oregon School Administrators). Sacramento, CA: Management Analysis & Planning Associates, L.L.C.

Stoddard, C. (2002). *Adjusting Teacher Salaries for the Cost of Living: The Effect on Salary Comparisons and Policy Conclusions.* Unpublished manuscript.

Wynne, M., and Sigalla, F. (1994). The Consumer Price Index. *Economic Review* (Research Journal of the Federal Reserve Bank of Dallas) 1994, 2nd quarter: 1–22.

Financing an Adequate Education: A Case Study of New York

William Duncombe
Syracuse University

Anna Lukemeyer
University of Nevada, Las Vegas

John Yinger
Syracuse University

About the Authors

William Duncombe is Professor of Public Administration, and Senior Research Associate, Center for Policy Research, the Maxwell School, Syracuse University. He is also the Associate Director of the Education Finance and Accountability Program at Syracuse University. He has published numerous articles on design of school finance systems, the estimation of educational costs, and state and local budgeting and financial management.

Anna Lukemeyer is an Assistant Professor of Public Administration in the Greenspun College of Urban Affairs at University of Nevada, Las Vegas. She received a J.D. from Southern Methodist University School of Law and a Ph.D. from the Maxwell School at Syracuse University. She is the author of Courts as Policymakers: School Finance Reform Litigation, which is part of the American Legal Institutions series published by LFB Scholarly Publishing LLC.

John Yinger is Professor of Economics and Public Administration and Associate Director of the Center for Policy Research at the Maxwell School, Syracuse University. He is Director of the Education Finance and Accountability Program at Syracuse University. He is a specialist in state and local public finance, including education finance, racial and ethnic discrimination in housing, and urban economics. Professor Yinger has published widely in leading scholarly journals, and has co-authored two books on local public finance.

Financing an Adequate Education: A Case Study of New York

William Duncombe

Syracuse University

Anna Lukemeyer

University of Nevada, Las Vegas

John Yinger

Syracuse University

Acknowledgments

This study would not have been possible without the financial assistance of The Research Foundation of State University of New York under the "Educational Finance Research Consortium" project funded by the New York State Department of Education. In particular, we would like to thank Professor James Wyckoff of the University of Albany for his support of this project. The conclusions and recommendations of this study are those of the authors only, and the authors are fully responsible for all errors and omissions.

Introduction

The New York State Board of Regents and Commissioner of Education have identified a set of clear performance standards for students in New York State. These standards represent the knowledge and skills students are expected to need in order to function successfully as productive citizens in the 21st century. These standards will be implemented through new "high-stakes" Regents examinations, which all students will be required to pass to graduate from high school, and supported by new examinations in the fourth and eighth grades, which will serve as important intermediate checkpoints in assessing student progress.

New York is not alone in setting higher standards for its students. Over the last decade, many states have implemented higher standards, and by 2004, almost half the states will require passage of exit exams for high school graduation (Meyer et al. 2002). Although this movement toward higher standards is driven primarily by state education departments and state elected officials, it has other roots as well. State courts often interpret the education clauses in their state constitutions as obligating the state to ensure that all children have the opportunity to reach an adequate level of content knowledge and skill (Lukemeyer 2003). New York's school finance system, for example, has been challenged in state court as unconstitutional because it does not provide a "sound basic education."[1]

[1] New York's highest court, the Court of Appeals, has interpreted article XI, section 1, of the state constitution as requiring the legislature to "ensure the availability of a sound basic education to all the children of the State." *Campaign for Fiscal Equity,* 655 N.E.2d 661 [*"CFE1"*] at 665; *Board of Education v. Nyquist* (1982). The two most recent decisions in the ongoing litigation include *Campaign for Fiscal Equity*, 719 N.Y.S.2d 475 (2001) (*"CFE2"*), and *Campaign for Fiscal Equity*, 744 N.Y.S.2d 130 (2002) (*"CFE3"*). In *CFE2*, the trial court found the system unconstitutional, but New York's intermediate appellate court reversed the trial court's decision in *CFE3*. The case has been appealed to New York's highest court, the Court of Appeals.

Moreover, the federal No Child Left Behind Act of 2001 requires states to implement annual testing from third through eighth grade as part of a broader accountability system that includes school report cards and state-set minimum performance standards (Robelen 2002).

Despite the clear trend toward higher standards in education, states have been slow to implement funding systems designed specifically to help students (and schools) reach new standards (Boser 2001). The objective of this paper is to provide state governments with tools to help them develop a school finance system that supports students and school districts trying to reach higher performance standards. The paper focuses on a well-known problem, namely, that schools with disadvantaged students must spend more than other schools to meet any given standard. This paper shows how to estimate each district's cost for achieving an adequacy standard and develops a foundation aid formula that adjusts for the higher costs in some districts.

The objective of this paper is to provide state governments with tools to help them develop a school finance system that supports students and school districts trying to reach higher performance standards.

The development of any adequacy-based school finance system involves three components, which correspond to the three substantive sections of this paper:

First, a state must select measures of adequacy, either in terms of resources or student performance. Such measures are necessary to identify school districts below the standard. Although these measures can be controversial and difficult to develop, this choice is unavoidable.

Second, a state must estimate the cost of reaching a given performance standard in each district. The cost function approach presented in this study relies on statistical methods to extract from actual data the impact of student needs, resource prices, and enrollment size on the spending required to reach a particular standard.

Third, a state must develop a school aid formula. This formula should provide all school districts the resources they need to reach the adequacy standard selected by the state.

This paper explains how each of these steps can be implemented, with illustrations based on data from New York State.[2] Our objective is to provide guidance for any state that wants to design an adequacy-based finance system.

Developing an Adequacy Standard

In setting an adequacy standard, a state must first decide whether the standard is intended to guarantee each district some minimum level of resources or to give all students the opportunity to reach a minimum level of student performance. A resource standard is typically represented in terms of a bundle of resources and course requirements that represent an opportunity for an adequate education. In contrast, a performance standard usually is expressed as a level of student performance on standardized exams. One set of examinations is unlikely to capture all dimensions of an adequate education, as defined by the courts or the general public; nevertheless, many states are setting adequacy standards by making the passage of specific tests either an objective or a graduation requirement.

In New York State, the debate over performance standards has not yet been resolved. Both the Board of Regents and Commissioner of Education have identified a clear set of performance requirements for students to graduate from high school. However, the courts have not yet identified the standards required by the New York State Constitution.

In a 1995 decision, New York's highest court defined the constitutional requirement that the state provide a "sound basic education" in terms of both student performance (knowledge and skills necessary to vote and serve on a jury) and resources (minimally adequate

[2] A more detailed discussion of data and methods used in this paper is available in Duncombe (2002), particularly appendix A (data sources and measures) and appendix B (statistical models and methods).

facilities, material, and teaching).[3] In later decisions, however, lower courts have differed as to the level of student performance that this definition requires. In January 2001, the trial court ruled that "a capable and productive citizen . . . is capable of serving impartially on trials that may require learning unfamiliar facts and concepts and . . . decid[ing] complex matters that require . . . verbal, reasoning, math, science, and socialization skills. . . ." (*CFE2* at 485) This implies that high school graduation from a reasonably demanding program is a requirement for productive citizenship. In contrast, in June 2002, an intermediate appellate court ruled that "The State submitted evidence that jury charges are generally at a grade level of 8.3, and newspaper articles on campaign and ballot issues range from grade level 6.5 to 11.7. . . . Thus, the evidence at trial established that the skills required to enable a person to obtain employment, vote, and serve on a jury, are imparted between grades 8 and 9, a level of skills which the plaintiffs do not dispute is being provided." (*CFE3* at 138) In other words, this court ruled that high school graduation is not mandatory for meeting the constitutional standard.

While translating these court decisions into specific performance measures is beyond the scope of this paper, it is clear that the level of student performance associated with "productive citizenship" as defined by the courts will have a large impact on the school finance system. In selecting a measure of performance to use in estimating the cost of adequacy, we have drawn from the measures developed by the New York State Education Department (SED). First, we average math and English exam scores in fourth grade, eighth grade, and high school. The measure used in this study is based on a weighted average of fourth- and eighth-grade exam scores, and high school Regents exam scores. Regents exam scores were weighted twice as heavily as fourth- and eighth-grade exam scores to reflect the fact that students are now required to pass these exams for high school graduation.[4] The resulting composite test scores can range from 0 to 200.

For comparison purposes, we are going to look at the costs associated with two standards, 130 and 160. A standard of 130 might be consistent with the third CFE decision (*CFE3*), because it implies adequate performance for all fourth- and eighth-grade students, but

[3] The Court of Appeals stated:

> Such an education should consist of the basic literacy, calculating, and verbal skills necessary to enable children to eventually function productively as civic participants capable of voting and serving on a jury. If the physical facilities and pedagogical services and resources made available under the present system are adequate to provide children with the opportunity to obtain these essential skills, the State will have satisfied its constitutional obligation. As we stated in *Levittown,*
>
> > The Legislature has made prescriptions (or in some instances provided means by which prescriptions may be made) with reference to the minimum number of days of school attendance, required courses, textbooks, qualifications of teachers and of certain nonteaching personnel, pupil transportation, and other matters. If what is made available by this system (which is what is to be maintained and supported) may properly be said to constitute an education, the constitutional mandate is satisfied. (57 N.Y.2d, at 48.)
>
> The State must assure that some essentials are provided. Children are entitled to minimally adequate physical facilities and classrooms which provide enough light, space, heat, and air to permit children to learn. Children should have access to minimally adequate instrumentalities of learning such as desks, chairs, pencils, and reasonably current textbooks. Children are also entitled to minimally adequate teaching of reasonably up-to-date basic curricula such as reading, writing, mathematics, science, and social studies, by sufficient personnel adequately trained to teach those subject areas.

(*CFE1* at 666 [footnote omitted])

[4] Newly developed examinations in mathematics and English language arts are required of all fourth- and eighth-grade students. SED has divided test results into four levels and reports the counts (and percent) of students reaching a given level. The levels are selected to reflect students with "serious academic deficiencies" (level 1), students needing "extra help to meet the standards and pass the Regents examinations" (level 2), students meeting "the standards and with continued steady growth, should pass the Regents examinations" (level 3), and students exceeding "the standards and are moving toward high performance on the Regents examination" (level 4). The percent of students reaching each level is first identified, and then a weighted average of these percents is calculated with a weight of 1 for level 2 and a weight of 2 for levels 3 and 4. With relatively few exceptions (e.g., severe disabilities), all students will have to pass a series of Regents examinations to receive a regular high school diploma. A similar process is used to aggregate results for the Regents examinations. The percent of students receiving between 55 and 64 on the Regents exams in math and English are given a weight of 1, and the percent of students receiving above a 64 are weighted at 2. Performance in high school is a more accurate reflection of the accumulated knowledge and skills of students than performance in earlier grades. Thus, a weight of 50 percent is applied to the Regents exams, 25 percent to fourth-grade exams, and 25 percent to eighth-grade exams in constructing an overall performance measure. Sensitivity analysis was also performed using equal weights on exams from all three grade levels. The results of the analysis are not highly sensitive to these weights. See Duncombe (2002), appendix A, for a more detailed discussion of these measures.

only basic competency for most students on the high school exams. Taken literally, the new Regents standards imply a score close to 200, because students are required to pass the Regents exams to receive a high school diploma. Very few districts would presently meet a standard of 200. A more realistic standard that still might be consistent with the second CFE decision (*CFE2*) would be the present state average of 160. Most districts in New York already meet this standard, but a standard of 160 would be a very ambitious standard for many urban districts.

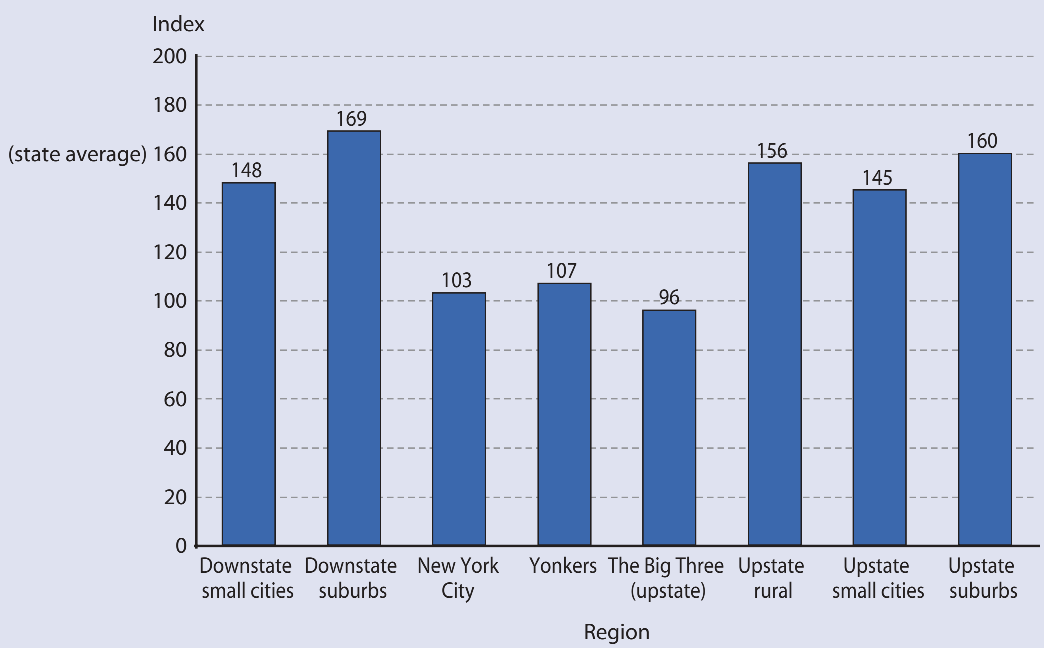
As indicated in figure 1, there are wide disparities in student achievement across districts in New York State, and these disparities are tied closely to school district size and urbanization. The five large city school districts have performance levels of approximately 100, which is well below both the current state average and our more modest standard of 130. Only 5 percent of the districts don't reach a standard of 130, but these districts serve close to half the students in the state. Most of the suburban districts and many rural districts exceed the state average of 160.

Estimating the Cost of Adequacy

The heart of any adequacy-based finance system is an estimate of the costs or spending required for each district to reach a particular resource or performance standard. This cost cannot be directly observed for a low-performing district, so this step requires a method to estimate the extent to which some districts must pay more than others for the same performance because of characteristics, such as student poverty, that are outside their control. This calculation leads to a cost index, which can then be used to determine how much money each district needs to boost its student performance. This approach is analogous to estimating and applying a cost-of-living index. If one location has a cost of living that is higher than average, then people living in that location must receive a higher income than people in the average location in order to achieve the same standard of living. Estimating a cost index is complicated, however, and several different approaches have been developed.[5] In this paper, we focus on one method, which is called the "cost function approach."

[5] For a review of these methods, see Guthrie and Rothstein (1999) and Duncombe and Yinger (1999).

Figure 1. Comparison of student performance index by New York region: 2000



SOURCE: New York State Education Department.

The cost function approach uses statistical methods to relate data on actual spending in school districts to student performance, resource prices, student needs, and other relevant district characteristics.[6] The resulting estimates are used to construct an education cost index, which measures how factors outside a district's control affect the spending required to reach a given resource or student performance level. The cost function approach is well suited to developing estimates of the cost of adequacy in individual districts, and the results can be used directly in aid formulas.

These benefits are contingent, however, on the quality of the data used in statistical analysis and the accuracy of the statistical results. Any researcher estimating an education cost function must make a number of choices. Each of these choices may affect the statistical results, in some cases significantly, and some of these choices are not "transparent" to policymakers and educators.[7] The onus is on a researcher using the cost function approach to explain the method in an intuitive fashion and to convince policymakers and other policy analysts that reasonable choices were made. In this section, we discuss the choices we made in applying the cost function approach to New York.

Because the primary resources used by school districts are teachers and other professional staff, adjusting for differences in the cost of hiring teachers is particularly important.

The first step in the cost function approach is to estimate a teacher cost index. As discussed below, a teacher cost index is sometimes used on its own as a measure of resource cost differences across school districts. In addition, however, a teacher cost index plays a critical role in an analysis of total educational costs, which must consider not only resource costs differences, but also differences in costs that arise because of district size or the presence of many disadvantaged students (also known as "at-risk" students). We begin this section, therefore, by explaining how to estimate a teacher cost index and by presenting teacher cost index results for New York. We then turn to our method for estimating a full education cost index, that is, for determining the resources each district needs to provide a given quality education given its resource costs, its enrollment, and its concentration of at-risk students. The section ends with a presentation of cost index results for New York school districts.

Estimating a Teacher Wage Model and a Teacher Cost Index

If a state's adequacy standard requires that all districts receive a minimum level of resources, then a state aid program needs to make some adjustment for the higher cost of purchasing educational resources in some school districts than others. Because the primary resources used by school districts are teachers and other professional staff, adjusting for differences in the cost of hiring teachers is particularly important.[8] Such differences could arise for several reasons. Specifically, some districts may have to pay significantly more than others to recruit teachers of equal quality because of a higher cost of living in the area, strong competition from the private sector for similar service-sector occupations, or more difficult working conditions facing teachers. Not all teachers consider the same factors in evaluating working conditions, but classroom discipline problems, violence in schools, and a general lack of student motivation are likely to make a teaching job less attractive to most teachers.

In developing a teacher cost index, it is important to distinguish between discretionary factors that a district can influence, and labor market or working con-

6 For other examples of this approach, see Downes and Pogue (1994), Reschovsky and Imazeki (1997), and Duncombe and Yinger (2000).

7 The cost function approach has been criticized and ultimately rejected by some researchers, because its technical complexity makes it difficult to explain to "reasonably well-educated policymakers" (Guthrie and Rothstein 1999, p. 223). In our view, this is an inappropriate criterion for selecting a method for estimating the cost of adequacy, because simpler approaches, even if they are easier to explain, may be grossly inaccurate. The main criteria in selecting a method should be accuracy, not transparency.

8 In principle, cost differences can also be calculated for other inputs, such as transportation, energy, and facilities, but this step is rarely included in practice. For a good introduction to methods for calculating input cost differences, see Fowler and Monk (2001).

dition factors that are outside a district's control.[9] Factors a district can influence include the experience and education of its teaching force, the certification level of its staff, the size of schools and classes, average student performance, and the general level of efficiency in the district. Factors outside a district's control include labor market factors, such as private sector salaries and unemployment rates, and factors related to working conditions, such as a concentration of at-risk students, juvenile crime rates, and pupil density. A teacher cost index that is used to help compensate high-need districts as part of a state aid system obviously should only reflect factors that a district cannot control. As a result, a teacher wage model accounts for factors influenced by a district but does not consider them in calculating the teacher cost index.

A teacher cost index that is used to help compensate high-need districts as part of a state aid system obviously should only reflect factors that a district cannot control.

Using information on individual teacher salaries and characteristics in 2000, along with school and district characteristics, we estimate a teacher wage model for New York State. The sample size is over 120,000 full-time classroom teachers, representing almost all the state's districts. The dependent variable is the teacher's salary, without fringe benefits or compensation for extracurricular activities.[10] The model is estimated with standard linear regression techniques.[11] The explanatory variables include a wide range of teacher, school, and district characteristics. The 2-year average share of K–6 students eligible for a free lunch, for example, is used as a measure of student poverty.[12] A complete list of the variables in the model is provided in appendix table A-1.

The results for the teacher wage model are reported in table 1. Looking first at teacher characteristics, most of the variables are statistically significant and have the expected sign. There is a positive relationship, for example, between teacher salaries and total teaching experience, whether the teacher has a graduate degree, whether she teaches math or science, and the percentage of assignments in which she is certified to teach. The two variables representing the quality of the college the teacher attended (as rated by *U.S. News & World Report*) have the expected positive sign, but they are not statistically significant.

Among the other discretionary factors, we found that working in a larger school and having larger classes are associated with higher wages, holding other factors constant, but the class-size effect is not statistically significant. Not surprisingly, we found that the more resources that a district has relative to its peer groups, the higher the wages are.[13] One unusual result is the positive coefficient for the student outcome measure, which implies that teachers require additional pay to work with high-performing students. Another possi-

[9] For a detailed discussion of the process of developing a teacher cost index and a cost of education index, see Chambers (1997).

[10] Following many other studies, the teacher salary variable is specified as the natural logarithm of the observed salary.

[11] Because the equation is estimated at the individual teacher level, it is reasonable to assume that teachers are price takers, that is, that they cannot influence the salary schedule they face or the underlying personnel policies of the school district. Thus, endogeneity of some of the independent variables is not likely to be a problem. However, the variables used in the model are from at least two different levels of aggregation, the individual teacher and the school district. This implies that the standard errors from an ordinary least squares regression (OLS) are biased, because the error terms are not independent across observations. In particular, the estimated standard errors on district-level variables may significantly understate the actual standard errors. We use a well-known method to correct for this problem. See Huber (1967) and White (1980). These corrections were made using the software package STATA, and clustering was assumed only at the district level. There are three variables at the county level—professional wage, unemployment, and crime rate. It is possible that the standard errors for these variables are underestimated. Finally, the model was initially estimated with a measure of high-cost special needs students, but the coefficient was not found to be statistically significant. The final model was estimated without this variable.

[12] One of the difficulties of estimating a "reduced form" teacher wage model is that variables, such as poverty, can pick up both working condition differences and fiscal capacity differences across districts. The coefficient on the percent of free-lunch students was consistently negative, suggesting that this variable is picking up fiscal capacity differences. To separate these two effects, we regressed the percent free-lunch students on the natural log of per pupil income and property values, and used the residual in the regression as the measure of poverty. This variable had the expected positive relationship with wages, holding other factors constant.

[13] This is one of the so-called efficiency variables, which are discussed later in the paper.

Table 1. Results of the teacher wage model: 2000[1]

Variables	Coefficient	t-statistics
Constant	7.84418	26.40
Teacher characteristics		
Total experience[2]	0.21596	10.13
Master's or higher	0.06403	2.51
Teacher of math/science	0.01261	6.00
Percent of assignments certified	0.03318	7.78
M.A. from top-rated school	0.00932	0.97
B.A. from top-rated school	0.00215	0.88
Factors under district control		
School enrollment[2]	0.01827	4.50
Class size	0.00006	1.39
Aid efficiency variable[3]	0.59311	2.55
Income efficiency variable[3]	0.00000	5.00
Full value efficiency variable[3]	0.00000	0.45
Average student performance	0.00348	7.50
Factors outside district control		
Labor market factors		
Average unemployment rate (1997–99)	–0.01626	–3.95
Pupil density[2]	0.03074	5.58
Professional wage[2]	0.14947	5.22
Share of county's teachers	–0.16798	–3.00
Working condition factors		
Average percent LEP[4] students	0.43459	2.03
Adjusted free lunch student rate[5]	0.23406	5.38
Juvenile violent crime rate	–45.71180	–3.72
District enrollment[2]	0.02708	2.50
Adjusted R-square	0.71400	

[1]Estimated with ordinary least-squares regression, with standard errors adjusted for nonindependence using Huber (White) method. Dependent variable is the natural logarithm of teacher salaries. Sample size is 121,203.

[2]Expressed as natural logarithm.

[3]Calculated as the difference between district level and average level in peer group. See Duncombe (2002), appendix B.

[4]"LEP" means limited English proficient.

[5]Residual from a regression of the average (1999–2000) share of free lunch students in elementary school regressed on the log of per pupil income and per pupil property values.

SOURCE: Calculations by authors.

bility is that this variable is picking up fiscal capacity differences across districts associated with unobserved teacher quality.

Turning to the factors outside of district control, we find that most of the variables fit expectations. More urbanized districts pay higher wages, for example, as do districts with higher private sector wages. The coefficient on the unemployment rate variable has the expected negative sign; lower unemployment rates lead to tighter labor markets and higher salaries. Salaries are negatively related to the share of a county's teachers in a district, indicating that districts with relatively large numbers of teachers may be more attractive to teachers because they provide more options.[14]

We also find, as expected, that salaries are affected by the working conditions in a district. To be specific,

[14] Another interpretation for this variable is that it measures the ability of the district to exercise market power over wages. If the variable is interpreted as a monopsony measure, then it would be a discretionary variable and would be held constant in constructing the teacher wage index.

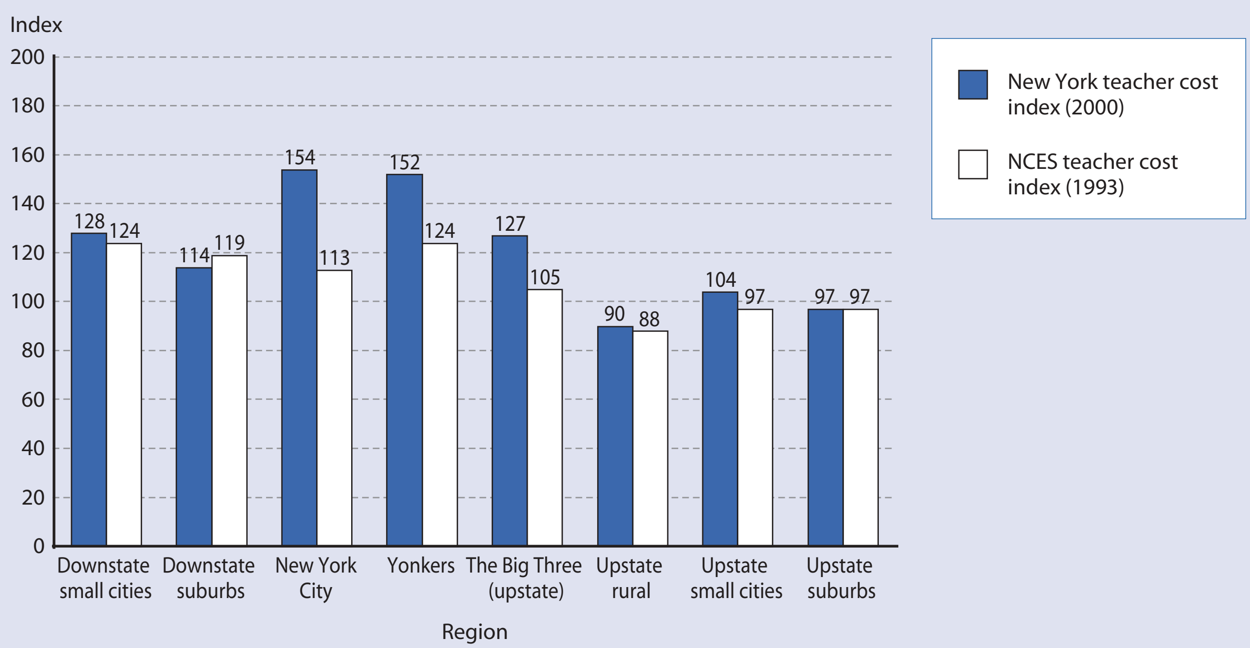
districts with higher shares of students with limited English proficiency or receiving free lunch pay higher salaries, holding other factors constant. Larger districts (in terms of enrollment) are associated with higher salaries, even controlling for school size and pupil density, suggesting that large district size may negatively affect working conditions. One of the variables included to measure working conditions, juvenile violent crime rate, is negatively related to wages. Possible explanations for this counterintuitive result include (1) teacher quality has not been adequately controlled for, so that this variable is picking up both working conditions and lower teacher quality, and (2) the crime rate is capturing omitted urbanization and fiscal capacity variables, and its coefficient reflects the fact that poorer urban areas tend to have lower fiscal capacity. In either case, the crime rate variable does not appear to be reflecting differences in working conditions.

This teacher wage model can be used to develop a measure of the underlying wage that a school district must pay to attract teachers with a given set of characteristics to a school district. As noted earlier, this predicted wage should only measure variation in factors outside a school district's control. Constructing the predicted wage involves three steps: (1) multiplying the regression coefficient associated with each discretionary variable by the state average for that variable, (2) multiplying the regression coefficient associated with each variable outside a district's control by the actual value for that variable in each district, and (3) summing for each district the results from the first two steps to obtain the predicted wage.[15] The teacher wage index is then defined as the ratio of the predicted wage for each district divided by the state average wage and multiplied by 100.

Our teacher cost index for New York is reported in figure 2. This index reveals a distinct difference in resource costs between upstate and downstate districts. Most of the downstate districts have above-average costs, and most of the upstate districts have below-average costs. New York City and Yonkers, for example, would have to pay over 50 percent more than the av-

[15] Because the wage is expressed as a logarithm, the expected wage is the antilog of this sum.

Figure 2. Comparison of teacher cost indexes for New York regions: 1993, 2000



SOURCE: Chambers, J. (1997). *A Technical Report on the Measurement of Geographic and Inflationary Differences in Public School Costs*; and calculations by authors.

erage district to attract similar teachers. These high index values reflect both the high cost of living in downstate New York and the challenging working environment in these two cities. Even though the other large cities, commonly called the Big Three, are located in upstate New York, where the cost of living is below average, their working conditions are so difficult that they still would have to pay salaries 25 percent higher than those in the average district to be able to recruit teachers with similar characteristics.

Figure 2 also presents results for the 1993 teacher cost index developed by Chambers (1997) for NCES.[16] This index shows the same general pattern as our index, but its values for large cities are significantly smaller. The NCES index values for New York City and Yonkers, for example, are only 10 to 25 percent higher than the state average, and only 5 percent higher than the state average for the upstate large cities (the Big Three). Because it is based on more detailed and more recent data and is specific to New York State, we believe that our index provides more credible results than the NCES index. To put it another way, the significant differences between our teacher cost index and the NCES index highlights the importance of careful state-by-state analysis of factors affecting resource costs.

> *A state adequacy standard that requires all districts to raise their students to a given level of student performance cannot be achieved with a standard foundation aid formula.*

Estimating Cost Functions and Full Cost Indexes

A standard foundation aid formula brings all districts up to a minimum level of spending per pupil, but does not ensure a minimum level of student performance. A state adequacy standard that requires all districts to raise their students to a given level of student performance cannot be achieved, therefore, with a standard foundation aid formula. Instead, the only way to ensure that all districts have the resources they need to meet this standard is to implement a foundation aid formula that includes adjustments both for resource cost differences across districts and for the higher level of resources required in some districts because of a concentration of at-risk students and other factors outside their control. The necessary adjustments can be determined by estimating an education cost function and using the results to calculate an overall education cost index.

An education cost function relates per pupil spending in a school district both to factors outside a district's control and to factors a district can influence. Only the former factors are considered, however, in calculating an education cost index. The logic behind a cost function begins with the observation that spending levels in a district are clearly affected by the level of student performance that school officials, and ultimately taxpayers, want to support, a key factor inside the district's control. The cost function we estimate, therefore, includes as an explanatory variable the student performance measure described earlier. Because additional resources are generally required to raise student performance, we expect a positive relationship between student performance and spending, holding other factors constant.

The relationship between spending and performance has to be tempered by the possibility of inefficiency in the use of resources, another factor within a district's control. Some school districts may have high spending relative to their level of student achievement not because of higher costs, but because of inefficient use of resources. Moreover, a cost model requires careful accounting for efficiency differences across districts, because the results may depend on which set of efficiency factors is included.

The literature on managerial efficiency and public bureaucracies suggests three broad factors that might be related to productive inefficiency: fiscal capacity, competition, and factors affecting voter involvement in monitoring government (Leibenstein 1966; Niskanen 1971; Wyckoff 1990; Duncombe, Miner, and Ruggiero 1997). Research on New York school districts suggests incentives for efficient use of resources may be lower in

[16] The NCES index developed by Chambers (1997) is based on a regression model fit to national data on teachers, schools, and districts from several NCES data sources, and other national data sources. While the basic structure of the teacher wage equation is similar, the measures of teacher salary, teacher characteristics, and school district characteristics differ substantially from those used in this study.

wealthier or higher income districts, or those receiving more state aid, because looser financial constraints diminish the incentive for taxpayers to put pressure on their school districts (Duncombe and Yinger 2000). Moreover, school officials have an incentive to compare their school's performance to that of similar districts and will work hard to keep from falling behind other districts at the same level of income or wealth. To measure the relative affluence of a district, we include the difference between a district and the average in its peer group for per pupil income, per pupil property values, and state aid as a percent of district income. In this context, a peer group is defined as one of the need/resource-capacity categories defined by SED, with the five large cities treated as one peer group.[17] We expect that the higher a district's resources relative to its peer group, the less efficient the district will be and thus the more it will spend, all else being equal.

We expect that the higher a district's resources relative to its peer group, the less efficient the district will be and thus the more it will spend, all else being equal.

The other variables in a cost function are factors that are outside a district's control. These cost factors can be divided into three categories, resource prices, student needs, and the physical characteristics of the district. As discussed above, some districts may have to pay significantly more to recruit teachers of equal quality. The average salary for full-time teachers with a graduate degree and 1 to 5 years of experience is used as the teacher salary measure.[18] Factors affecting students' school readiness, motivation, and behavior influence not only the working conditions facing a teacher, and hence competitive salaries, but also the quantity of resources required to reach any given student performance standard. We expect, for example, that students whose native language is not English will require additional resources in the form of bilingual education classes and other support to help them obtain mastery of English and to stay on track in the curriculum. The cost function in this study includes two student need factors: the share of district enrollment that consists of limited English proficient (LEP) students, and the percentage of the district's children between 5 and 17 years old living below the poverty line. Finally, education costs may be affected by certain physical characteristics of a district, including enrollment size and physical terrain. Our cost model includes a set of variables indicating the enrollment level in the district to reflect the fact that costs are likely to be higher in very small school districts (Duncombe and Yinger 2001b).

The dependent variable in the cost model is per pupil operating expenditure for fiscal year 2000.[19] The sample size is 678 school districts. Descriptive statistics for the variables in the cost model are provided in appendix table A-2. One technical complexity arises in estimating this model. Budget decisions involve tradeoffs between desired student performance levels, constraints on local property tax rates, and decisions over teacher salaries. In other words, spending levels, performance targets, and teacher salaries are set simultaneously in the budget process, which implies that the performance measure and teacher salaries are likely to be endogenous and standard regression techniques are likely to yield biased results. Consequently, we estimate the cost model with the appropriate simultaneous-equations procedure.[20]

[17] The categories include New York City, other large cities, high-need urban/suburban, high-need rural, average need, and low need. These districts are classified based on a comparison of fiscal capacity (property values and income) and student needs (students receiving reduced-price lunch, limited English proficient [LEP] students, and students in sparsely populated districts). New York City and the other large cities were combined as one category. See New York State Education Department (2001), appendix, for a description of this classification.

[18] As before, this variable is expressed as a natural logarithm.

[19] Expressed as a natural logarithm.

[20] The cost model was estimated with two-stage least squares regression (2SLS), with instruments selected from characteristics of adjacent school districts. We calculated the average, minimum, and maximum values of adjacent districts for a set of student characteristics, performance levels, physical characteristics, and fiscal capacity measures. These potential instruments are then tested, and those that meet the requirements of an instrument are used in the cost model. Instruments include the log of the pupil density, the average of LEP students in adjacent districts, the maximum for income and performance on the grade 8 exams, and the minimum of performance on grade 8 exams for adjacent districts. See Duncombe (2002), appendix B, for a detailed discussion of the process of selecting instruments.

The cost model results are reported in table 2. In general, the coefficients in the regression models have the expected signs. The student performance variable has a positive coefficient and is statistically significant, indicating that higher performance requires more resources. The precision of this coefficient is important, because it is used in the adequacy calculations discussed below. As anticipated based on our analysis of district inefficiency, the more resources a district has relative to its peers, the higher its spending. Teacher salaries are positively related to per pupil spending and the salary coefficient is sensible; a 1 percent increase in predicted salaries is associated with a 1 percent increase in per pupil spending.

The results for student characteristics also follow expectations. As the proportion of poor students or LEP students increases, the level of spending also increases, controlling for performance. Both of these coefficients are statistically significant at conventional levels. The coefficient on the child poverty variable (LEP variable) indicates that a 1 percentage point increase in the child poverty rate (share of LEP students) is associated with a 0.98 (1.075) percent increase in per pupil spending, all else being equal. Finally, the coefficients for the enrollment class variables indicate that, relative to very small districts (under 1000 students), costs per pupil are generally lower for most enrollment categories except the largest (over 15,000 students). The coefficient on the 1000-to-2000-student variable, for example, indicates that these districts spend, on average, 9.3 percent less than districts with fewer than 1000 students, holding other variables constant. In other words, the smallest districts have the highest costs.

Once an education cost function has been estimated, an education cost index can be calculated in simple steps. For each variable that a district can influence,

Table 2. Results of the education cost models: 2000[1]

Variables	Coefficient	t-statistics
Constant	–2.58360	–2.29
Performance index	0.00752	3.57
Efficiency variables[2]		
Full value	0.00000	10.55
Aid	1.12073	3.83
Income	0.00000	0.61
Average teacher salary[3]	0.99296	7.65
Percent child poverty (1997)[4]	0.97819	5.46
2-year average LEP[5] students[4]	1.07514	2.30
Enrollment classes[6]		
1,000–2,000 students	–0.09342	–4.20
2,000–3,000 students	–0.07956	–2.72
3,000–5,000 students	–0.09500	–2.68
5,000–7,000 students	–0.07944	–2.01
7,000–15,000 students	–0.09579	–2.08
Over 15,000 students	0.05404	0.51
Adjusted R-square	0.493	

[1]Estimated with linear two-stage least squares regression, with the student performance and teacher salaries treated as endogenous. See Duncombe (2002), appendix B for discussion of instruments.

[2]Calculated as the difference between district value and the average in peer group. (See Duncombe 2002, appendix B.)

[3]For full-time teachers with 1 to 5 years of experience. Expressed as natural logarithm.

[4]Variables expressed as percent of enrollment.

[5]"LEP" means limited English proficient.

[6]The base enrollment is 0 to 1,000 students. The coefficients can be interpreted as the percent change in costs from being in this enrollment class compared to the base enrollment class.

SOURCE: Calculations by authors.

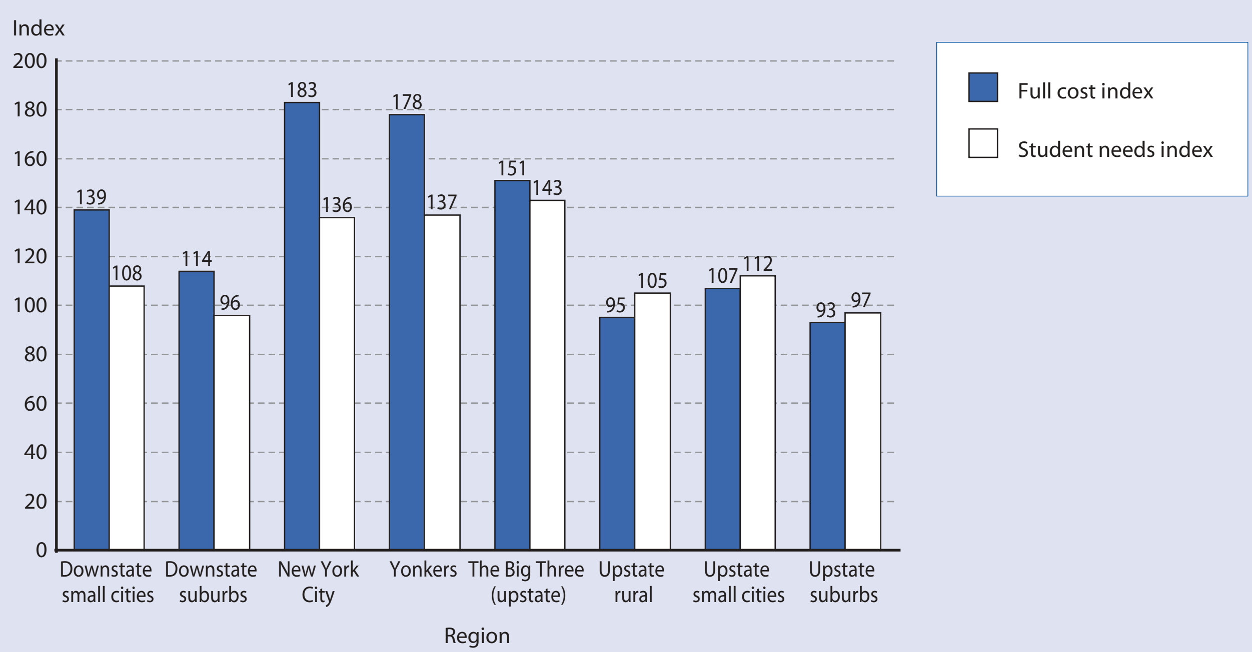
the estimated coefficient from the cost model is multiplied by some constant value for the variable, usually the state average, and these products are summed across all such variables. This approach holds these variables constant across school districts; that is, it does not allow factors inside a district's control to influence its relative educational costs. For each variable outside a district's control, the estimated coefficient from the cost model is multiplied by the actual value for the variable in each district. These products are then summed across all such variables. The variation in these variables across districts is, of course, the source of the variation in the cost index. These two sums (based on factors inside and outside a district's control, respectively) are then added, resulting in a prediction of the amount each district must spend per pupil to obtain an average performance level, assuming that it has the efficiency level in the average district.

The final step is to transform this predicted spending into an index. This step involves dividing predicted spending in each district by predicted spending in a district with average characteristics (including those inside a district's control) and then multiplying the result by 100. This index reveals how much more or less than the average district each district must spend to achieve any given performance standard. An index value of 200 indicates, for example, that a district must spend twice as much as the average district to obtain any given performance standard, whereas an index value of 50 indicates that a district needs to spend only half as much.

We also calculate a student need index, which has the same form as the overall education cost index except that it holds all factors at the state average except for the poverty and LEP variables. A value of 150 for this index, for example, indicates that a district must spend 50 percent more than the average district to achieve any given performance standard simply because of the high needs among its students (as measured by poverty and LEP).

Figure 3 presents our education cost index and student need index. The full cost index, which reflects variation in both resource costs and student needs, has a value of 183 for New York City, which indicates that even if operating at an average efficiency level, New York City would have to spend 83 percent more than a district with average cost characteristics to reach the

Figure 3. Cost and student needs indexes for New York regions: 2000



SOURCE: Calculations by authors.

same level of student performance. In addition, child poverty and LEP levels in New York City raise the costs of achieving any adequacy target by 36 percent compared to a district with average poverty and LEP rates. This index also indicates that to reach the same student performance level as the average district, Yonkers would have to spend almost 80 percent more per pupil, and the upstate Big Three would have to spend 51 percent more per pupil. Moreover, student needs alone have about the same impact on required spending for Yonkers and for the Big Three as they do for New York City. The only other districts with costs significantly above average are the "downstate small cities," which have to pay above-average teacher salaries but do not have above-average student needs.

The typical approach for including student-need adjustment in aid formulas is to weight some students more heavily than others in the distribution of aid. If aid is distributed on a per pupil basis, then counting some types of students twice, for example, will assure that districts with these types of students receive more resources. While most states use the weighted-pupil approach to adjust for student needs, the origins of most of these weights remain obscure. At best, some are based on professional judgments about the extra costs associated with certain types of students; others appear to be ad hoc political compromises. Rarely are pupil weights determined through careful analysis of the actual relationship between student characteristics and costs. This is unfortunate, because an education cost model, such as the one estimated for this paper, can be used to calculate these weights.

We now illustrate this principle by using our cost model to calculate cost weights for both students in poverty and LEP students. The first and third columns of table 3 provide estimates of the extra costs associated with a student with certain characteristics in different types of districts. We find that each student in poverty requires a district to spend between $7,000 and $9,000 in additional resources to maintain the average performance level in the state. For LEP students, the extra costs are even higher, namely, in excess of $10,000 per student.

Pupil weights are calculated by dividing these additional costs by the spending required to bring non-LEP and poverty students up to average student performance. The resulting weights are presented in the second and fourth columns of table 3.[21] For both types of students the weights are approximately equal to 1. A weight of 1 can be interpreted as indicating that it is twice as expensive to bring a student of this type up to any given performance level as it is to bring other types of students up to that performance level. While there exists no definitive list of the pupil weights used by various states, the available evidence suggests that weights of 0.5 or below for at-risk students are the norm (Alexander and Salmon 1995, table 9.2). Our results indicate that the typical weight is far too low for New York State.

[21] See Duncombe (2002), appendix B, for a discussion of the methodology used to calculate pupil weights from cost function results.

Table 3. Cost impact of student needs: 1999–2000*

Regions	Extra cost per child in poverty (in dollars)	Child poverty weight	Extra cost per LEP student (in dollars)	LEP student weight
Downstate small cities	8,002	0.98	10,571	1.13
Downstate suburbs	7,941	0.98	10,343	1.10
New York City	7,945	0.98	10,762	1.15
Yonkers	7,606	0.94	11,008	1.18
The Big Three (upstate)	8,985	1.10	10,440	1.12
Upstate rural	8,086	0.99	10,170	1.09
Upstate small cities	7,715	0.95	10,260	1.10
Upstate suburbs	7,951	0.98	10,129	1.08

*Pupil weight is defined as the percent increase in costs associated with a student of a certain type. For example, the limited English proficient (LEP) student weight in New York City is 1.15. This indicates that bringing a typical LEP student in NYC up to an average performance level (160) will cost 115 percent more than a non-LEP student with otherwise similar characteristics.

SOURCE: Calculations by authors.

Estimating the Cost of Adequacy

The bottom line in developing a school finance system to support adequacy is determining what it will cost in each school district to reach the adequacy standard (assuming average efficiency). As explained earlier, we consider student performance standards of 130 and 160 to illustrate the effects of different adequacy standards on costs. For each performance standard, we first use our cost model to calculate the per pupil spending required to reach the standard in a district with average characteristics. This required per pupil spending in the average district is then multiplied by the cost index (divided by 100) to estimate the cost of adequacy in other districts.

To estimate the cost of adequacy with a resource standard, one must select a minimum bundle of resources and then estimate its cost. One technique for carrying out these steps is commonly called the "resource cost model" (RCM), which is a "bottom-up" approach to estimating the cost of adequacy (Chambers and Parish 1982; Management Analysis 1997). The RCM method involves designing prototypical classrooms, schools, and districts by asking professional educators what resources are required for a school to meet a particular standard. These resources are multiplied by resource prices to estimate the cost of resource adequacy in a prototypical district. The cost in the prototypical district is then multiplied by the resource cost index to estimate adequacy costs for other districts. For simplicity, we use the cost of adequacy in a district with average characteristics to identify a prototypical district's cost, instead of identifying a bundle of resources and determining its cost. We then multiply the spending required in this district by different resource cost indexes rather than by the full cost index.

Table 4 provides estimates of the per pupil spending required to reach different adequacy standards using different cost indexes for New York school districts. Comparisons are made to actual per pupil expenditures in the 1999–2000 fiscal year. As expected, we find that estimated required spending levels depend heavily on which standard and which cost index are used. With a standard of 130 and the teacher cost index produced for this study (New York teacher cost index), achieving adequacy requires significant increases over actual spending only in New York City and the large upstate cities (top panel of table 4).[22] Using the NCES teacher cost index, actual spending in New York City is estimated to already be adequate to reach a standard of 130. Using the 130 standard and a full cost index, which adjusts for resource prices and student needs, adequacy cannot be achieved without significant spending increases in all the large cities. We estimate, for example, that per pupil spending in New York City would have to increase by 56 percent, from $8,823 to $13,758.

With a standard of 130 and the teacher cost index produced for this study (New York teacher cost index), achieving adequacy requires significant increases over actual spending only in New York City and the large upstate cities.

If the more ambitious 160 standard is selected, then spending increases would be required in New York City and the upstate Big Three using any cost index. Using the NCES index, modest spending increases would have to occur in all the large cities except Yonkers and in the downstate small cities. When either the teacher cost index or the full cost index developed for this study is used, however, achieving adequacy would require sizeable spending increases in all the large cities and downstate small cities. Using the full cost index, for example, we estimate that spending would have to double in New York City, increase by 35 percent in Yonkers, and increase by 53 percent in the large upstate cities (the Big Three). Clearly, the level of the standard and the type of adjustment for cost differences across districts can have a large impact on the estimated costs of reaching an adequacy standard.

State Aid Formulas to Fund Adequacy

Basic operating aid formulas should be designed primarily to assist state governments in accomplishing their educational equity objectives. In most states, school districts differ widely in property wealth, income, resource prices, and student needs, and these

[22] Because regional averages are presented, the results in table 4 obscure the fact that some districts in other regions are estimated to require significant spending increases to reach the adequacy standard.

Table 4. Required spending per pupil for adequacy for different cost indexes*

Regions	1999–2000 per pupil expenditure	Standard of 130: New York teacher cost index (2000)	Standard of 130: NCES teacher cost index (1993)	Standard of 130: New York full cost index (2000) (all cost factors)
	In dollars			
State average (per pupil)	9,781	7,606	7,606	7,606
Downstate small cities	10,400	9,765	9,458	10,502
Downstate suburbs	11,723	8,642	9,038	8,573
New York City	8,823	11,701	8,597	13,758
Yonkers	12,437	11,569	9,430	13,384
The Big Three (upstate)	9,289	9,627	7,990	11,372
Upstate rural	9,509	6,842	6,693	7,181
Upstate small cities	9,335	7,902	7,357	8,054
Upstate suburbs	8,307	7,361	7,348	7,028
	2000 average performance index	**Standard of 160: New York teacher cost index (2000)**	**Standard of 160: NCES teacher cost index (1993)**	**Standard of 160: New York full cost index (2000) (all cost factors)**
		In dollars		
State average (per pupil)	160	9,532	9,532	9,532
Downstate small cities	148	12,236	11,852	13,161
Downstate suburbs	169	10,829	11,326	10,774
New York City	103	14,663	10,773	17,241
Yonkers	107	14,497	11,817	16,772
The Big Three (upstate)	96	12,036	10,012	14,251
Upstate rural	156	8,574	8,387	8,999
Upstate small cities	145	9,903	9,220	10,093
Upstate suburbs	160	9,224	9,208	8,808

*Calculated by estimating the cost in district with average cost to reach the given standard multiplied by the cost index (divided by 100).

NOTE: Large city districts are shaded.

SOURCE: Calculations by authors.

differences can lead to equally large differences in student performance. Most states have long recognized that variation in fiscal capacity can play an important role in creating large disparities in spending and student performance across districts. The equally significant impact on student performance of variation in resource costs and student needs has received far less attention. Educational cost indexes are important largely because they make it possible to design school aid formulas that effectively target resources to districts with the highest costs and greatest student needs. This section will illustrate how a cost index can be used in conjunction with fiscal capacity measures to develop simple but effective operating aid formulas for funding adequacy standards.[23]

Designing a Cost-Adjusted Foundation Formula

The majority of states use some form of a foundation grant system, which is designed to ensure that all districts meet some minimal standard.[24] For the most part, however, these systems express their standard in terms of spending, not student performance, so they do not bring the most disadvantaged districts up to a reasonable performance standard. In other words, these sys-

[23] This section draws heavily from Ladd and Yinger (1994), and Duncombe and Yinger (1998, 2000).

[24] For the most recent compilation of school finance systems, see U.S. Department of Education (2001).

tems are not consistent with the current focus on minimum adequacy standards for student performance.

In designing a *traditional foundation formula*, a state government needs to set a statewide minimum level of spending (E^*) and the minimum amount of local effort. The latter is often defined in terms of a state-determined minimum local property tax rate (t^*). The amount of revenue raised at this rate depends on the actual property values per pupil in a school district (V_i). Once these are defined, the per pupil aid (A_i) received by a district is simply the difference between the minimum spending level and the sum of the revenue raised by the district at the minimum local effort.[25] In short,

$$A_i = E^* - t^* V_i .$$

While the minimum spending level is constant statewide, the amount raised at the minimum level of local effort will vary across districts in direct proportion to their fiscal capacity. Thus, a foundation formula expects wealthier districts to contribute more taxes per pupil than poorer districts. If the traditional foundation formula is to successfully bring districts up to the minimum spending level, then a minimum level of local effort must be enforced; that is, no district should be allowed to levy a tax rate below t^*. Taken literally, this formula also could lead to "negative aid" or "recapture" of local property taxes in wealthy districts. In practice, however, the minimum aid amount is usually set to zero, and we use this aid design in the rest of our analysis.[26]

A traditional foundation formula with a minimum-tax-rate requirement should be successful in bringing spending in all districts up to the desired minimum level. However, the same minimum spending will be much more successful in raising student performance in some districts than in other districts, due in part to factors outside a district's control. Thus, a traditional foundation formula will generally not be successful in raising student performance in all districts up to an adequate performance level unless the minimum spending level is set very high, and the performance adequacy standard is set very low.

A traditional foundation formula will generally not raise student performance in all districts up to an adequate level unless the minimum spending level is very high and the performance adequacy standard is very low.

To convert a traditional foundation formula into a *cost-adjusted foundation formula* requires the basic tools that have been developed in this study.[27] First, the state must select an adequacy standard defined as a minimum level either of resources or of student performance, not simply of spending. Second, the adequacy standard must be converted into the spending required to meet the adequacy standard, an amount that obviously varies across districts because of variations in costs. One approach to these two steps is, of course, developed in this paper. Specifically, we estimate the cost of adequacy by multiplying the spending required in the district with average cost characteristics by a cost index. For a resource adequacy standard, the cost index reflects differences in the resource costs across the state that arise because higher salaries must be paid to attract teachers in some districts than in others. For a performance adequacy standard, the cost index captures both variation in resource prices and the greater quantity of inputs required in some districts because of higher student needs.

These steps make it possible to define cost-adjusted foundation aid per pupil, which is the difference between the spending per pupil necessary to reach the

[25] Some states consider other local revenue sources or certain types of federal aid as part of the local contribution. To minimize the required state aid, we counted all federal aid as part of the local effort.

[26] A few states have turned the local property tax into a state tax, which is an indirect way to include recapture in a foundation formula.

[27] This could also be called a performance-based foundation when the cost adjustment is for resource costs, sparsity, and student needs (our full cost index). The aid formula with full cost adjustment is designed to provide adequate resources for a district to have the opportunity to reach a particular performance standard (Duncombe and Yinger 2000). We have used the more general term, cost-adjusted foundation, to reflect either resource cost adjustment or full cost adjustment.

adequacy standard in a given district and the amount raised in the district by the minimum local tax effort and federal aid:

$$A_i = E^* c_i - t^* V_i ,$$

where E^* is required spending in the district with average characteristics, and c_i is an education cost index (centered on the district with average characteristics). The cost of adequacy calculated previously is represented by $E^* c_i$.

This cost-adjusted foundation formula is simple enough to be transparent to most school personnel and to the average voter; the logic of adjusting for costs is compelling and easy to understand. Moreover, the available evidence indicates that it would be effective. Duncombe and Yinger (1998) tested a number of aid formulas using New York data to determine which ones are the most effective in accomplishing specific educational equity objectives. They conclude:

> Our simulations of the impacts of . . . outcome-based [foundation] plans indicate that such plans can be an effective tool for promoting educational adequacy, at least when they include a required minimum tax rate. Indeed, by requiring contributions from local taxpayers, these plans can bring the vast majority of districts up to any standard policymakers select. The districts that remain below the standard are relatively inefficient. (p. 258)

Such a cost-adjusted foundation aid program would result in an increase of $6 billion, substantial aid increases in the larger cities, and significant aid cuts in many downstate districts and rural districts.

As with a traditional foundation formula, the success of a cost-adjusted foundation aid formula in significantly raising resources and student performance depends on enforcing a minimum-local-tax-rate provision and on the efficiency with which needy school districts use the additional resources.

Example of Aid Distribution With a Cost-Adjusted Foundation System

To illustrate a cost-adjusted foundation formula, we use the estimates of spending required to reach particular adequacy standards in table 4. In addition, we impose a minimum local effort equivalent to a property tax rate of $15 per $1,000 of market value, which is equal to the 1999–2000 state average.[28]

By design, a cost-adjusted foundation focuses aid on districts that face the most severe constraints in reaching the performance standard. However, table 5 makes it clear that the distribution of aid across districts depends significantly on the standard chosen and the type of cost adjustment made. This table compares the current aid distribution with aid that is distributed entirely through a cost-adjusted foundation formula. With a standard of 130 and the NCES teacher cost index, switching to a cost-adjusted foundation program would actually cut aid by over $2 billion, and even the large cities would receive little, if any, aid increases. In contrast, using the teacher cost index developed in this study would raise aid by $3 billion, and would result in large aid increases in the large cities. A cost-adjusted foundation aid program based on the full cost index developed in this study would result in an increase in the overall aid budget of $6 billion, substantial aid increases in the large cities, and significant aid cuts in many downstate districts and in rural districts.

Not surprisingly, the results for a performance standard at the current state average of 160 are more dramatic. In this case, switching to a cost-adjusted foundation aid program would result in substantial aid increases for the large cities using any cost index. Aid increases in New York City would range from about $2,000 per pupil (a 52 percent increase) with the NCES teacher cost index, to $8,500 per pupil (a 215

[28] Although this minimum effort is expressed as a property tax rate, the revenue could be raised through some other source, such as a local income tax. In this case, the local property tax rate would not have to be this high.

Table 5. Distribution of cost-adjusted foundation aid for different cost indexes[1]

Regions	2000–2001 per pupil school aid [2]	Standard of 130: New York teacher cost index (2000)	Standard of 130: NCES teacher cost index (1993)	Standard of 130: New York full cost index (2000) (all cost factors)
Total aid budget (in millions of dollars)	11,145	13,332	9,702	15,458
	In dollars			
State average (per pupil)	4,053	2,856	2,784	2,836
Downstate small cities	3,205	2,291	1,971	2,828
Downstate suburbs	2,419	1,312	1,531	1,204
New York City	3,949	6,922	3,817	8,979
Yonkers	3,112	5,837	3,697	7,652
The Big Three (upstate)	5,835	6,516	4,879	8,261
Upstate rural	5,203	3,099	2,877	3,397
Upstate small cities	4,937	4,321	3,800	4,496
Upstate suburbs	4,031	3,365	3,358	3,039

	Standard of 160: New York teacher cost index (2000)	Standard of 160: NCES teacher cost index (1993)	Standard of 160: New York full cost index (2000) (all cost factors)
Total aid budget (in millions of dollars)	19,762	15,223	22,395
	In dollars		
State average (per pupil)	4,448	4,440	4,397
Downstate small cities	4,340	3,887	5,145
Downstate suburbs	2,505	2,834	2,334
New York City	9,884	5,993	12,462
Yonkers	8,765	6,084	11,040
The Big Three (upstate)	8,953	6,901	11,140
Upstate rural	4,680	4,351	5,066
Upstate small cities	6,289	5,626	6,497
Upstate suburbs	5,133	5,108	4,716

[1]Cost-adjusted foundation aid is calculated by taking the estimated per pupil spending to reach the standard, and subtracting from it the required minimum local tax contribution (1.5 percent of property values) and federal aid. If the calculated aid is negative, it is set equal to 0.

[2]Includes all formula aid except Building Aid, Transportation Aid, and Reorganization Building Aid. Based on estimates of aid distribution in May 2001.

NOTE: Large city districts are shaded.

SOURCE: Calculations by authors.

percent increase) with the full cost index developed for this study. Aid increases would be even higher in Yonkers and would range from 18 percent to 91 percent in the other large cities. If one of the cost indexes developed in this study is used, aid increases would also occur in many small city districts. The significant aid increases in large urban districts would be financed from two sources: aid reductions, particularly in some rural and suburban districts, and large increases in state aid budgets (assuming minimum local effort is kept at the current state average of $15 per $1,000). For a standard of 160, the aid budget would increase between $4.1 billion (37 percent) and $11 billion (101 percent), depending on the cost index used.

Policy Choices in Financing an Adequate Education

Our estimates of the cost of achieving adequacy imply that adequacy cannot be achieved in New York without dramatic changes in the state's school finance system. In particular, spending levels in the high-need urban districts would have to rise significantly to provide the resources these districts need to bring their students up to any reasonable standard. Part of that required spending increase would cover higher teacher salaries so that these districts could compete with their suburbs for the best teachers. In addition, this required spending increase could fund class-size reductions, additional staff to support intense instruction in reading and math, and programs to address the social and health needs of at-risk children. When interpreting these large required spending increases, it is important to keep in mind that reaching the current statewide student average performance (160) in New York would require raising student performance in New York's large cities to levels that have seldom been achieved in large cities anywhere in the nation.

This study has presented estimates of the spending required for a district to have the opportunity to reach an adequacy standard. Another central policy question is how this spending should be financed. To answer this question, that is, to design a school finance system, state policymakers must address two key issues: the relative contributions of state and local governments and the impact of aid changes on school district efficiency.

State Versus Local Contribution to School Funding

The amount of state aid required to support an adequacy objective is directly related to two key policy decisions: how high to set the standard and how high to set the minimum local contribution. The advantage of a simple aid formula, such as the cost-adjusted foundation, is that it makes clear the impact of these two decisions on the required state aid budget. With any reasonable minimum local tax effort, the state aid budget would have to increase significantly to finance the adequacy standards presented in this report, and the only way to lower the required state aid budget for a given standard is to raise the required local tax effort. This analysis requires the state to enforce the minimum local tax effort as a legal requirement for receiving state aid. Otherwise, financially strapped districts, such as the large cities, will be tempted to cut local school tax rates and siphon state school aid into other services or tax cuts.[29] This type of behavior obviously undermines an adequacy standard.

The minimum local tax effort must be a legal requirement for receiving state aid. Otherwise, financially strapped districts, such as the large cities, will be tempted to cut local school tax rates and siphon state school aid into other services or tax cuts.

Before making a decision about the required minimum local tax effort, a state needs to consider several issues. The first issue is that there are some good arguments for keeping local property taxes low. While a well-administered property tax is not as regressive as is commonly believed, it can impose a significant burden on some low-income households. Moreover, a substantial property tax increase may undermine the competitiveness of a community, particularly a large city, in attracting or retaining residents and business. In our simulations, some of the largest required local tax increases would be in Buffalo and Syracuse and other upstate cities, which have experienced little economic growth in the last decade.

Some states have tried to minimize the burden of local property taxes without increasing state education aid by passing a property tax relief program, such as a homestead exemption. These programs help to ease the property tax burden on homeowners, but they often do not help renters or businesses. Moreover, these programs do not focus tax relief (and the state funds that support it) on homeowners in the school districts that need help the most. If a state is concerned about school finance equity, it should keep local property taxes low by increasing state aid to education, not by implementing direct property tax relief programs (Duncombe and Yinger 2001a).

An alternative to enforcing a minimum-tax-effort requirement is to use matching grants for operating aid.

[29] For a good review of the evidence on local tax effort in New York, where no minimum local effort is required, see New York State Education Department (2000). The study shows that several of the large upstate cities, Buffalo and Syracuse, used most of the school aid increases in the 1990s to lower school taxes rather than improve education.

A matching grant can be adjusted for fiscal capacity and educational costs, so that the state matching rate will be much higher in large cities and other high-need districts. These high matching rates are designed to encourage local spending on schools without requiring any particular local contribution. There is no guarantee, however, that high-need districts will significantly increase local tax effort in response to such a grant, let alone that they will increase local effort enough to achieve an adequate performance, however defined. In fact, a recent analysis using New York data shows that for any given state aid budget, even well-designed matching grants will not be as effective as cost-adjusted foundation grants in reaching an adequacy standard (Duncombe and Yinger 1998). While enforcing a minimum-local-effort provision may be politically unpopular with some local officials, it is a more cost-effective strategy than a matching grant for assuring adequate educational performance.

A final issue that arises in deciding on the state's share of education spending is that any increase in this share may lower productive efficiency in school districts. Indeed, some recent research based on New York data finds evidence supporting this possibility (Duncombe and Yinger 2000). This effect could arise, for example, because citizens are more apt to put pressure on school boards and superintendents, and thereby keep school districts efficient, when they must finance education through local taxes than when money for education is provided from state aid. A substantial increase in state aid to high-need districts could increase inefficiency by (1) putting pressure on already strained teacher labor markets; (2) encouraging rapid expansion of teacher salaries without accountability; (3) raising local construction costs through a large building program; and (4) straining the capability of district personnel to efficiently manage finances, to monitor private contracts, and to evaluate the success of existing or new programs.

A substantial increase in state aid to high-need districts could increase inefficiency.

These efficiency effects are not so large that they eliminate the benefits of higher state aid to school districts, but they do indicate that some of the benefits of state aid "leak out" in the form of higher inefficiency. As a result, states should be leery of setting the required minimum local tax effort too low.

Improve School Efficiency

An alternative approach to the issue of school district efficiency is to devise policies that boost school district efficiency directly, and thereby offset to some degree the efficiency-lowering effects of increased state aid. This approach is appealing, because it allows a state to minimize the required local tax effort for any given state aid budget (or to minimize state aid at any given required local tax effort), but it is also risky, because the impacts of direct policies to boost school district efficiency appear to be modest but are not well understood. Indeed, it is highly unlikely that any policies currently known could generate efficiency improvements sufficient to raise low-performing districts up to a reasonable adequacy standard. Nevertheless, these policies have the potential to make a significant positive contribution to a state education finance system, and in particular, to help high-need districts cope with large aid increases, and they are clearly worthy of more investigation.

Among the policies that appear most promising is technical assistance provided by a state education department on a variety of topics, including

- personnel functions, such as planning and forecasting future staffing needs, teacher recruitment and retention policies, and teacher evaluation methods, etc.;
- the use of program evaluation methods and student performance data to help guide program decisions made by school districts;
- the development of long-range capital plans, and evaluation of alternative capital financing options; and
- financial management practices, such as the use of cost accounting techniques, and school-based budgeting.

Another set of promising policies concerns the training of school district administrators. The recent selection of superintendents from noneducation backgrounds by some large-city districts may reflect in part the lack of training in basic management functions in many educational administration programs. State education departments can help shape the training that education administrators receive through both certification requirements and promoting innovative education management programs. While state governments may be loath to expand state education departments, particularly during an era of declining revenues, assisting districts to improve their management capacity may require an expanded staff and a diversification of specializations within these departments. In some cases, investing in increased capacity in state education departments to provide technical assistance in school management and improved administrator training programs may do as much to promote an adequacy standard as investing in higher state aid.

Conclusions

The trend toward higher student performance standards, which is backed by elected officials, education departments, and courts in many states, is clearly here to stay. It is time for state education finance systems to catch up, and in particular, to implement state aid systems that explicitly recognize that some districts must spend more than others to achieve any given performance standard.

The objective of this study is to assist state governments in developing this type of education finance system. In particular, we explain that an adequacy-based finance system involves three components. First, states must clearly define the type and level of the adequacy standard. They must decide, for example, whether to focus on resource adequacy or performance adequacy. As illustrated in the *CFE* decisions in New York, the distinction between these two types of standards is not always clarified by the courts; nevertheless, this distinction is crucial because it determines whether the state aid system must make adjustments for cross-district differences in student needs.

Second, a state government must estimate the spending required to reach adequacy in each district. This step is consistent with the court decisions in most states, which focus on resource or performance standards, not spending. This estimated cost of adequacy varies across districts in line with education costs. We illustrate the use of two statistical models, namely, a teacher wage equation and an education cost function, to develop education cost indexes. These indexes play a crucial role in estimating the cost of adequacy by measuring differences in resource costs and student needs across school districts. Using New York as a case study, we illustrate how the estimated cost of adequacy, particularly in large cities, is affected by choices about the stringency of the adequacy standard and the cost index. Given the importance of cost adjustments to estimating the cost of adequacy, all state governments would be well advised to support research on educational costs in their state and how these costs vary across districts.

Third, a state must develop a state aid formula that focuses aid on the districts with the highest costs and the lowest fiscal capacities. In New York, these districts include the large cities, which also have some of the lowest levels of student performance in the state. A simple modification of a traditional foundation formula to incorporate the estimated cost of adequacy provides a simple, but powerful aid system for reaching an adequacy standard. The simplicity of this formula helps to focus attention on the key questions in designing a school finance system: What is the adequacy standard? How should costs be accounted for? What should be the state share of educational spending?

Table A-1. Variables in a teacher wage equation

Variable name	Variable description	Source	Level	Mean[1]	Standard deviation[1]
Dependent variable:					
Lnsalary	Natural log of basic salary (no fringes or extra pay)	PMF	teacher	10.82305	0.30820
Discretionary factors					
Teacher quality measures:					
Lexper	Log of total teaching experience	PMF	teacher	2.38441	0.97610
Gradsch	1 if have Ph.D. or M.A.	PMF	teacher	0.74533	0.43568
Mathsci	1 if major assignment is in math or science	PMF	teacher	0.14258	0.34108
Sumcert	Share of assignments teacher has permanent certification.	PMF	teacher	0.88374	0.30213
MA_USN	1 if M.A. college is in *U.S. News* 1st tier	TCERT/*U.S. News*	teacher	0.03037	0.17161
BA_USN	1 if B.A. college is in *U.S. News* 1st tier	TCERT/*U.S. News*	teacher	0.04543	0.20824
Working condition measures:					
Lschenr	Log of enrollment in school where teacher teaches	IMF	school	6.61511	0.63250
Clsize	Average class size for teacher's assignments	PMF	teacher	23.75623	19.49249
Outcomes	Average district student performance	SED	district	141.52944	30.97875
Efficiency measures:[2]					
Aiddif	Difference in aid per dollar of income in this district and average district in similar need-capacity category	State aid	district	–0.01208	0.02283
Fvdif	Difference in per pupil property value in this district and average district in similar need-capacity category	State aid	district	13845	65578
Incdif	Difference in per pupil income in this district and average district in similar need-capacity category	State aid	district	–49726	251518
Factors outside district control					
Labor market variables:					
Lprofwage	Log of average county payroll for professional, scientific and technical sector (1997)	Census	county	10.59301	0.35579
Avgunemp	Average unemployment rate (1997–1999)	BLS	county	4.63639	1.44679
Tchshare	District share of county's full-time teachers	IMF	district	0.41629	0.34830
Working condition variables:					
Lpupden	Log of enrollment per square mile	IMF	district	5.83664	1.96455
Ldisenr	Log of district enrollment (average enrollment)	IMF	district	9.85490	2.65105
Flunres[3]	Adjusted 2-year average of percent K–6 enrollment receiving free lunch (1999–2000)	SED	district	–0.03499	0.26970
Avglep	2-year average of percent LEP[4] students (1999–2000)	SED	district	0.05142	0.05515
Crrate2	Violent crime rate for juveniles (under 18 years old) per 100,000 people (1998)	FBI	county	0.00275	0.00199

[1]Average of values associated with individual teachers. Sample size is 121,203. For county- or district-level variables, this is equivalent to a weighted average, weighted by the relative number of teachers. All data are for 2000 (or the 1999–2000 school year or fiscal year) unless otherwise noted.

[2]Need-capacity categories are defined by the New York State Education Department based on property, wealth, and student characteristics in the district.

[3]Residual from a regression of the average (1999–2000) share of free lunch students in elementary school regressed on the log of per pupil income and per pupil property values.

[4]"LEP" means limited English proficient.

SOURCE: PMF = New York State Education Department Personnel Master File; TCERT = New York State Education Department teacher certification data base; IMF = New York State Education Department Institutional Master File; State aid = New York State Education Department state aid files; Census = U.S. Bureau of the Census, 1997 Economic Census for Service Industries; BLS = U.S. Bureau of Labor Statistics, Local Area Unemployment Statistics; *U.S. News* = *U.S. News & World Report* rankings of undergraduate colleges; FBI = U.S. Department of Justice, FBI Uniform Crime Reporting system; and SED = Provided directly by New York State Education Department staff.

Table A-2. Descriptive statistics for variables in cost model: 1999–2000

Variables	Mean	Standard deviation
Per pupil spending[1]	9.106	0.231
Performance index	159.43	17.58
Efficiency variables[2]		
Full value	0.00000	623613
Aid	0.00000	0.02723
Income	0.00000	73010
Average teacher salary[3]	10.5137	0.1342
Percent child poverty (1997)[4]	0.1580	0.0978
2-year average LEP[5] students[4]	0.0129	0.0307
Enrollment classes[6]		
1,000–2,000 students	0.3201	0.4668
2,000–3,000 students	0.1608	0.3676
3,000–5,000 students	0.1431	0.3504
5,000–7,000 students	0.0605	0.2385
7,000–15,000 students	0.0516	0.2214
Over 15,000 students	0.0103	0.1012
Downstate small city or suburb	0.2589	0.4383

[1]Total spending without transportation, debt services, or tuition payments for students in private placements. Sample size is 678 school districts.

[2]Calculated as the difference between district value and the average in peer group. See text for discussion of peer group.

[3]For full-time teachers with 1 to 5 years of experience. Expressed as natural logarithm.

[4]Variables expressed as a percent of enrollment.

[5]"LEP" means limited English proficient.

[6]The base enrollment is 0 to 1,000 students. Variable equals 1 if district is this size, or else it equals 0.

SOURCE: Calculations by authors.

References

Alexander, K., and Salmon, R. (1995). *Public School Finance*. Boston: Allyn and Bacon.

Board of Education, Levittown Union Free School District v. Nyquist, 439 N.E.2d 359 (N.Y. 1982).

Boser, U. (2001, January 11). Pressure Without Support. *Education Week, Quality Counts 2001, 20,* 68–74.

Campaign for Fiscal Equity, Inc., v. State, 655 N.E.2d 661 (N.Y. 1995). (*"CFE1"*)

Campaign for Fiscal Equity, Inc., v. State, 719 N.Y.S.2d 475 (N.Y. Sup. Ct. 2001). (*"CFE2"*)

Campaign for Fiscal Equity, Inc., v. State, 744 N.Y.S.2d 130 (N.Y. App. Div. 2002). (*"CFE3"*)

Chambers, J. (1997). *A Technical Report on the Measurement of Geographic and Inflationary Differences in Public School Costs*. Prepared for the National Center for Education Statistics, Washington, DC.

Chambers, J., and Parish, T. (1982). *The Development of a Resource Cost Model Funding Base for Education Finance in Illinois*. Report prepared for the Illinois State Board of Education.

Downes, T., and Pogue, T. (1994). Adjusting School Aid Formulas for the Higher Cost of Educating Disadvantaged Students. *National Tax Journal, 47*(1): 89–110.

Duncombe, W. (2002). *Estimating The Cost of An Adequate Education in New York.* Center for Policy Research Working Paper, No. 44. Syracuse, New York: Syracuse University. Available: *http://cpr.maxwell.syr.edu/efap*

Duncombe, W., Miner, J., and Ruggiero, J. (1997). Empirical Evaluation of Bureaucratic Models of Inefficiency. *Public Choice, 93,* 1–18.

Duncombe, W., and Yinger, J. (1998). School Finance Reform: Aid Formulas and Equity Objectives. *National Tax Journal, 51*(2): 239–262.

Duncombe, W., and Yinger, J. (1999). Performance Standards and Educational Cost Indexes: You Can't Have One Without the Other. In H. Ladd, R. Chalk, and J. Hansen (Eds.), *Equity and Adequacy in Education Finance* (pp. 260–297). Washington, DC: National Academy Press.

Duncombe, W., and Yinger, J. (2000). Financing Higher Student Performance Standards: The Case of New York State. *Economics of Education Review, 19*(4): 363–386.

Duncombe, W., and Yinger, J. (2001a). Alternative Paths to Property Tax Relief. In W.E. Oates (Ed.), *Property Taxation and Local Government Finance* (pp. 243–394). Cambridge, MA: The Lincoln Institute of Land Policy.

Duncombe, W., and Yinger, J. (2001b). *Does School District Consolidation Cut Costs?* Center for Policy Research Working Paper No. 33, The Maxwell School. Syracuse, NY: Syracuse University.

Fowler, W., and Monk, D. (2001). *A Primer for Making Cost Adjustments in Education* (NCES 2001–323). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Guthrie, J., and Rothstein, R. (1999). Enabling "Adequacy" to Achieve Reality: Translating Adequacy Into State School Finance Distribution Arrangements. In H. Ladd, R. Chalk, and J. Hansen (Eds.), *Equity and Adequacy in Education Finance: Issues and Perspectives* (pp. 209–259). Washington, DC: National Academy Press.

Huber, P.J. (1967). The Behavior of Maximum Likelihood Estimates Under Non-Standard Conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1,* 221–233.

Ladd, H., and Yinger, J. (1994). The Case for Equalizing Aid. *National Tax Journal, 47*(1): 211–224.

Leibenstein, H. (1966). Allocative Efficiency vs. X-Efficiency. *American Economic Review, 56*(3): 392–415.

Lukemeyer, A. (2003). *Courts as Policymakers: School Finance Reform Litigation*. New York: LFB Scholarly Publishing.

Management Analysis & Planning Associates, L.L.C. (1997). *A Proposed Cost-Based Block Grant Model for Wyoming School Finance*. Report prepared for the Joint Appropriations Committee of the Wyoming Legislature.

Meyer, L., Orlofsky, G., Skinner, R., and Spicer, S. (2002, January 10). The State of the States. *Education Week*, *Quality Counts 2002*, *21,* 68–70.

New York State Education Department. (2000). *Analysis of School District Fiscal Response, 1993–94 to 1998–99.* Paper prepared by the State Aid Work Group, November 2000.

New York State Education Department. (2001). *The Regents Proposal on State Aid to School Districts for School Year 2002–03.* Albany: Author.

Niskanen, William A. (1971). *Bureaucracy and Representative Government*. Chicago: Aldine-Atherton.

Reschovsky, A., and Imazeki, J. (1997). The Development of School Finance Formulas to Guarantee the Provision of Adequate Education to Low-Income Students. In W.J. Fowler (Ed.), *Developments in School Finance, 1997* (NCES 98–212) (pp. 123–147). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Robelen, E. (2002, January 9). An ESEA Primer. *Education Week, 21,* 28–29.

U.S. Department of Education. (2001). *Public School Finance Programs of the United States and Canada: 1998–99*. Washington, DC: U.S. Department of Education.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica, 48*(3): 817–830.

Wyckoff, P. (1990). The Simple Analytics of Slack-Maximizing Bureaucracy. *Public Choice, 67*(1): 35–67.

Bond Ratings and Bond Insurance: Market and Empirical Analysis for School Districts

Mary H. Harris

Cabrini College

About the Author

Mary H. Harris is an Assistant Professor of Finance at Cabrini College in Radnor, Pennsylvania. She earned her Ph.D. in January 2002 at Lehigh University and was awarded the 2001 AEFA/NCES New Scholar's award for her dissertation topic of school district state funding programs and bond finance for capital expenditures. Prior to earning her doctorate degree, she spent 13 years in the banking industry specializing in the areas of accounting and finance. Her research interests are school finance and education policy.

Bond Ratings and Bond Insurance: Market and Empirical Analysis for School Districts

Mary H. Harris

Cabrini College

Introduction

A school district capital expenditure project typically begins with the issue of a bond to raise the required local revenue.[1] As the district prepares to issue a bond, it must determine whether or not to have it rated by an independent bond rating agency. This determination requires the district to do a cost-benefit analysis because rating agencies charge a fee to conduct bond ratings. The stated purpose of the rating is to provide potential bond buyers a measure of the risk that the district will default on future payments. In practice, the bond market uses the rating as information about the creditworthiness of the district, and this information in turn influences the yield (interest rate) at which the bond offering can be issued. The role of the rating agencies and the bond rating effect on the supply and demand of capital funds play a significant part in the process of determining the final yield on the bond issue.

Once a bond has been rated, the school district must determine whether to improve the bond rating by purchasing bond insurance. This determination requires an explicit cost-benefit analysis by the district because there is a premium charged to cover this guarantee. The bond insurance companies are key players in this part of the process, and evaluate the districts in a different way than the rating agencies.

Based on the sequential nature of this decisionmaking process, a three-stage empirical model was tested to estimate the significant factors at each stage (Harris and Munley 2002). A summary of the significant findings of those empirical results will be presented in this article, following the market analysis of the key players in the bond issuance process. The data sample used in the three-stage empirical models consisted of 148 bond issues, representing 10 different states.[2] Only bonds that were sold from July 1, 1993, through June

[1] In this paper we focus only on locally raised revenues. See Harris (2001) for a comprehensive discussion of the different state aid programs that exist to support capital spending by school districts, as well as the state-specific rules about referenda requirements for bond issues and overall debt limitations.

[2] These 10 states were chosen because as a group they provide a cross-sectional representation of the different institutional structures, in particular with regard to referenda requirements and debt limitation rules, that govern the bond issuing abilities of schools throughout the United States. See Harris (2001) for a complete discussion of why these states were selected.

30, 1994, and whose proceeds were used for capital expenditures were included in the sample.[3] This article will begin with a focus on the role of the bond rating agency in the process of issuing a bond.

The Role of the Rating Agency

Moody's Investors Service, Inc. (Moody's) and Standard & Poor's Corporation (S&P) are the two major rating agencies.[4] A list of their ratings and definitions can be found in table 1, panel A. They rate bonds upon request by the issuing district for an agreed upon fee. The fee is usually based on time and effort required to do the bond rating and averaged $7,000 per rating for the 1993–94 time period under observation in this research. Once the rating agency rates the new bond issue, it then continues to maintain and renew the rating until the bond has been redeemed. These rating agencies were originally developed to assist investors in comparing different bond issues by utilizing an easily recognizable set of symbols (Lamb and Rappaport 1980). The perception of investors is that all rating agencies grade all types of bond issues on the same criteria. Both Moody's and S&P have fundamental differences in their bond rating philosophies and policies. The two agencies do have similar criteria when evaluating the municipalities by examining the entity's debt level, economic base, and finances and management. The difference is that Moody's focuses more on the district's debt level and S&P focuses more on the district's economic base. The following section will analyze in greater detail the focus of the bond rating criteria of S&P and Moody's.

Standard and Poor's

S&P bases its bond rating criteria on four major factors: a district's economic base, financial position, debt

[3] We thus exclude issues used to refinance existing debt. The school year 1993–1994 was chosen because at the time this research was undertaken it was the most recent year for which all the data used in this empirical model were available.

[4] Fitch Investors is the third largest player in this market.

Table 1. Bond rating categories and yield averages

Panel A: Bond rating categories by Moody's and Standard & Poor's

Moody's rating	S&P's rating	Descriptions
Aaa	AAA	Highest Quality (low default risk)
Aa	AA	High Quality
A	A	Upper Medium Grade
Baa*	BBB*	Medium Grade
Ba	BB	Lower Medium Grade
B	B	Speculative
Caa	CCC or CC	Poor (high default risk)
Ca	C	Highly Speculative
C	D	Lowest Grade

Panel B: Moody's municipal bond yield averages over time from 1950 to 1994 (In percent)**

Month and year	Average municipal	Aaa	Aa	A	Baa
January 1950	2.03	1.61	1.82	2.23	2.46
January 1960	3.92	3.49	3.73	4.02	4.43
January 1970	6.74	6.38	6.60	6.88	7.13
January 1980	6.98	6.58	6.72	7.04	7.60
January 1990	7.02	6.81	6.93	7.01	7.35
January 1993	6.10	5.91	6.05	6.17	6.28
January 1994	5.33	5.14	5.19	5.36	5.60

*Bonds rated Baa (Moody's) and BBB (S&P) and above, are considered investment-grade bonds.

**The above yields are for long-term bonds.

SOURCE: Moody's Financial Government Manual, 1995, Volume 1.

levels, and administrative management strategies. Since the rating is an analysis of the district's long-term ability to pay, it must focus on both current and future economic conditions. Also, any state credit enhancement programs in which the state offers certain guarantees on debt payments may result in a certain minimum rating, usually an A, based on the strength of the state aid support (Hitchcock 1992). Insurance, if by a reputable insurance company, will also improve the rating.

Analysis of the economic base focuses on the district's wealth and income levels, employment by sector, government transfer payments, economic concentration and volatility, location in relation to other cities and employment centers, infrastructure, major area employers, and tax base composition. The analysis of financial position will depend on the level and volatility of operating revenues, expenditures, fund balance reserves, financial reports with proper accounting, and state revenue sources (Hitchcock 1992).

The analysis of debt levels will determine the size of the debt burden, the debt structure for the bond issue, and any future financing needs. This is accomplished through some debt ratios, including overlapping municipal debt to market value of property tax base, debt per capita, debt service expense to budget. The administrative management factor is the hardest to measure because it includes long-term administration, finance planning and goals, long-term capital improvement plans with sources and uses, future debt issuance plans, budgeting procedures, financial management policies, labor contracts, and pension policies (Hitchcock 1992).

Moody's

Moody's bases its bond rating criteria on the same four major factors: economic base, debt levels, financial position, and administrative management strategies. The analysis of the economic base concentrates on the regional economy, and more specifically, the expected tax revenues used to repay the bond obligations. Indicators of the economic stability of the region include unemployment level, diversity of employers, retail sales, number of new building permits, median income, and full valuation of taxable property per capita (Lipnick, Rattner, and Ebrahim 1999).

Analysis of the debt levels focuses on indicators such as the impact of the new debt on the existing credit quality of the school district, overlapping debt, and the structure of the bond issue. Analysis of the financial position factor focuses on the general fund balance as a percentage of revenues, and as an indicator of any potential revenue generating problems within the district (Lipnick, Rattner, and Ebrahim 1999). Analysis of the administrative management strategies is not always easy, but tends to become apparent from the analysis of the other three factors.

The national statistics show that the majority of rated bonds (58 percent) are rated by both Moody's and S&P.

Table 2 summarizes the national and regional market statistics based on this data sample of 148 bonds. The national statistics show that the majority of rated bonds (58 percent) are rated by both Moody's and S&P, followed by a rating only by Moody's of 29 percent. The regional statistics confirm similar results for the Southeast (bonds from Georgia, Kentucky, and Louisiana), the Southwest (bonds from Arizona and New Mexico), and the Plains (bonds from Kansas and Nebraska) categories of bonds in the sample. For smaller bond issues, it is sufficient to receive a bond rating from only one rating agency. However, the larger bond issues typically receive ratings from two or three rating companies. Many districts may stay with a particular rating agency for subsequent bond issues where updated information is required instead of an initial evaluation requiring past and present data information.

Bond Rating Effect on Demand for Capital Funds

It has been estimated that the bond rating is inversely related to the bond financing costs for a school district bond. If a school district receives a high bond rating on its issue, then the result will be a lower bond financing cost. When bond financing costs are reduced, the Local Education Agency (LEA) will have an increased demand for capital funds. The other impact the rating has on the bond issue is that it increases the

Table 2. Descriptive statistics for bond rating agencies

Panel A: National market statistics

Rating agency	Number of bonds	Percentage of total
Only Moody's	36	29
Only S&P	17	13
Both Moody's and S&P	**73**	**58**
Total	126	100

NOTE: Statistics are based on the 126 rated bonds in this sample. Information was only available on Moody's and S&P, but some of the bonds may also have been rated by Fitch.

Panel B: Regional market statistics

Rating agency	Southeast		Southwest		Plains	
	Number	Percent	Number	Percent	Number	Percent
Only Moody's	21	51	5	23	4	18
Only S&P	2	5	1	4	1	4
Both Moody's and S&P	**18**	**44**	**16**	**73**	**18**	**78**
Total	41	100	22	100	23	100

NOTE: The Mideast was biased toward S&P, while the Great Lakes and Far West were biased toward Moody's.

SOURCE: Information obtained from official bond statements for all bonds in data sample.

marketability of the issue. Once a bond has been rated, the district has a representation of its creditworthiness which can attract a larger pool of investors. This should also reduce the price of the bond which will, in turn, reduce the total bond financing costs. This is assuming that the bond rating is a good rating. The other factor to consider is whether or not the district should purchase a bond rating.

There are three principal reasons why a school district may decide not to obtain a rating for a bond issue. One is that the district anticipates that the issue will receive such a poor rating that not having any rating at all is just as attractive. Since a bad bond rating would hurt the bond's marketability and result in high bond financing costs, such a rating would lead to a reduced demand for capital funds by the district. A second is that the district expects the issue to be marketed locally, so that investors purchasing the bond already have sufficient information about the creditworthiness of the district, and thus there is no need to incur the extra expense of paying an independent agency to conduct a rating. A third is that the amount of debt being issued is small enough that the potential interest savings from a good rating are not large enough to offset the cost of obtaining one. See table 1, panel B, for a listing of the historical yield differentials by rating category for Moody's.

A bond rating is necessary when trying to attract nonlocal investors or institutional investors. For a small local school district bond issue, the lack of a rating might not significantly impact the bond financing costs, marketability, and demand for capital funds. However, for a large bond issue, the lack of a rating would cause the perception that the district had poor creditworthiness, and there would be a negative impact on the bond financing costs and marketability of that bond issue, resulting in a decrease in the demand for capital funds by that district. Based on a study in 1999 by Fitch IBCA, the education sector had the lowest cumulative default rate (.05 percent for $143,115,000 worth of defaulted par) of all sectors considered, indicating that school districts overall display a high level of creditworthiness (Litvack 1999).

A bond rating will affect the bond financing costs through the underwriting profit. If a bond receives an investment-grade rating, several underwriters will enter the bid process, which will keep the pricing competitive and the bond financing costs down for the school district. However, when the bond rating is of speculative or low investment grade, there will not be many underwriters interested in bidding, which will result in a higher price to compensate for the additional risk. The end result will be higher bond financing costs for the district. The next section will discuss how the bond rating impacts the investors in their supply of capital funds.

Bond Rating Effect on the Supply of Capital Funds

Due to federal and state regulations, many institutional investors, particularly banks and pension funds, are restricted to purchasing investment-grade bonds. Although the ratings are meant to be a relative measure, they are viewed more often as an absolute measure of credit quality. The federal government uses these ratings as standards for bank portfolio audits (Lamb and Rappaport 1980). If the rating agencies do not place a bond issue in the investment-grade category, the issue will be unable to attract the institutional investors required for a successful large bond issue. There has been some evidence that the standards at the rating agencies have tightened, and that there are fewer school district bonds in the investment-grade category now than there were 10 years ago.

Due to federal and state regulations, many institutional investors, particularly banks and pension funds, are restricted to purchasing investment-grade bonds.

Investors utilize the bond ratings as a measure of the default risk for the bond issue. If the bond rating is increased, then the risk is assumed to be reduced, which may increase the supply of capital funds. When bond rating changes are announced, the market price of that bond reacts immediately. If the bond rating is lowered, then the price of the bond will drop and the yields will increase. If the bond rating is raised, then the price of the bond will rise and the yields will drop. Therefore, the market is utilizing the bond ratings as important information on the risk of the district. As in the optimal portfolio theory, if the rating or risk of the district changes, then there should also be a change in the supply of capital funds.

Bond Insurance

Once a district has obtained a rating for a particular bond issue, it may proceed to issue the debt with this rating. It may decide, on the other hand, to purchase private bond insurance to improve the bond's rating. In issuing this type of policy, an insurance company agrees to stand behind the debt obligations of the district. This financial assurance will result in a higher rating for the bond issue—based on the credit quality of the insurance company. The original fee incurred by the district to obtain the initial bond rating is not impacted by the purchase of insurance. However, if insurance is purchased, the district must pay the additional cost of the insurance premium.

Insurance premiums are based on an assessment of the financial condition of the school district and the associated risk of default.[5] Because each insurance company uses its own assessment criteria to evaluate each district, a preliminary rating from a rating agency is not necessarily required. The insurance premiums are typically quoted as basis points (bp) for negotiated bond issues and converted to a flat dollar amount for competitive issues.[6] The basis point price is multiplied by the bond issue's total principal and interest to calculate the total fee. As of March 2002, a $20 million school district bond with an underlying (preliminary) A rating would have an average premium of between 15 and 25 basis points.[7] A district would

[5] This description of the insurance market for school district bond issues is based on conversations with industry officials.

[6] Bond issues in which school districts solicit bids from all interested underwriting firms are known as competitive bond issues. Bond issues in which school districts select one underwriter without soliciting competitive bids are known as negotiated bond issues.

[7] Therefore, the total premium would be .25 percent (or 25 bp) times the total principal plus interest of the bond issue. These averages were quoted by an insurance industry official in telephone conversation.

choose to purchase this type of insurance only if the higher rating would result in a reduction in the overall bond financing cost—net of the cost of the insurance premium. This would typically be the case if the reduction in interest cost is substantial, because the presence of insurance results in a steep upgrade in the bond's credit rating.

Some school districts that would benefit from purchasing insurance may not be able to do so. If the preliminary rating is below investment grade, then there might not be any insurance company of reputable credit quality willing to underwrite the policy. Also, if the size of the issue is too small, then the insurance company may refuse to undertake the risk associated with an unsuccessful marketing of the bond issue. These points highlight the different considerations when evaluating the creditworthiness of the district by the rating agencies and the insurance companies. Bond insurance is a long-term commitment, since the insurer cannot change the guarantee once it has been issued. On the other hand, the rating agencies can downgrade the bond ratings when a district's creditworthiness deteriorates. The three leaders in the municipal[8] bond insurance market are American Municipal Bond Assurance Corporation (Ambac), Municipal Bond Insurance Association (MBIA), and Financial Guaranty Insurance Co. (FGIC).[9]

Some school districts that would benefit from purchasing insurance may not be able to do so.

Ambac was founded in 1971 as a subsidiary of MGIC Investment Corp, and was the founder of the municipal bond insurance industry. In 1974, MBIA was formed as a consortium of four major insurance companies. In 1983, the third-largest player, FGIC, was formed. In 1975, Ambac and MBIA had a combined market share of 1.8 percent of municipal bonds issued for that year. By 1992, the percent of insured municipal issues reached over 30 percent of new bonds issued for that year. Ambac is a subsidiary of Ambac, Inc., which became a publicly traded company on the New York Stock Exchange in 1991.[10] FGIC is a GE Capital Company. Any district purchasing insurance from one of these companies will receive an automatic AAA rating from Moody's, S&P, and Fitch.[11] Although this rating is guaranteed by the approved insurance policy, the school district at this point does not know how this will translate into the final yield of the bond, which is also impacted by other factors.[12] To illustrate, table 1, panel B, shows the average differences in yields for the rating categories.

Table 3 provides national and regional statistics on the market share for the insurance companies based on this sample. On the national level, there was an even division in market share among the four insurance companies, FGIC, Ambac, MBIA, and Financial Security Assurance, Inc. (FSA).[13] However, there were market leaders on a regional basis. MBIA insured 62 percent of bonds in the Southeast; FGIC insured 67 percent of bonds in the Southwest; and FSA insured 63 percent of bonds in the Plains. Ambac consistently insured the second highest percentage of bonds in each regional segment. Table 3 also illustrates the percentage of rated bonds that are insured.

Empirical Results

A summary of the findings of the empirical research will focus only on the significant determinants of each stage of the bond rating process.[14] The following in-

8 Municipal refers to issues including all taxing entities such as cities, counties, school districts, townships, etc.

9 This information was found on the web site *http://www.southwest.msus.edu/RDIC/rdic1999/index.html*

10 Found on Ambac's web site, *http://www.ambac.com/aboutus_history.html*

11 Confirmed by Moody's as of March 27, 2002.

12 Harris (2001) presents an analysis of the determinants of market yields for school district bond issues.

13 Although FSA is not a current market leader, it was utilized along with the other three insurance companies during the 1993–94 period represented by this data set.

14 See Harris and Munley (2002) for a detailed explanation of the empirical analysis.

Table 3. Descriptive statistics for bond insurance companies

Panel A: National market statistics

Rating agency	Number of bonds	Percentage of total
FGIC	16	29
Ambac	14	25
MBIA	11	20
FSA	14	26
Total	55	100

NOTE: Based on the 126 rated bonds in this sample, 44 percent were insured.

Panel B: Regional market statistics

Rating agency	Southeast		Southwest		Plains	
	Number	Percent	Number	Percent	Number	Percent
FGIC	3	23	8	67	2	12
Ambac	2	15	3	25	4	25
MBIA	8	62	0	0	0	0
FSA	0	0	1	8	10	63
Total	13[1]	100	12[2]	100	16[3]	100

[1]The 13 insured bonds in the Southeast represent 32 percent of the rated bonds in the sample for this region.

[2]The 12 insured bonds in the Southwest represent 55 percent of the rated bonds in the sample for this region.

[3]The 16 insured bonds in the Plains represent 70 percent of the rated bonds in the sample for this region.

NOTE: The Mideast was divided evenly among Ambac, FGIC, and FSA, with 14 percent of its rated bonds insured. The Great Lakes was divided between MBIA and Ambac, with 40 percent insured; and the Far West was divided evenly among all four companies, with 53 percent insured.

SOURCE: Information obtained from official bond statements for all bonds in data sample.

dependent variables are utilized in the three estimating equations. They comprise measures found in other empirical studies of the bond rating process for both corporate and municipal issues.[15]

ECONOMIC AND DEMOGRAPHIC CHARACTERISTICS

INC—median household income of the district's population

ENROLL—number of students enrolled in the district

NW—percentage of the district's student population that is non-White

URBAN—binary variable equal to one for urban districts

RURAL—binary variable equal to one for rural districts[16]

FINANCIAL CHARACTERISTICS

LTE—local tax effort, defined as local tax revenue per student divided by median household income in the district

CASH—the district's end of the year cash fund balance

INGVT—intergovernmental revenues, defined as the percentage of a school district's total revenue coming from all federal and state grants[17]

[15] See, for example, Kaplan and Urwitz (1979); Aronson and Marsden (1980); Linda Ravelle (1990); Ziebell and Rivers (1992); and Moon and Stotsky (1993). See also Moody's Investors Service (2000) for their own discussion of the factors taken into account in their bond rating procedure.

[16] The omitted category serving as the reference for urban and rural districts is suburban school districts.

[17] In estimating the model we include both this variable and its squared term (INGVTSQ) to allow for the potential of a nonlinear relationship.

GRDBT—gross debt, defined as a district's per capita sum of long-term debt outstanding at end of year plus short-term debt at end of year plus the par value of the bond issue under study

BOND ISSUE CHARACTERISTICS

PAR—size of the bond issue, defined as its par value

RATED—binary variable equal to 1 if the bond is rated, and 0 if not. This is the dependent variable in the first-stage estimation equation.

INS—binary variable equal to 1 if the bond is insured and zero if not. This is the dependent variable in the second-stage estimation equation.

^INS—the predicted value of the probability of purchasing insurance from the second-stage equation. This is an independent variable in the third-stage estimation equation.

HIGH—binary variable equal to 1 if the bond is rated AAA or AA and equal to zero if the bond is rated A or Baa[18]

Table 4 presents summary statistics for these variables for the entire sample of 148 school district bond issues.[19] *In toto,* these variables capture a variety of factors that should enter the decision calculus of districts and rating agencies as they interact through the bond rating process. The mean values of the total data set are compared to the regional summary statistics and discussed throughout this section as applicable.

[18] For the bonds rated by both Moody's and S&P, only the Moody's ratings were used as the RATING category for this research. There were only a few circumstances where a bond was only rated by S&P, in which case those ratings were used.

[19] The table of correlation coefficients shows that only the values relating par and enrollment (.63) and the values relating local tax effort and intergovernmental revenues (–.68) exceed 0.5. The full table of correlation coefficients is available from the authors upon request.

Table 4. Descriptive statistics for 148 school district bond issues (mean values)

Variable	National	Southeast	Southwest	Plains
PAR (in dollars)	11,928,294	17,233,158	12,341,111	9,554,816
ENROLL	9,772	15,952	13,326	3,936
LTE (in dollars)	0.07	0.05	0.07	0.05
NW (in percent)	17.01	21.48	26.09	10.84
INC (in dollars)	31,768	26,058	27,138	29,168
INGVT (in percent)	56.9	67.16	62.45	64.33
CASH (in dollars)	986	845	1,075	1,074
GRDBT (in dollars)	844	496	971	1,230
Number of bonds issued	148	45	27	32

NOTE: These Southeast, Southwest, and Plains regions encompass 7 out of the 10 states that make up this data set. Variables are defined as follows:

PAR	size of the bond issue, defined as its par value
ENROLL	number of students enrolled in the district
LTE	local tax effort, defined as local tax revenue per student divided by median household income in the district
NW	percentage of the district's student population that is non-White
INC	median household income of the district's population
INGVT	intergovernmental revenues, defined as the percentage of a school district's total revenue coming from all federal and state grants
CASH	the district's end of the year cash fund balance
GRDBT	gross debt, defined as a district's per capita sum of long-term debt outstanding at end of year plus short-term debt at end of year plus the par value of the bond issue under study

SOURCE: Information obtained from U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), 1993–94; U.S. Bureau of the Census; and official bond statements.

Table 5 presents the mean values of the model's independent variables for the four bond rating classifications. It is interesting to note that the values for the three variables that measure a school district's economic vitality—INC, NW, INGVT—are most positive for those districts rated AA, not AAA. As the empirical analysis below will show, this is because most districts that obtain a AAA rating do so as a result of purchasing private insurance to upgrade an initially less favorable rating.

Table 6 presents maximum likelihood estimates for the three equations that make up the sequential bond rating model.[20] Column one of table 6 reports the results for whether or not the district chose to have the bond rated by a rating agency. Of the 148 bonds in this data sample, 126 (85 percent) were rated and 22 (15 percent) were not. The size of the bond issue (PAR) is positive and significant at the 1 percent level, which indicates that for large capital projects, the fixed cost associated with obtaining a credit rating can easily be offset by the savings potential of lower interest costs over the life of the bond, if the rating is favorable. The binary variable representing rural districts is negative and significant at the 10 percent level. This result suggests that because of their distance from financial markets, rural districts may be more likely to market their bonds locally, so that potential investors are already familiar with the district's financial situation and do not need the (costly) additional information that a credit rating provides. This variable may, however, also be picking up some of the effect of bond issue size, since most rural bond issues are smaller in par value than their urban or suburban counterparts. The final variable that is significant in this equation, also at the 10 percent level, is the district's end of the year cash fund balance. The sign of this estimated coefficient is negative. This result is somewhat surprising. Because it seems reasonable for rating agencies to interpret a large cash balance as a positive signal about a district's financial condition, we would expect this variable to increase the probability that a bond would be rated.

[20] Because all three dependent variables are dichotomous in nature, ordinary least squares regression will not yield efficient parameter estimates for these equations. The parameter estimates in this model are based on the LOGIT estimating procedure.

Table 5. Mean values for rating categories

Variable	AAA	AA	A	BAA	Unrated
PAR (in dollars)	13,957,000	21,506,000	7,867,000	1,458,000	2,111,000
ENROLL	8,267	15,493	9,650	11,510	4,422
LTE (in dollars)	0.06	0.09	0.06	0.02	0.07
NW (in percent)	16.74	13.61	15.89	13.99	25.47
INC (in dollars)	30,947	46,070	25,776	22,716	24,065
INGVT (in percent)	57.81	36.54	63.42	82.10	66.86
CASH (in dollars)	869	1,034	792	643	1,602
GRDBT (in dollars)	1,129	863	524	374	738

NOTE: Variables are defined as follows:

PAR	size of the bond issue, defined as its par value
ENROLL	number of students enrolled in the district
LTE	local tax effort, defined as local tax revenue per student divided by median household income in the district
NW	percentage of the district's student population that is non-White
INC	median household income of the district's population
INGVT	intergovernmental revenues, defined as the percentage of a school district's total revenue coming from all federal and state grants
CASH	the district's end of the year cash fund balance
GRDBT	gross debt, defined as a district's per capita sum of long-term debt outstanding at end of year plus short-term debt at end of year plus the par value of the bond issue under study

SOURCE: Information obtained from U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), 1993–94; U.S. Bureau of the Census; and official bond statements.

Table 6. LOGIT estimation results from three stages of bond rating model

	First Stage		Second Stage		Third Stage	
Stage	Rating obtained		Insurance purchased		Observed rating	
variable	RATED		INS		HIGH	
Intercept	–.792	(–.180)	–7.26	(–2.26)**	–1.25	(–.207)
PAR	+.000000573	(3.140)*	+.00000000761	(.560)	+.0000000469	(1.105)
ENROLL	–.0000116	(–.200)	–.0000344	(–1.33)	+.0000592	(1.62)***
LTE	+4.16	(.470)	–2.099	(–.296)	+22.2	(2.42)**
NW	–.7079	(–.570)	+2.68	(1.78)***	–5.99	(–1.65)***
INC	+.0000434	(.899)	–.0000321	(–.950)	+.000229	(3.49)*
INGVT	–1.52	(–.178)	+25.5	(3.13)*	–39.4	(–1.54)
INGVTSQ	+3.15	(.475)	–22.4	(–3.17)*	+35.3	(1.60)
CASH	–.000591	(–1.806)***	+.0000979	(.350)	–.00113	(–2.83)*
GRDBT	–.000276	(–.379)	+.00188	(3.77)*	–.00364	(–1.72)***
^INS	†		†		+15.90	(2.62)*
URBAN	+.00230	(.002)	+.187	(.300)	–.631	(–.742)
RURAL	–1.23	(–1.66)***	–.653	(–1.14)	+2.18	(2.10)**
Log-Likelihood	–38.54		–63.19		–42.70	
Chi-Square	47.34*		46.25*		70.52*	
Correctly Predicted (in percent)	89.90		77.8		81.8	

†Not applicable.

*Significant at the .01 level.

**Significant at the .05 level.

***Significant at the .10 level.

NOTE: Numbers in parentheses are *t*-statistics for the null hypothesis of no association. Variables are defined as follows:

- RATED binary variable equal to 1 if the bond is rated and 0 if not. This is the dependent variable in the first-stage estimation equation.
- INS binary variable equal to 1 if the bond is insured, and 0 if not. This is the dependent variable in the second-stage estimation equation.
- HIGH binary variable equal to 1 if the bond is rated AAA or AA, and equal to 0 if the bond is rated A or Baa
- PAR size of the bond issue, defined as its par value
- ENROLL number of students enrolled in the district
- LTE local tax effort, defined as local tax revenue per student divided by median household income in the district
- NW percentage of the district's student population that is non-White
- INC median household income of the district's population
- INGVT intergovernmental revenues, defined as the percentage of a school district's total revenue coming from all federal and state grants
- INGVTSQ this variable represents the squared INGVT variable and was utilized due to the non-linear relationship with the dependent variable
- CASH the district's end of the year cash fund balance
- GRDBT gross debt, defined as a district's per capita sum of long-term debt outstanding at end of year plus short-term debt at end of year plus the par value of the bond issue under study
- ^INS the predicted value of the probability of purchasing insurance from the second-stage equation. This is an independent variable in the third-stage estimation equation.
- URBAN binary variable equal to 1 for urban districts
- RURAL binary variable equal to 1 for rural districts

SOURCE: Data from U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), 1993–94; U.S. Bureau of the Census; and official bond statements were utilized to run this LOGIT estimation model.

Table 7 describes national and regional market statistics on the percentage of bonds that are rated and not rated. On a national level, 85 percent of the bonds in this sample were rated. The regional statistics were varied, with the Southeast bonds rated 91 percent of the time, the Plains 72 percent, and the Southwest 81 percent. Based on the regional descriptive statistics in table 4, the Southeast has the highest mean PAR value, followed by the Southwest and then the Plains. The higher PAR value is consistent with the empirical findings that the PAR value is the leading determinant for having a bond rated.

Column two of table 6 reports the results for whether or not the school district chose to purchase private insurance to upgrade the initial rating. As illustrated in table 3, of the 126 bonds in this data sample, insurance was purchased on 55 (44 percent) of them and not purchased on 71 (56 percent) of them.

The percentage of the school district's population that is non-White is positive and significant at the 10 percent level, consistent with the notion that poor districts have, on average, a higher portion of residents who are non-White. Gross debt is also positive and significant at the 1 percent level, which suggests that districts already carrying high levels of debt are more likely to need the help of insurance to upgrade a bond rating to finance additional capital projects.

The effect of a greater reliance on total intergovernmental grants in a school district's financial profile on its need to purchase insurance is particularly interesting. Although the programs in place by the federal government and the states to provide funds to local school districts are many and varied, the overall pattern is clearly need based. This variable, therefore, presents a comprehensive measure of school district financial need that depends on a variety of social and economic characteristics of the district. The coefficient of its squared term is negative and significant at the 1 percent level. These combined results suggest that the propensity of districts to purchase private insurance increases with this measure of district "neediness" but at a decreasing rate, reaches a maximum, and then decreases. A possible explanation for this result is that the neediest of districts receive sufficient support from the state programs, described above, that have been put in place to help them secure better bond ratings so that they do not need to purchase private insurance.

Table 7. Descriptive statistics for bond rating decisions

Panel A: National market statistics

Decision	Number of bonds	Percentage of total
Rated	126	85
Not Rated	**22**	**15**
Total	148	100

Panel B: Regional market statistics

Rating agency	Southeast		Southwest		Plains	
Decision	Number	Percent	Number	Percent	Number	Percent
Rated	41	91	22	81	23	72
Not Rated	**4**	**9**	**5**	**19**	**9**	**28**
Total	45	100	27	100	32	100

NOTE: Mideast had 95 percent rated bonds, Great Lakes had 80 percent rated bonds, and Far West had 82 percent rated bonds.
SOURCE: Information obtained from official bond statements for all bonds in data sample.

According to table 3, panel A, 44 percent of bonds were insured on a national level. As illustrated in table 3, panel B, the percentages of bonds insured varied by region. For example, 32 percent were insured in the Southeast, 55 percent in the Southwest, and 70 percent in the Plains. The Plains had the highest mean gross debt level, and its mean gross level was more than twice that of the Southeast's. The reliance on intergovernmental funding was above the national mean and similar for all three regions. Although the Plains had the lowest percentage of non-White population, the increased financial leverage would have led to a lower bond rating without insurance. Therefore, these regional results also represent intuitive and consistent results when compared to the empirical findings for this stage of the bond rating process.

Column three of table 6 reports the third-stage results for whether a district's bond receives a high (AAA or AA) rating (dependent variable equal to 1) or a medium (A or Baa) rating (dependent variable equal to zero). Of the 126 bonds in this data sample, 87 (69 percent) received a AAA or AA rating while 39 (31 percent) received the medium rating.[21] As noted above, districts that purchase insurance from a reputable underwriter automatically receive a AAA rating from Moody's, S&P, and Fitch.[22] Whether or not a district purchases private insurance, therefore, is a clear determinant of the rating it receives and must be included in any model where rating is the dependent variable. The purchase of insurance is a choice variable of the school official, however, and thus endogenous to the model. For this reason we include in this equation the instrumental variable that is the predicted value of the probability of purchasing insurance from the previous equation. The coefficient of this variable, ^INS, is positive and significant at the 1 percent level. Several other variables that also exhibited statistical significance are described in the next paragraph.

Apparently, rating agencies also give weight to how much residents are currently willing to provide support for district spending.

Total enrollment is positive and marginally significant at the 10 percent level, which suggests that school district size may be an advantage in the bond rating process. The percentage of the district's population that is non-White is negative and significant at the 10 percent level, even though we have controlled for the fact that a higher non-White percentage of the population increases the likelihood of purchasing insurance. Not surprisingly, median household income is positive and significant at the 1 percent level. The rating agencies clearly take the ability of a district's population to make future tax payments into account when providing a bond rating. The existing level of local tax effort is also positive and significant at about the 1 percent level. Apparently, rating agencies also give weight to how much residents are currently willing to provide support for district spending. It is worthwhile to note that the par value of the bond being issued exhibits no statistically significant effect in explaining the rating that a bond receives. The size of an issue apparently does not influence how it will be rated.

Gross debt is negative and significant at the 10 percent level. Districts already carrying high levels of debt are viewed as posing a greater risk of defaulting on new issues than those not so encumbered. The size of a district's year-end cash fund balance is negative and significant at the 5 percent level, which suggests that cash-rich districts actually receive a less favorable rating. This result is counterintuitive, as was the result for this variable in the first equation for whether or not a district had a bond issue rated in the first place. Either this is not a correct interpretation of what this variable actually measures within the context of a district's financial profile, or the role that it plays in the bond rating process is too complicated for this basic model to capture.

On a regional basis, 39 percent of the Southeast's bonds received the high rating, and 61 percent received the

[21] The yield differential in January 1994 between AAA and AA bonds was .05; between AA and A bonds was .17; and between A and Baa bonds was .24. Only 8 of the 126 rated bonds in our sample, however, are classified Baa. This, together with Moody's own designation of high (AAA or AA) versus its medium (A or Baa) investment grade, provides the rationale for the dichotomous rating classification used here.

[22] In this data sample, all of the issuers of AAA rated bonds had purchased insurance to secure the rating.

medium rating. This is consistent with the fact that their percentage of insured bonds was the lowest at 32 percent. According to the other descriptive statistics, the mean household income was the lowest, as well as the average tax rate for the Southeast region. The Southwest had 68 percent of its bonds in the high rating category and 32 percent in the medium rating category. The Plains had 88 percent of its bonds in the high rating category and 12 percent in the medium rating category. This is also consistent with this region insuring the highest number of bonds (70 percent) and maintaining the highest mean household income when compared to the other two regions. The Plains also had a non-White population percentage of 11 percent compared to the Southeast's percentage of 21 percent, which is also a significant determinant of the bond rating.

Conclusion

The market information on the key players in the rating agencies and insurance companies provides interesting results at both the national and regional level. The empirical findings confirm the significance of analyzing each stage of the bond rating process when considering a rating assigned to a specific bond issue.

The first stage explains whether or not districts choose to obtain a rating for a new bond issue. The finding that par value of the bond issue is the most statistically significant determinant in this decision supports the supposition that districts may choose not to have their bonds rated due to the transaction costs of the rating process, and not necessarily because the districts are of poor credit quality. Likewise, the finding that rural districts are more likely not to obtain a rating lends credence to the supposition that a local marketing strategy may also be a contributing factor in this decision. There were no significant indicators that poor credit quality was a factor in choosing whether or not to have the bond rated. These findings were consistent with the descriptive statistics analyzed on a regional basis.

The second stage of the bond rating model explains the choice to purchase insurance. In contrast to the results from the first equation, at this stage, measures of district economic need and financial danger signals are all that seem to matter. A higher concentration of non-White population and a higher proportion of district revenues derived from intergovernmental grants both raise the likelihood that a district will purchase insurance to enhance a bond rating. A greater amount of pre-existing debt also increases this likelihood. Again, these findings were consistent with the descriptive statistics analyzed on a regional basis.

The final stage of the bond rating model deals with estimating the rating categories themselves. Due to limitations of the sample data and prior information on bond yield differentials, we classify the ratings as either high (AAA or AA) or medium (A or Baa) investment quality. The districts that are the strongest financially appear to be in the AA category. This is expected since it is not cost advantageous for a district with a bond rated initially AA to purchase insurance to improve the rating to AAA in exchange for a slightly lower interest cost. The descriptive statistics for the AAA bonds in this sample suggest that it is the purchase of private insurance coverage that leads to the high rating, not the financial condition of a school district. Nonetheless, it is in this third equation that we find the greatest number of statistically significant explanatory variables. Measures of the underlying economic condition of the district's population, the district's financial profile, and characteristics of the bond issue itself all appear, in ways that make intuitive sense, to contribute to a rating agency's determination of creditworthiness. Several of the descriptive statistics on a regional basis proved to be consistent with the results from this stage. Further extensive research into regional variations is worth pursuing.

Appendix: Regional Classifications

The following states are included in the empirical data samples and were classified into regions. The empirical comparisons are based on these regional classifications:

Southeast	Southwest	Plains	Mideast	Great Lakes	Far West
Georgia	Arizona	Kansas	New Jersey	Illinois	Oregon
Kentucky	New Mexico	Nebraska			
Louisiana					

References

Aronson, J.R., and Marsden, J.R. (1980). Duplicating Moody's Municipal Credit Ratings. *Public Finance Quarterly, 8*(1): 97–106.

Harris, M.H. (2001). School District Bond Finance: A Study on Bond Ratings and State Capital Aid Programs (Ph.D. diss., Lehigh University).

Harris, M.H., and Munley, V.G. (2002). The Sequence of Decisions Facing School District Officials in the Bond Rating Process: A Multistage Model. *Journal of Education Finance, 28*(1).

Hitchcock, D. (1992). How School District Bonds Are Rated, *School Business Affairs, February,* 15–18.

Kaplan, R.S., and Urwitz, G. (1979). Statistical Models of Bond Ratings: A Methodological Inquiry. *Journal of Business, 52*(2): 231–261.

Lamb, R., and Rappaport, S.P. (1980). *Municipal Bonds: The Comprehensive Review of Tax-Exempt Securities and Public Finance.* New York: McGraw-Hill Book Company.

Lipnick, L.H., Rattner, Y., and Ebrahim, L. (1999). The Determinants of Municipal Credit Quality. *Government Finance Review, 15*(6): 35–41.

Litvack, D. (1999). Measuring Municipal Default Risk. *Government Finance Review, 15*(6): 19–21.

Moody's Investors Service (2000). "Rating Approach," *http://www.Moodys.com*.

Moon, C.G., and Stotsky, J.G. (1993). Municipal Bond Rating Analysis: Sample Selectivity and Simultaneous Equations Bias. *Regional Science and Urban Economics, 23,* 29–50.

Ravelle, L. (1990). A Model of Revenue Bond Ratings for Water and Sewer Authorities in Pennsylvania (Ph.D. diss., Lehigh University).

Ziebell, M.T., and Rivers, M.J. (1992). The Decision to Rate or Not to Rate: The Case of Municipal Bonds. *Journal of Economics and Business, 44*(4): 301–16.

GASB Update

Randal Finden

Governmental Accounting Standards Board

About the Author

Randal Finden has been a project manager with the Governmental Accounting Standards Board (GASB) in Norwalk, Connecticut, for 7 years. He has worked on the reporting model project, Statement No. 34, *Basic Financial Statements—and Management's Discussion and Analysis—for State and Local Governments*, and its Implementation Guide and Statement No. 38, *Certain Financial Statement Note Disclosures.* He continues to work on financial instrument projects, such as the current Deposit and Investment Risks project, derivatives, and hedging. He guided Statement 31, *Accounting and Financial Reporting for Certain Investments and for External Investment Pools,* from exposure to its publication.

He has spoken at numerous conferences and workshops and authored many articles. Mr. Finden has been an editorial advisor to the *Journal of Accountancy* and a past contributing author and reviewer of Harcourt Brace's *Miller Governmental GAAP Guide* and newsletter. Before coming to the GASB, Mr. Finden served 18 years in the Washington State Auditor's Office developing accounting guidelines for local governments. He is a member of the American Institute of CPAs. He is a graduate of California State University, Sacramento.

GASB Update

Randal Finden

Governmental Accounting Standards Board

Introduction

This article summarizes my remarks at the 2002 National Center for Education Statistics (NCES) Summer Data Conference. Although other topics could have been addressed, this article and the discussion at the conference are limited to comments about the Governmental Accounting Standards Board's (GASB's) new reporting model, affiliated organizations, the deposit and investment risk project, and the other postemployment benefits project.

The New Reporting Model

GASB Statement No. 34, *Basic Financial Statements—and Management's Discussion and Analysis—for State and Local Governments,* substantially changes the format of school financial statements. Because it has been a topic of many earlier NCES sessions, its details are beyond the scope of this article. In broad terms, school financial statements will now include a statement of net assets and a statement of activities. Government-wide statements in this format will, for the first time, provide a means to evaluate a government's overall financial position and its activities on an economic basis.

As part of school financial statements, a management's discussion and analysis will be required that describe a school's financial events in a narrative format. Finally, revenue and expenditure information (that is, fund-based information) that has been available historically will continue with little change. Much more information is available at our web site (*http://www.gasb.org*), including links to the financial statements of schools that have implemented the Statement's requirements early.

Affiliated Organizations

The Board issued Statement No. 39, *Determining Whether Certain Organizations Are Component Units,* which addresses the relationship of affiliated organizations to schools, in May 2002. Affiliated organizations include parent-teacher-student organizations, booster clubs, and foundations. Development of this Statement has been a difficult project, which included two exposure drafts. The chief concern of the Board has been the creation of a standard that captures for inclusion the large organizations, such as university and large school district foundations, while at the same time excluding the many very small organizations that are associated with most schools.

The Board settled on three rules to establish inclusion. An included organization will most likely be reported as a discretely presented component unit:

> Organizations that are legally separate, tax-exempt entities and that meet all of the following criteria should be discretely presented as component units. These criteria are:
>
> 1. The economic resources received or held by the separate organization are entirely or almost entirely for the direct benefit of the primary government, its component units, or its constituents.
>
> 2. The primary government, or its component units, is entitled to, or has the ability to otherwise access, a majority of the economic resources received or held by the separate organization.
>
> 3. The economic resources received or held by an individual organization that the specific primary government, or its component units, is entitled to, or has the ability to otherwise access, are significant to that primary government. [Excerpt from GASB Summary of Statement No. 39]

Note that the focus is not limited to financial resources, but includes economic resources. An organization that benefits multiple organizations, such as United Way, would not be considered for inclusion.

Deposit and Investment Risk Project

The Board issued a proposed Statement (also referred to as an Exposure Draft, or ED), *Deposit and Investment Risk Disclosures,* in June 2002.* This project includes a review of existing deposit and investment disclosure requirements. It is important to emphasize that this project is not the result of a round of depository or investment losses, although there have been some recent, localized depository losses. Instead, the finance literature, investment professionals, and financial statement users have been consulted to determine the effectiveness of existing requirements. New disclosures are proposed and existing requirements are reduced.

The Board held a public hearing on the ED on October 1, 2002. People who are interested in information about the hearing should check the GASB's web site at *http://www.gasb.org.*

Interest Rate Risk

Because investments are reported at fair value, as interest rates change, investment fair values vary. Interest rate risk is the risk that changes in interest rates may adversely affect an investment's fair value. Generally, the longer an investment's maturity, the greater its exposure to interest rate risk. In practice there are several ways of managing interest rate risk. The Board identified five methods, proposing that any one of the five may be selected:

Specific Identification. The easiest method and the most attractive to small governments would be a list of investments, their maturities, and any call options, as shown in the following example:

As of December 31, 2003, the district's pooled investments were as follows:

Investment	Fair value	Maturity
State investment pool	$1,506,980	6.5 months average
U.S. Treasury bills	452,980	January 2004
Federal National Mortgage Association	282,230	March 2004
ABC Corporation commercial paper	350,000	January 2004
DEF Corporation bonds	50,000	March 2005
Total	**2,642,190**	

Weighted average maturity. When there are numerous individual investments and investment types, listing every investment is usually not practical. Summarization methods are available. The weighted average maturity method summarizes investments by type and dollar-weights their maturities, as shown in the following example:

* In March 2003, the GASB approved a final statement—Statement No. 40, *Deposit and Investment Risk Disclosures.* Although the basic premise of the proposed standard was unchanged, there were substantive changes to the proposal. The final statement should be consulted for an understanding of the final disclosure requirements.

As of December 31, 2003, the city had the following investments:

Investment type	Fair value	Weighted average maturity (months)
Repurchase agreements	$215,000	0.20
U.S. Treasury	119,864	4.21
U.S. agencies	23,614	3.21
Certificates of deposit	55,493	12.85
Corporate bonds	160,500	17.48
Total	**574,471**	**7.21**

Duration. Similar to the weighted average maturity method, duration uses discounted present values of cash flows. There are different versions of duration in practice: Macaulay, modified, and effective. Any version would be acceptable.

Simulation models. For sophisticated governments, the proposed standard permits use of simulation models. Changes in a portfolio's fair value would be estimated given hypothetical changes in interest rates, as shown in the following example:

The following table summarizes the estimated effects of hypothetical increases in interest rates on investment fair values. It assumes that the increases occur immediately and uniformly to each type of investment. The hypothetical changes in market interest rates do not reflect what could be deemed best- or worst-case scenarios. Variations in market interest rates could produce significant changes in the timing of repayments due to any prepayment options. For these reasons, actual results might differ from those reflected in the table.

	Fair value
December 31, 2002	**$3,000,000**
Impact on Fair Value of Basis Point Increase of:	
100 Points	2,915,979
200 Points	2,834,756
300 Points	2,756,226

Segmented time distributions. In our field test, the most popular method was depicting maturities by aggregating by selected time periods, as shown in the following example:

As of December 31, 2003, the city had the following investment types and maturities. (Amounts are in thousands.)

		Investment maturities (in years)			
Investment type	**Fair value**	**Less than 1**	**1–5**	**6–10**	**More than 10**
Repurchase agreements	$15,000	$15,000			
U.S. Treasury	119,864	62,000	$42,864	$15,000	
U.S. agencies	23,614		15,000	8,614	
Commercial paper	50,697	50,697			
Corporate bonds	35,493		10,000	20,493	$5,000
Mutual bond funds	74,420	74,420			
Certificates of deposit	1,000		1,000		
Total	**320,088**	**202,117**	**68,864**	**44,107**	**5,000**

Highly Sensitive Investments. In the context of interest rate risk, some investments are highly sensitive to changes in interest rates. The Board felt that these required additional disclosure. These are investments with contract terms that make the investments' fair values highly sensitive to interest rate changes. Because new securities are constantly being brought to market, the concept is deliberately without specifics. However, examples are provided: inverse floaters; an investment's variable coupons, which include a multiplier (for example, coupon varies by 125 percent of London Interbank Offered Rate); and collateralized mortgage obligations, interest-only or residual tranches.

Credit Risk

Credit risk is the possibility that an issuer or other counterparty will not fulfill its obligations. It is most commonly realized when a debtor defaults on its debt. Many, but not all, governments are limited by statute to corporate debt that has the highest two credit ratings (for example, Aaa or AAA) issued by nationally recognized statistical rating organizations. These organizations—for example, Fitch, Moody's Investors Service, and Standard & Poor's—enjoy special status in federal securities law. The proposed standard would require disclosure of credit ratings as of the end of the reporting period. Investments with the guarantee of the U.S. government would be exempt from this disclosure requirement. If an investment is not rated, the disclosure would indicate that fact.

Custodial Credit Risk

The Board reconsidered existing custodial credit risk requirements. Depository custodial credit risk is the risk of loss arising from the inability to recover deposits if the financial institution fails. Investment custodial credit risk is the risk of loss arising from the inability to recover the value of investment or collateral securities in the possession of an outside party if the counterparty to the transaction fails.

Investment custodial credit risk is the risk of loss arising from the inability to recover the value of investment or collateral securities.

Custodial credit risk requirements were established in 1986 when the Board issued Statement No. 3, *Deposits with Financial Institutions, Investments (including Repurchase Agreements), and Reverse Repurchase Agreements.* Some believe, however, that although in its day Statement No. 3 was very helpful, reduced custodial credit losses, in part the result of increased regulation, argue for reduced disclosures. The federal Government Securities Act of 1986 required all government securities dealers to be supervised, reducing the number of unregulated dealers.

The Board's proposed changes would not eliminate custodial credit risk disclosures. However, such disclosures would be reduced to what has become the "category 3" deposits and investments. Category 3 deposits are uninsured and uncollateralized. Category 3 investments are uninsured investments that are held by either the counterparty or the counterparty's trust department, but not in the name of the government.

Concentration of Credit Risk

When a portfolio has a disproportionate investment in one debtor, there is an above-the-ordinary amount of credit risk. Additional disclosures would be required in this situation. The proposed standard indicates that concentration risk is present when 5 percent of a portfolio's investments are in any one issuer. Investments issued or guaranteed by the U.S. government would not be included in this calculation.

Foreign Currency Risk

Investments not denominated in U.S. dollars expose the investment to foreign currency risk. The proposed standard would require the currency denomination to be disclosed. Like interest rate risk, the longer the term-to-maturity of the investment, the greater the exposure to foreign currency risk. Time horizon disclosures, similar to interest rate disclosures, would be required for debt investments.

Investment Policies

Investment policies indicate a government's risk tolerance. For example, even though a portfolio's weighted average maturity is less than 1 year, is the government willing to go out 2 or more years? Investment policies are an indication. The Exposure Draft would require disclosure only of those policies that are relevant to the risks that are disclosed. In other words, the focus would be on risk first, followed by any relevant investment policies. Because investment policies commonly include topics not directly relevant to deposit or investment risks, the Board wished to avoid unnecessary disclosures.

Level of Detail

The new reporting model provides new guidance on the level of disclosure. Consistent with the general requirements of Statement No. 34:

> The disclosures required by this Statement should focus on the governmental activities, business-type activities, major funds, nonmajor funds in the aggregate, internal service funds in the aggregate, and fiduciary fund types of the primary government. [GASB Statement No. 34, paragraph 5]

Effective Date

The proposed standard would be effective for fiscal years beginning after June 15, 2004. Earlier application would be encouraged.

Other Postemployment Benefits

A current project of the Board is the Other Postemployment Benefits (OPEB) project. OPEB refers to postemployment benefits other than retirement benefits, such as medical, dental, vision, and hearing benefits. OPEB also refers to other forms of postemployment benefits when they are provided separately from a pension plan. Examples include life insurance and long-term care.

The Board has tentatively concluded that postemployment benefits are part of compensation for services rendered by employees. That is, they are part of an exchange transaction. (Someone has done something in expectation of payment.) Benefits are earned, and obligations accrue or accumulate, during employment. However, payment is deferred until after employment. The tentative decision is to require recognition of OPEB costs generally over an employee's years of service. Expressed in oversimplified terms, current OPEB expenses would be determined by projecting total OPEB liability, discounting using present value principles, and then allocating current costs and prior service costs over an employee's years of service, not to exceed 30 years. This methodology would be consistent with current pension reporting requirements.

Required note disclosures would include relevant information about the accrued OPEB obligation and the progress made in funding the plan.

The GASB staff is working on the possibility of a method for small employers to calculate OPEB liability and expense without the use of an actuary. This spreadsheet-based method would simplify the selection and handling of assumptions, such as longevity, life expectancies, and health care cost trends. The Board currently is field-testing the feasibility of the method as an alternative to actuarial valuations.

Views expressed are those of the author and are not official representations of the GASB. The views of the GASB are established after due process.

High Performance of Minority Students in DoDEA Schools: Lessons for America's Public Schools

Claire E. Smrekar
Debra E. Owens
Vanderbilt University

About the Authors

Claire E. Smrekar is Associate Professor of Public Policy and Education at Peabody College, Vanderbilt University. She received her doctorate in Educational Administration and Policy Analysis from Stanford University. Her research focuses upon the social context of education and public policy.

Debra E. Owens is a former Assistant Superintendent of schools with over 20 years of experience in public education and administration. She is currently a Ph.D. candidate in Education Policy in the Department of Leadership, Policy, and Organizations, Peabody College, Vanderbilt University. Her research focuses upon state and national education policy.

High Performance of Minority Students in DoDEA Schools: Lessons for America's Public Schools

Claire E. Smrekar
Debra E. Owens
Vanderbilt University

Introduction

The debate among scholars continues regarding the degree to which an array of economic, social, cultural, psychological, and institutional factors influences student achievement. Most agree that differences in students' performance on standardized tests are related to a set of school conditions and family characteristics (Alexander and Entwisle 1996; Jencks and Phillips 1998; Natriello, McDill, and Pallas 1990).

These issues and concerns create a complicated achievement equation. Many critical questions persist regarding how and why school environments (e.g., academic rigor, academic grouping, teacher quality, teacher expectations) and family environments (e.g., family income, level and quality of parental education, occupational status, family size and structure, parents' perceived self-efficacy, parenting style) differentially impact student achievement. We agree that this issue is complex, controversial, and unresolved.

DoDEA System: Background Briefer

The U.S. military established elementary, middle, and high schools for the children of service men and women overseas and in the United States shortly after World War II. The schools were organized in two distinct but similar systems: The Department of Defense Dependents Schools (DoDDS) overseas, and the Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) in the United States. (Almost all the DDESS schools are located in the southeastern United States.) The two systems were united under the umbrella Department of Defense Education Activity (DoDEA) in 1994. Military personnel must live on base in order to enroll their dependents in the DDESS system.

Today, the Department of Defense Education Activity (DoDEA) enrolls approximately 112,000 students in schools located in the United States (DDESS system) and overseas (DoDDS system)—or about the same number of students as the Charlotte-Mecklenburg, North Carolina, school district, or the state of North Dakota, with the same proportion of minority students as in New York state schools (average 40 percent minority) (see table 1). Another approximately 600,000 school-age children of U.S. active military personnel attend school in one of the more than 600 civilian public school districts located near military installations in the continental United States (Military Family Resource Center 2001).

This study, conducted by researchers at the Peabody Center for Education Policy, was designed to provide a descriptive analysis of one school system—the DoDEA schools—that has demonstrated high minority student achievement and high achievement overall, as measured by the 1998 National Assessment of Educational Progress (NAEP) (see table 2). The study focuses upon a set of systemwide governance structures, school conditions, instructional policies, teacher characteristics, and administrative practices that are related to a school's capacity (Cohen and Ball 1999; Cohen and Spillane 1992; Corcoran 1995; Ferguson 1998) to produce student learning. We also explore school climate to examine whether or not DoDEA schools reflect the properties of "communally organized" schools that recent research suggests produce higher achievement (Bryk and Driscoll 1988; Bryk, Lee, and Holland 1993; Coleman and Hoffer 1987).

We visited a total of 15 middle schools located in 10 different school districts across the United States, Germany, and Japan (5 domestic districts and 5 overseas districts). The schools in our study reflect the average minority student enrollment for the DoDDS and DDESS systems, although some schools in the study reflect a higher than average minority enrollment. We deliberately selected schools that vary somewhat in size, mobility rates, installation deployment and training patterns, pay and rank composition of parents, and in the percentage of children who are eligible for free and reduced-price lunch. Students from these schools have parents in various military services (see table 3). This selection decision produced a group of schools that reflects the depth, range, and diversity of DoDDS and DDESS schools.

Approximately 130 interviews were completed over the course of the 4-month data collection period. We conducted in-depth interviews with the principal and language arts teachers at each school. At each district, military commanders and liaisons, curriculum specialists, assistant superintendents, and the superintendent were interviewed. Our interest focused upon issues of financial support, resource allocation, personnel recruitment and selection, teacher quality, accountability, leadership styles, program diversity, and academic policy priorities.

Table 1. Number of districts, schools, teachers, and students in the DoDEA[1] system, 2000–01

	DoDDS[2]	DDESS[3]	Total
Districts	12	12	24
Schools	157	70	227
Teachers	5,747	3,675	9,422
Students	77,912	34,294	112,206

[1]DoDEA is Department of Defense Education Activity, the umbrella agency under which DoDDS and DDESS were united in 1994.

[2]DoDDS is the Department of Defense Dependents Schools located overseas.

[3]DDESS is the Department of Defense Domestic Dependents Elementary and Secondary Schools located in the United States.

SOURCE: Department of Defense Education Activity, Annual Accountability Profiles, 2000–01.

Table 2. Ranking of DoDEA[1] minority students on the 1998 NAEP compared to other states

	Eighth-grade reading	Eighth-grade writing
DoDDS[2] African American students	First	Second
DoDDS Hispanic students	Second	First
DDESS[3] African American students	Second	First
DDESS Hispanic students	First	First

[1]DoDEA is Department of Defense Education Activity, the umbrella agency under which DoDDS and DDESS were united in 1994.

[2]DoDDS is the Department of Defense Dependents Schools located overseas.

[3]DDESS is the Department of Defense Domestic Dependents Elementary and Secondary Schools located in the United States.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 Writing Assessment and 1998 Reading Assessment.

Table 3. Percentage makeup of DoDEA[1] student population by sponsor's service, 2000–01

Sponsor's Service	DoDDS[2] (percent)	DDESS[3] (percent)
Army	35	60
Navy	14	10
Marine Corps	6	16
Air Force	32	7
National Guard	0	1
Civilian	12	5

[1]DoDEA is Department of Defense Education Activity, the umbrella agency under which DoDDS and DDESS were united in 1994.

[2]DoDDS is the Department of Defense Dependents Schools located overseas.

[3]DDESS is the Department of Defense Domestic Dependents Elementary and Secondary Schools located in the United States.

SOURCE: Department of Defense Education Activity, Annual Accountability Profiles, 2000–01.

In addition to interviews, we collected an array of school and district documents, including curriculum guides and benchmark standards, staff development plans, accountability reports, student/family demographic data, school handbooks, and parent newsletters. At each military installation, we collected information on housing, health services, recreation services, and social services on the base. An extensive school and base tour, and multiple classroom observations (e.g., language arts classes, computer classes, industrial drawing) were an essential part of each full-day site visit.

Findings

What Accounts for These High Levels of Performance?

"Your study is looking at why minority students do better. I think the answer to that question is that all our students do better. There are no 'minority' students here." (Teacher, DoDEA, 2001)

I. Assessment Systems in DoDEA

"We get benchmarks and we determine what assessments we want to use. You need a few leaders that are curriculum-minded and change-minded in the school to make it work." (Teacher, DoDEA, 2001)

Our analysis of test scores across multiple assessment systems confirms that students in the DoDEA schools perform at a high achievement level in reading and writing. DoDEA uses three assessments systems to measure reading and writing achievement of DoDEA students: their NAEP scores along with their scores on the Terra Nova Achievement Test and the DoDEA Writing Assessment.

NAEP

NAEP, sponsored by the U.S. Department of Education and administered by the National Center for Education Statistics (NCES), is known as the "Nation's Report Card" and is the only continuing assessment of the nation's students in various subject areas (Pellegrino, Jones, and Mitchell 1999). Since 1969, periodic assessments have been conducted in reading, mathematics, science, writing, U.S. history, civics, geography, and the arts. The population is sampled for the three types of NAEP: national NAEP, state NAEP, and long-term NAEP.

Our study focuses upon the state NAEP data that provide state/jurisdiction comparisons but cannot be disaggregated by individual students or schools. However, results of the state NAEP can be disaggregated by subgroups (e.g., race). In 1998, between 40 and 44 jurisdictions voluntarily participated in the state NAEP reading and writing assessments.

NAEP results have been increasingly used by policymakers as indicators of the nation's educational health (Pellegrino, Jones, and Mitchell 1999). NAEP policy is determined by the nonpartisan, independent National Assessment Governing Board. NAEP has earned the reputation as the nation's best measure of student achievement over time.

The 1998 NAEP scores in reading and writing for DoDEA schools are impressively high (see table 4).* Although this study focuses upon the performance of minority students in DoDEA schools, the overall DoDEA NAEP results are worthy of review as well. In writing, students in DDESS were second in the nation, with 38 percent scoring at or above the *Proficient* level; DoDDS students were fourth in the nation, with 31 percent scoring at or above the *Proficient* level. This compares favorably to the national rate of 24 percent scoring at or above the *Proficient* level. In reading, only two states had a greater percentage of students at or above the *Proficient* level than either DDESS (37 percent) or DoDDS (36 percent). Again, DoDEA schools are scoring well above the nation in the number of *Proficient* or above level readers.

Black and Hispanic students in DoDEA schools rank either first or second in the nation for reading and writing (see table 2). Although achievement gaps in writing exist between White students and minority students in DoDEA schools, the gaps between Black and White students and Hispanic and White students are far smaller in DoDEA schools than in the nationwide comparative results in writing (see table 5). All groups in DoDEA schools report higher scaled scores in writing than the national averages. Note that the DDESS system has a much higher percentage of Black students and Hispanic students than the national average.

Reading scores for DoDEA students show a similar pattern of above-average scores and smaller racial gaps (see table 6). There is no significant gap in reading between White and Hispanic students in DDESS schools. However, a gap exists between Black and White students in DDESS schools. Again, all reading scaled scores are higher than the national average for comparable groups.

When a parent's level of education is considered, a greater percentage of students in DoDEA schools are scoring at or above the *Proficient* level in writing and reading than are students nationwide in all but one category (see table 7). Among the category of students with a parent who has "some education after high school," 37 percent of DDESS students obtained writing scores at or above the *Proficient* level, compared to only 19 percent of the students in the national sample. In this same category, 40 percent of DDESS students obtained reading scores at or above the *Proficient* level, compared to 35 percent of the

* The term *Proficient* refers to one of the three achievement levels used by NAEP: *Basic*, *Proficient,* and *Advanced*. *Basic* denotes partial mastery of the knowledge and skills that are fundamental for proficient work at each grade level in a particular subject matter; *Proficient* represents solid academic performance—students reaching this level have demonstrated competency over the subject matter; *Advanced* signifies superior performance on NAEP in the particular subject matter.

Table 4. Percent of eighth-graders in top achievement levels on 1998 NAEP writing and reading assessments in DoDEA[1] schools and public schools in selected states
(In percent)

	Writing			Reading		
Jurisdiction	***Proficient***	***Advanced***	**Total**	***Proficient***	***Advanced***	**Total**
Connecticut	40	5	45	38	4	42
DDESS[2]	32	6	**38**	31	6	**37**
Maine	30	2	32	38	4	42
DoDDS[3]	30	1	**31**	33	3	**36**
Nation[4]	23	1	24	28	2	30

[1]DoDEA is Department of Defense Education Activity, the umbrella agency under which DoDDS and DDESS were united in 1994.

[2]DDESS is the Department of Defense Domestic Dependents Elementary and Secondary Schools located in the United States.

[3]DoDDS is the Department of Defense Dependents Schools located overseas.

[4]The national results are based on the national assessment sample, which includes the DoDEA schools.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 Writing Assessment and 1998 Reading Assessment.

Table 5. Average scaled scores on the 1998 NAEP writing assessment, by race/ethnicity

Race/ethnicity	Percent of total population	Average scale score	White versus Black gap	White versus Hispanic gap
DDESS[1]				
White	41	167	†	†
Black	26	150	17	†
Hispanic	27	153	†	14
DoDDS[2]				
White	46	161	†	†
Black	18	148	13	†
Hispanic	17	153	†	8
Nation[3]				
White	65	156	†	†
Black	15	130	26	†
Hispanic	14	129	†	27

†Not applicable.

[1]DDESS is the Department of Defense Domestic Dependents Elementary and Secondary Schools located in the United States.

[2]DoDDS is the Department of Defense Dependents Schools located overseas.

[3]The national results are based on the national assessment sample, which includes the DoDEA schools.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 Writing Assessment.

Table 6. Average eighth-grade scaled scores on the 1998 NAEP reading assessment, by race/ethnicity

Race/ethnicity	Percent of total population	Average scale score	White versus Black gap	White versus Hispanic gap
DDESS[1]				
White	42	279	†	†
Black	26	253	26	†
Hispanic	27	268	†	11*
DoDDS[2]				
White	46	276	†	†
Black	19	259	17	†
Hispanic	15	263	†	13
Nation[3]				
White	66	270	†	†
Black	15	241	31	†
Hispanic	14	243	†	33

*Not significantly different.

†Not applicable.

[1]DDESS is the Department of Defense Domestic Dependents Elementary and Secondary Schools located in the United States.

[2]DoDDS is the Department of Defense Dependents Schools located overseas.

[3]The national results are based on the national assessment sample, which includes the DoDEA schools.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 Reading Assessment.

students in the national sample. This level ("some education after high school") describes the educational backgrounds of the majority of enlisted men and women with children in DoDEA schools; enlisted men and women account for approximately 80 percent of all DoDEA parents. (See Section IV, "Policy Recommendations," of this report for a complete description of the educational levels of parents in the DoDEA system.)

Terra Nova

The pattern of high or above-average student achievement with some persistent gaps between White and minority students is reflected in the annual Comprehensive Test of Basic Skills Fifth Edition (CTBS/5) Terra Nova, Multiple Assessment (Terra Nova), an achievement test administered to all DoDEA students in grades 3 through 11 (see table 8) since the 1997–1998 school year. The Terra Nova is a norm-referenced achievement test that is typically administered to all students in a state. Scores are reported at the student, school, district, and national levels. When a system has more than 25 percent of its students in the top quarter, it is considered to be performing above the national quarter.

A greater percentage of DoDEA students score in the top quarter of the Terra Nova than the nation as a whole. In the 2000 Terra Nova, 39 percent of all students in DoDEA schools scored in the top quarter in language arts and 32 percent of all DoDEA students scored in the top quarter in reading, while only 7 percent and 8

Table 7. Percentage of eighth-grade students at or above the *Proficient* level on the 1998 NAEP writing and reading assessments, by parents' level of education
(In percent)

System	Did not finish high school	Graduated from high school	Some education after high school	Graduated from college	I don't know
Writing					
Nation[1]	6	18	19	33	3
DDESS[2]	**	**	37	39	**
DoDDS[3]	**	23	29	35	**
Reading					
Nation	—	21	35	42	—
DDESS	—	32	40	39	—
DoDDS	—	23	39	43	—

—Not available.

**Sample size is insufficient to permit reliable estimate.

[1]The national results are based on the national assessment sample, which includes the DoDEA schools.

[2]DDESS is the Department of Defense Domestic Dependents Elementary and Secondary Schools located in the United States.

[3]DoDDS is the Department of Defense Dependents Schools located overseas.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 Writing Assessment and 1998 Reading Assessment.

Table 8. Percentage of eighth-grade DoDEA* students in top and bottom quarters of the 2000 Terra Nova Tests in language arts and reading

	All DoDEA students		White		African American		Hispanic	
2000 Terra Nova Tests	Percentage of students in top quarter	Percentage of students in bottom quarter	Percentage of students in top quarter	Percentage of students in bottom quarter	Percentage of students in top quarter	Percentage of students in bottom quarter	Percentage of students in top quarter	Percentage of students in bottom quarter
Language arts	39	7	48	5	26	12	29	8
Reading	32	8	41	5	16	16	22	10

*DoDEA is Department of Defense Education Activity, the umbrella agency under which DDESS and DoDSS were united in 1994.

SOURCE: Department of Defense Education Activity (DoDEA), Office of System Accountability.

percent, respectively, scored in the bottom quarter. In table 8, the scores for DoDEA minority students (subgroups) are compared with the scores for all DoDEA students, as represented by the quarters established by the total, national sample.

The 2000 Terra Nova Tests for eighth-graders in language arts show that 48 percent of White students score in the top quarter of the nation, while 26 percent and 29 percent of African American and Hispanic, respectively, fall into this top quarter. In the bottom quarter, 12 percent of African Americans and 8 percent of Hispanics score in this bottom range, while only 5 percent of White students score in the lowest quarter.

In reading, fewer minority students score in the top quarter and more in the bottom quarter than in language arts. Sixteen percent of African American students and 22 percent of Hispanic students had a score in the top quarter, while 16 percent African American and 10 percent Hispanic scored in the bottom quarter.

DoDEA Writing Assessment

In 2000, 74 percent of the eighth-graders scored *Distinguished* or *Proficient* on the DoDEA Writing Assessment (see table 9). Only 5 percent were in the lowest category, *Novice*. The DoDEA Writing Assessment is a hand-scored essay that was patterned after the National Writing Project. Each student's writing level is assessed, but there are no national norms for this assessment. The percentage of students scoring at each level are aggregated by school, district, and system.

Students across all subgroups achieve at high levels on the DoDEA Writing Assessment although there are persistent achievement gaps between White students and minority students. Overall, between 67 percent and 77 percent of students score at or above the *Proficient* level in writing. The DoDEA Writing Assessment results mirror the superior writing performance of DoDEA students on the NAEP Writing exam.

Use of Standardized Test Scores

Studies of accountability systems highlight the focus on student performance (Fuhrman 1999). Schools, not school districts, are often the unit of improvement within individual school improvement plans. Setting student achievement goals for a school provides a focus for work and increases energy devoted to instruction. Effective educational systems clarify content standards and utilize tests that are consistent with content standards (CORE 1998). The alignment among standards and assessment in DoDEA schools follows research recommendations.

The mission of DoDEA is to "provide, in military communities worldwide, exemplary education programs that inspire and prepare all students for success in a global environment" (DoDEA Community Strategic Plan 2001). Toward this goal, DoDEA monitors student progress and promotes student success regularly through the use of standardized tests. The policy of assessing the achievement of DoDEA students every year through standardized testing is required by law (see Annual Education Assessment 2000 and Systemwide Assessment Program 2001). DoDEA outlines three purposes of standardized tests (DoDEA Assessment Program 2001):

1. To help teachers determine the strengths and needs of students in order to work with them to improve their individual academic skills.

Table 9. Performance-level percentages of 2000 DoDEA* writing assessment of eighth-grade students, by race/ethnicity

Performance level	Percent of all DoDEA students	Percent of White DoDEA students	Percent of Black DoDEA students	Percent of Hispanic DoDEA students
Distinguished	33	38	25	27
Proficient	41	39	42	44
Apprentice	21	18	25	23
Novice	5	5	8	6
Proficient or above	74	77	67	71

*DoDEA is Department of Defense Education Activity, the umbrella agency under which DDESS and DoDSS were united in 1994.
SOURCE: Department of Defense Education Activity (DoDEA), Office of System Accountability.

2. To let parents know how their children scored in different academic subjects.

3. To provide accountability for DoDEA schools. The testing information used to help determine how well DoDEA schools work includes norm-referenced tests, which provide a comparison of the basic skills of DoDEA students with the achievements of students in non-DoDEA stateside schools.

Our analysis of DoDEA's testing measures provides compelling evidence of the benefits of linking assessment with strategic intervention for school improvement and systemwide reform. DoDEA assessment systems are embedded within a coherent policy structure that links instructional goals with accountability systems, supported by professional training and development programs.

The process begins with information exchange that is systematic, clear, and comprehensive. First, DoDEA provides every school and each district with detailed assessment results. These test results are analyzed in multiple ways, including performance by grade level, by gender, and by race. Each school utilizes the school improvement plan process to analyze student improvement needs, select student improvement goals, develop assessment instruments such as pre- and post-tests, identify interventions, monitor change in student performance, and document change in student performance. Student outcomes are specifically tied to strategic goals. Staff training and curricular intervention are coordinated with the school site plan. The ability and disposition to notice and act on instructional problems, and to use resources to help solve problems, are critical elements of school improvement (Cohen and Ball 1999). DoDEA exemplifies these school improvement principles.

First, DoDEA provides every school and each district with detailed assessment results. These test results are analyzed in multiple ways, including performance by grade level, by gender, and by race.

A vivid illustration of the alignment across curriculum standards, assessment, and training is the writing program and DoDEA Writing Assessment. Clear standards and expectations for writing performance are outlined in the DoDEA Standards Book for faculty and staff. The DoDEA Writing Assessment reflects the standards of writing performance outlined in the curricular goals. By effectively "teaching to the test," writing instruction embraces the performance standards for good writing evaluated by the DoDEA Writing Assessment. In this sense, the writing assessment becomes the means *and* the ends.

Professional development activities focus upon effective writing instruction and student performance. School and overall district performance levels in writing are reviewed each year by the Office of Accountability in DoDEA headquarters. Threshold levels of achievement are established by DoDEA, and districts are held accountable for meeting these established benchmarks (e.g., 75 percent of all students must perform at or above the *Proficient* level on the DoDEA Writing Assessment). In the end, if support and intervention do not improve writing achievement, other additional resources and assistance will be provided for schools. Recently, a handful of DoDEA sites, known as Framework Schools, were targeted for intervention and enhanced resources after years of low student achievement. Teachers met to identify problems and develop comprehensive reform proposals, assisted by a DoDEA instructional leader. These teams focused upon a package of resources and training that were essential for school improvement and enhanced student performance. The problem identification process and strategic planning utilized in the Framework School program suggest a bottom-up/top-down linked strategy that produces positive results for students and staff alike.

II. Financial Resources

Financial resources are vital to an effective school system. The DoDEA schools enjoy sufficient funding to implement instructional goals. The cost per pupil is higher than the national average. Teacher salaries are competitive and schools are well staffed. Instruction is enhanced by state-of-the-art equipment and well-maintained facilities.

Costs per Pupil

DoDEA has a higher average per-pupil expenditure than the national average. For 1998–1999, DoDEA reports that the total expenditures per pupil were $8,908. The overseas system has higher expenditures ($9,055) than the domestic system ($8,586). According to DoDEA, the funding levels for both systems are higher than the national average of $7,290. However, these reported figures may be misleading.

DoDEA schools' costs are not directly comparable to U.S. public schools' costs due to an important difference in organizational structure between DoDEA schools and their civilian counterparts. DoDEA schools lack the support of a state department of education. Public school districts in the United States are under the jurisdiction of a state and obtain various forms of support from state departments of education. This support is not reflected in the per-pupil expenses of United States public school districts. DoDEA headquarters provides many services to its districts, but these costs are added to DoDEA schools' per-pupil expenditures. When DoDEA district superintendents were interviewed, many reported that DoDEA headquarters provided services similar to state departments of education.

Teacher Salaries

Highly qualified teachers are considered to be vital to the operation of the DoDEA school system. Thus, maintenance of competitive teacher salaries is a top priority of DoDEA. Administrators believe that DoDEA still has the ability to attract and retain effective teachers, though the employment pool is more limited today than in the past. Salaries are viewed as a means of promoting this practice. The salary schedules of comparable (e.g., by size, demographics) school districts in the United States are reviewed regularly by DoDEA to establish a competitive salary schedule. A goal of the organization is to keep pace with the salaries offered by these comparable school districts.

The teacher salaries for both DoDDS and DDESS are displayed below in table 10, along with teacher salaries for a district of similar size, Charlotte-Mecklenburg in North Carolina. Two DoDEA school districts are located in North Carolina and they compete with Charlotte-Mecklenburg for the top teachers.

III. Curriculum and Instruction

"We spend a massive amount of time on our curriculum. Now of course people said, isn't that teaching to the test? No. We are testing what we are teaching." (Principal, DoDEA, 2001)

Well-qualified teachers, high expectations, and academic focus characterize the DoDEA schools. At a time when many school districts have large numbers of vacancies among the teacher ranks and uncredentialed staff, DoDEA has a fully staffed teaching force. The teachers in the DoDEA system have many years of experience and high levels of education, receive extensive ongoing training, and exhibit a strong commitment to teaching.

Table 10. Lowest and highest salaries on the 2000–01 teacher salary schedules for DDESS,[1] DoDDS,[2] and the Charlotte-Mecklenburg, NC, school district

System	Starting salary—teachers with a bachelor's and no years of experience (in dollars)	Highest salary—teachers with a doctorate and the most years of experience (in dollars)
Overseas–DoDSS teacher salary	30,700[3]	63,550[3]
Domestic–DDESS teacher salary	29,276	71,026
Charlotte-Mecklenburg, NC, teacher salary	28,068	60,104
Charlotte-Mecklenburg, NC, salary for national board teachers	—	67,013

—Not available.

[1]DDESS is the Department of Defense Domestic Dependents Elementary and Secondary Schools located in the United States.

[2]DoDDS is the Department of Defense Dependents Schools located overseas.

[3]Salary does not include housing allowance.

SOURCE: Department of Defense Education Activity (DoDEA) web site and Charlotte-Mecklenburg, NC, school district, web site.

Teachers and students share high expectations. The focus on academics is evident in the disciplinary procedures, scheduling, heterogeneous groupings, student supports, assessment, and innovative practices.

Teacher Quality

"We know what we are doing. We are good and we are dedicated." (Teacher, DoDEA, 2001)

Common indicators of teacher quality point to a strong teaching force in DoDEA schools. These teachers tend to have many years of teaching experience, high levels of education, and full qualifications to teach their subjects. In addition to these attributes, DoDEA teachers participate in integrated and extensive professional development, and exhibit a strong commitment to and enthusiasm for teaching.

Teaching Experience and Degrees Attained

Research has linked teacher qualifications and ability to student achievement. Robert Mendro tracked student performance in math and reading from grade 1 to 12 in the Dallas school system (Archer 1998). He found a 41 percent drop in average standardized test scores for students who had ineffective teachers for 3 years. A Harvard study indicated that spending more on highly qualified teachers produced greater gains in student performance than spending on any other item (Ferguson 1991). Another study found that the percentage of teachers with master's degrees accounted for 5 percent of the variation of student achievement scores (Berliner 1993). A significant problem in urban districts, where there are high concentrations of minority students, is that many newly hired teachers have no teaching license or emergency credential (Olson and Gerald 1998).

In DoDEA schools, a licensed teacher fills nearly every position and many teachers have extensive work experience and hold graduate degrees. As indicated below (see table 11), 73 percent of teachers in DoDEA schools have over 10 years of experience while only 10 percent of teachers have fewer than 3 years of experience. It is important to note that 64 percent of DoDEA teachers hold master's degree and 2.5 percent have doctorates.

Professional Development

"We probably have the best staff development program I have ever seen or read about. I truly believe that the success we have with kids is because of the training we give teachers. We have to train, train, train. . . . You have to have a teacher who wants it. And we do." (Principal, DoDEA, 2001)

"It is almost like an extended family when you come here. The teachers are very friendly, willing to cooperate with each other, willing to share information." (Teacher, DoDEA, 2001)

Education literature contends that professional development can be more effective by closely linking training to school initiatives to improve teaching strategies, offering intellectual, social, and emotional engagement with ideas and colleagues, and providing time and follow-up support for teachers to integrate new strategies into practice (Corcoran 1995). In addition, a RAND study concluded that professional learning is critically influenced by organizational factors at the school site and district, such as active involvement of the administration (McLaughlin and Marsh 1990). Furthermore, the study found that teacher efficacy, that is, a belief that the teacher can help even the most difficult student, was positively related to the number of goals achieved, amount of instructional change, and improved student performance. It is not surprising that DoDEA teachers believe they receive effective training.

Table 11. Percentage of DoDEA* teachers, by years of experience and highest level of education (based on the 1999–2000 DoDEA Profiles)

	Years of teacher experience				Teacher education		
	0 to 2	3 to 9	10 to 20	More than 20	BA or BS	MA or MS	Doctorate
Percent of DoDEA teachers	10	17	31	42	34	64	2.5

*DoDEA is Department of Defense Education Activity, the umbrella agency under which DDESS and DoDSS were united in 1994.

SOURCE: Department of Defense Education Activity (DoDEA), Annual Accountability Profile, 1999–2000.

Professional development is strongly supported in DoDEA schools. At DoDEA schools throughout the world there are opportunities to take university continuing education courses. In addition, every district that we visited had an array of professional training options available to teachers.

All districts in the study reported extensive staff training linked to school goals that occurs over extended periods of time. Staff development primarily reflects school goals. Teachers attend training workshops in various cities, but much staff development occurs at the school site. When the school, district, or DoDEA places a priority on a certain area, well-organized training activities that address that area are routinely made available to staff. In many cases, the training takes place over many weeks or months, so teachers can practice strategies in the classrooms. Curriculum specialists, principals, and fellow teachers provide coaching for new skills. Sharing ideas among teacher teams and grade levels is a regular activity in which teachers receive helpful ideas. Teachers uniformly praised the top quality of relevant training opportunities at DoDEA schools.

All districts in the study reported extensive staff training linked to school goals that occurs over extended periods of time.

DoDEA encourages its teachers to earn continuing education units. DoDEA teachers based in the United States and overseas reported that their school was linked to at least one university where they could continue to gain college credit while they maintained their full-time position. Some overseas teachers found access to college classes easier overseas than in the U.S. (civilian) school districts. U.S.-based teachers must maintain their state teaching license, while overseas teachers must comply with DoDDS continuing education requirements. However, training for DoDEA teachers is not limited to university offerings.

High Expectations

"I think that the school has to accept responsibility to make the difference for kids, not expect the kids to conform to make the difference for us. That is my belief. It is our job to teach the children in the way that will fit the kids best. And no excuses." (Superintendent, DoDEA, 2001)

High expectations are the norm in DoDEA schools. These high expectations are manifested in DoDEA's use of elevated academic standards, DoDEA teachers' sense of personal accountability, and their proactive approach to educating a highly transient student population.

Students in DoDEA schools confirm that teachers hold high expectations for them. As part of the school climate survey administered to students who took the 1998 NAEP reading test, respondents were asked to rate teacher expectations for student achievement (response scale included: very positive/somewhat positive/somewhat negative/very negative). In DDESS, 81 percent of the students reported that teachers' expectations of students are "very positive," compared to 58 percent in the national public school sample (see table 12). When disaggregated by race, the results are even more remarkable and relate significantly, we believe, to the linkage between high minority achievement and teacher expectations in DoDEA schools. In the DDESS system, 85 percent of Black students and 93 percent of Hispanic students reported that teachers' expectations for student performance are "very positive," compared to 52 percent and 53 percent, respectively, in the national sample.

IV. Policy Recommendations

Some observers contend that the high achievement in DoDEA schools, particularly for minority students, is a function of the middle class family and community characteristics of such students. We believe that such a view is overly simplified. Approximately 80 percent of all DoDEA students have a DoDEA parent/military sponsor who is enlisted. Most enlisted personnel have a high school diploma *only* and have income levels at or near the poverty line. Many enlisted personnel and their families do not live in comfortable housing. We argue that DoDEA schools simultaneously "do the right things," and "do things right." This statement applies both to what happens in schools and to what happens in a DoDEA out-of-school environment that reinforces rather than dilutes academic learning.

Table 12. Percentage of students who rated teacher expectations of student achievement "very positive" on the 1998 NAEP reading test
(In percent)

Race/ethnicity	Students in DDESS[1]	Students in nation[2]
All	81	58
White	70	60
Black	85	52
Hispanic	93	53

[1]DDESS is the Department of Defense Domestic Dependents Elementary and Secondary Schools located in the United States.

[2]The national results are based on the national assessment sample, which includes the DoDEA schools.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 Reading Assessment.

Small Schools. A larger proportion of middle schools and high schools in the DoDEA system have small enrollments compared to most other state systems. This fact stands in stark contrast to many urban school districts in the United States—the environments in which most minority students attend school (NCES 1998). In the DoDEA system, small school size contributes to teachers' and administrators' greater familiarity and personal knowledge of students, their instructional needs and strengths, and their unique family situations.

Policy recommendation: Research evidence and successful practice continually reinforce the utility of small schools, particularly in constructing an effective education for low income, minority students. A small school is defined as an elementary school with fewer than 350 students, a middle school with fewer than 600, and a high school with an enrollment of 900 or fewer (Lee and Smith 1997; Wasley et al. 2000). Creating smaller "learning communities" (Carnegie Council on Adolescent Development 1989) or schools-within-schools (Wasley et al. 2000) may very well facilitate the attainment of the organizational and social conditions evidenced in DoDEA schools, and could lead to enduring educational benefits for minority students in civilian schools.

Centralized direction-setting balanced with local decisionmaking. DoDEA's management strategy merges effective leadership at the topmost levels (e.g., establishing systemwide curriculum standards) with school- and district-level discretion in determining day-to-day operations such as instructional practices and personnel decisions.

Policy recommendation: Our findings suggest that state and local policymakers should utilize a management structure that functions as a "headquarters" for creating a blueprint for expected student learning and academic performance. DoDEA centrally establishes clear directions, goals, and targets without dictating methods for achieving results. This mix of top-down and bottom-up decisionmaking creates local capacity and professional confidence. It also serves as a basis for clear accountability. Principals and teachers know what they are expected to accomplish and are held responsible for accomplishing those goals. A similar civilian state-level priority setting strategy can serve as a springboard to propel higher academic achievement in U.S. public schools.

Policy coherence, structural alignment, and efficient flow of data. DoDEA schools reflect a strong and consistent alignment of curricular goals, instructional strategies, teacher supports, and performance assessment results. This is particularly evident in the area of writing, a subject area identified by DoDEA as a curricular priority and educational concern over 20 years ago.

Policy recommendation: DoDEA assessment systems are embedded within a coherent policy structure that links instructional goals with accountability systems supported by professional training and development programs. State and local policymakers can begin by adopting a performance-oriented information exchange that is systematic, clear, and comprehensive. States should provide every school and each district with detailed student performance assessment results. Using DoDEA as a model, each school should engage in a school improvement process to analyze student improvement needs and select student improvement goals. In DoDEA, student outcomes are specifically tied to downstream performance improvement goals. Staff training and curricular intervention are coordinated with a school's individual improvement plan. The ability and disposition to notice and act on instructional problems, and to deploy resources to solve problems, are critical elements of school improvement (Cohen and Ball 1999).

Sufficient financial resources. DoDEA provides a high level of support in terms of district and school staffing, instructional materials, facilities, and technology. The level of support for teachers is generous and well recognized throughout the system.

Policy recommendation: Money can matter, particularly when financial support is linked to specific, coordinated, and instructionally relevant strategic goals. State and local public education officials must acknowledge the crucial importance of sufficient resources. These resources enhance local capacity and strengthen the local districts' and individual schools' ability to implement school improvement goals. Sufficient resources enable districts to offer competitive salaries that attract and retain high-quality teachers.

DoDEA provides a high level of support in terms of district and school staffing, instructional materials, facilities, and technology.

Staff development. DoDEA professional development is linked to an individual school's pattern of student performance. It is tailored teacher by teacher, carefully structured to address a teacher's identified deficiencies, and sustained over time.

Policy recommendation: Professional development activities should be job-embedded; consistent with an individual school's improvement goals; based upon student needs and teacher interests; and modeled, repeated, and practiced over a long period of time. Professional training should include regular monitoring by peers or supervisors, sustained support, and regular feedback.

Academic focus and high expectations for all. DoDEA schools emphasize individual student achievement. High expectations are the norm in DoDEA schools. These high expectations are manifested in the use of elevated standards, teachers' sense of personal accountability, and a proactive approach to educating a highly transient student population. DoDEA schools do not generally group students by academic ability (i.e., tracking). Educational programs are provided that target lower achieving students for in-school tutoring and homework assistance after school.

Policy recommendation: Miles and Darling-Hammond (1997) found that high performing schools reflect a set of common strategies used to improve academic success. States should adopt these strategies, including: (1) a common planning time at each school to cooperatively develop curriculum; (2) a reduced number of specialized programs replaced by an integrated plan to serve students in regular classrooms (e.g., heterogeneous grouping); (3) targeted student groupings designed to meet individual needs and enable personal relationships; (4) modified school schedules to permit more varied and longer blocks of instructional time; and (5) creatively redesigned roles and work hours for staff to help meet goals. High academic rigor, supported by appropriate professional development, restores a system's focus on high academic performance.

Continuity of care for children. DoDEA schools are linked to an array of nationally recognized preschool programs and after-school youth service centers. This "continuity of care" commitment is evidenced by the high level of investment in these top-ranked programs in terms of staffing, training, and facilities. The DoDEA programs are widely recognized as a national model among child care providers in the United States in terms of staff training, educational programming, and facilities. The programs meet all standards established by the National Association for the Education of Young Children (NAEYC), the National Association of Family Child Care (NAFCC), and the National School-Age Care Association (NSACA).

Policy recommendation: State and local policymakers should utilize the DoDEA pre-school and after-school programs (e.g., youth service centers) as model programs that reflect the highest quality standards in the world. Many of these early and "out-of-school" educational activities contribute to enhanced student learning, self-esteem, and achievement.

"Corporate" commitment to public education. DoDEA schools reflect an elevated "corporate commitment" from the U.S. military that is both material and symbolic. This commitment includes an expectation of parent involvement in school- and home-based activities (e.g., soldiers are instructed that their "place of duty" is at their child's school on parent-teacher conference day, and are relieved of work responsibilities to volunteer at school each month). This commitment to promoting a parental role in education far surpasses the level of investment or involvement embraced by mentoring/tutoring models found in most business-education partnerships.

Policy recommendation: States and communities can gain similar levels of corporate commitment for public school students by making more visible the facets of the workplace that limit the ability of employee-parents (particularly the ability of hourly workers) to participate in school-based activities. Schools tend to structure school-based activities for traditional, stay-at-home mothers. At the same time, a large number of households have parents who are employed in full-time occupations that provide little flexibility and opportunity for parents to leave work during school hours. As schools begin to rethink the purpose and organization of their parent involvement activities, employers should re-evaluate workplace policies that hinder the kind of parental commitment to educational excellence that organized business groups are demanding in the current debate on the quality of our nation's schools.

References

Alexander, K., and Entwisle, D. (1996). Schools and Children at Risk. In A. Booth and J. Dunn (Eds.), *Family-School Links* (pp. 67–88). Mahwah, NJ: Lawrence Erlbaum Associates.

Annual Education Assessment, 20 U.S.C. § 924 (2000). Available: *http://www.access.gpo.gov/uscode/uscmain.html* (retrieved January 6, 2003).

Archer, J. (1998, April) Rotten Apples. *Teacher Magazine*. Available: *http://www.teachermagazine.org* (retrieved January 19, 2003).

Berliner, D. (1993). Mythology and the American System of Education. *Phi Delta Kappan, 74*(8): 634–640.

Bryk, A., and Driscoll, M. (1988). *The High School as Community: Contextual Influences for Students and Teachers*. Madison, WI: National Center on Effective Secondary Schools.

Bryk, A., Lee, V., and Holland, P. (1993*). Catholic Schools and the Common Good*. Cambridge, MA: Harvard University Press.

Carnegie Council on Adolescent Development (1989). *Turning Points: Preparing American Youth for the 21st Century.* Waldorf, MD: Carnegie Corporation of New York.

Cohen, D., and Ball, D. (1999). *Instruction, Capacity, and Improvement*. CPRE Research Report Series (RR-43). University of Pennsylvania: CPRE.

Cohen, D., and Spillane, J.P. (1992). Policy and Practice: The Relations Between Governance and Instruction. In G. Grant (Ed.), *Review of Research in Education, 18,* 4–49. Washington, DC: American Educational Research Association.

Coleman, J., and Hoffer, T. (1987). *Public and Private High Schools: The Impact of Communities*. New York: Basic Books.

Corcoran, T. (1995). *Helping Teachers Teach Well: Transforming Professional Development.* CPRE Research Report Series (RB-16). University of Pennsylvania: CPRE.

CORE (Consortium on Renewing Education). (1998). 20/20 Vision: A Strategy for Doubling America's Achievement by the Year 2020. Nashville, TN: Vanderbilt University, Peabody Center for Education Policy.

DoDEA Assessment Program. (2001, September). Available: *http://www.odedodea.edu/instruction/support/accountability/assessment.htm* (retrieved January 6, 2003).

DoDEA Community Strategic Plan. (2001, March). Available: *http://www.odedodea.edu/csp.mission.htm* (retrieved January 6, 2003).

Ferguson, R. (1991). Paying for Public Education: New Evidence on How and Why Money Matters. *Harvard Journal on Legislation, 28*(2): 465–498.

Ferguson, R. (1998). Teachers' Perceptions and Expectations and the Black-White Test Score Gap. In C. Jencks and M. Phillips (Eds.), *The Black-White Test Score Gap* (pp. 217–317). Washington, DC: Brookings Institution Press.

Fuhrman, S. (1999). *The New Accountability*. CPRE Research Report Series (RB-27). University of Pennsylvania: CPRE.

Jencks, C., and Phillips, M. (1998). The Black-White Test Score Gap: An Introduction. In C. Jencks and M. Phillips (Eds.), *The Black-White Test Score Gap* (pp. 1–51). Washington, DC: Brookings Institution Press.

Lee, V., and Smith, J. (1997). High School Size: Which Works Best and for Whom? *Educational Evaluation and Policy Analysis, 19*(3): 205–228.

McLaughlin, M., and Marsh, D. (1990). Staff Development and School Change. In A. Lieberman (Ed.), *Schools as Collaborative Cultures: Creating the Future Now* (pp. 213–232). New York: Falmer Press.

Miles, K., and Darling-Hammond, K. (1997). *Rethinking the Allocation of Teaching Resources: Some Lessons from High Performing Schools*. CPRE Research Reports Series (RR-38). University of Pennsylvania: CPRE.

Military Family Resource Center. (2001). *Profile of the Military Community: 2000 Demographics*. Arlington, VA: Author.

National Center for Education Statistics (1998). The NCES Common Core of Data (CCD), "Public Elementary/Secondary School Universe Survey, 1996–1997." U.S. Department of Education. Washington, DC: Author.

Natriello, G., McDill, E., and Pallas, A. (1990). *Schooling Disadvantaged Children: Racing Against Catastrophe.* New York: Teachers College Press.

Olson, L., and Jerald, C.D. (1998, January 8). The Teaching Challenge. *Education Week*: pp. 16–17.

Pellegrino, J.W., Jones, L.R., and Mitchell, K.J. (Eds.) (1999). *Grading The Nation's Report Card: Evaluating NAEP and Transforming the Assessment of Educational Progress.* Washington, DC: National Academy Press.

Systemwide Assessment Program, DoDEA Regulation 2000.6. (2001). Available: *http://www.odedodea.edu/foia/iod/pdf/2000_6.pdf* (retrieved January 6, 2003).

Wasley, P., Fine, M., Gladden, M., Holland, N., King, S., Mosak, E., and Powell, L. (2000). *Small Schools: Great Strides: A Study of New Small Schools in Chicago.* New York: The Bank Street College of Education.