Volume Title: The Economics of School Choice

Volume Author/Editor: Caroline M. Hoxby, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-35533-0

Volume URL: http://www.nber.org/books/hox03-1

Conference Date: February 22-24, 2001

Publication Date: January 2003

Title: School Vouchers. Results from Randomized Experiments

Author: Paul Peterson, William Howell, Patrick J. Wolf, David Campbell

URL: http://www.nber.org/chapters/c10087

4

School Vouchers
Results from Randomized Experiments

Paul E. Peterson, William G. Howell, Patrick J. Wolf, and David E. Campbell

In the past decade much has been learned about the way in which school vouchers affect low-income families and their children. Ten years ago, the empirical information available about this widely debated question came primarily from a flawed public school choice intervention attempted in Alum Rock, California during the 1960s (Bridge and Blackman 1978; El-

Paul E. Peterson is the Henry Lee Shattuck Professor of Government and director of the Program on Education Policy and Governance at Harvard University. William G. Howell is assistant professor of government at Harvard University. Patrick J. Wolf is assistant professor of public policy at Georgetown University. David E. Campbell is assistant professor of political science at Notre Dame University.

The authors wish to thank the principals, teachers, and staff at the private schools in Dayton, Washington, and New York City who assisted in the administration of tests and questionnaires. We also wish to thank the SCSF, PACE, and WSF for cooperating fully with these evaluations. Kristin Kearns Jordan, Tom Carroll, and other members of the SCSF staff assisted with data collection in New York City. John Blakeslee, Leslie Curry, Douglas Dewey, Laura Elliot, Heather Hamilton, Tracey Johnson, John McCardell, and Patrick Purtill of the Washington Scholarship Fund provided similar cooperation. T. J. Wallace and Mary Lynn Naughton, staff members of Parents Advancing Choice in Education, provided valuable assistance with the Dayton evaluation. Chester E. Finn, Bruno Manno, Gregg Vanourek, and Marci Kanstoroom of the Fordham Foundation, Edward P. St. John of Indiana University, and Thomas Lasley of the University of Dayton provided valuable suggestions throughout various stages of the research design and data collection. We wish to thank especially David Myers of Mathematical Policy Research, a principal investigator of the evaluation of the New York School Choice Scholarship Program; his work on the New York evaluation has influenced in many important ways the design of the Washington and Dayton evaluations. We thank William McCready, Robin Bebel, Kirk Miller, and other members of the staff of the Public Opinion Laboratory at Northern Illinois University for their assistance with data collection, data processing, conduct of the lottery, and preparation of baseline and year-one follow-up data. We are particularly grateful to Tina Elacqua, Matthew Charles, and Brian Harrigan for their key roles in coordinating data collection efforts.

We received helpful advice from Paul Hill, Christopher Jencks, Derek Neal, Donald Rock, and Donald Rubin. Daniel Mayer and Julia Chou were instrumental in preparing the New York City survey and test score data and executing many of the analyses reported in the paper.

more 1990). In the early and mid-1990s, however, new voucher programs sprouted across the country in such cities as Milwaukee, Cleveland, Indianapolis, and San Antonio. Initially, the evaluations of these innovations were limited by the quality of the data or the research procedures employed. Often, planning for the evaluation began after the experiment was under way, which made it impossible to gather baseline data or to ensure the formation of an appropriate control group. As a result, the quality of the data collected was not as high as researchers normally would prefer.[1]

Despite their limitations, these early evaluations provided program operators and evaluation teams with opportunities to learn the problems and pitfalls accompanying the study of school vouchers. Subsequent evaluations of voucher programs in New York, Washington, D.C., and Dayton, Ohio have been designed in such a way as to allow for the collection of higher-quality information about student test score outcomes and parental assessments of public and private schools. Because vouchers in these cities were awarded by lot, program evaluations could be designed as randomized field trials. Prior to conducting the lotteries, the evaluation team collected baseline data on student test scores and family background characteristics. One, two, and three years later, the evaluation team again tested the students and asked parents about their children's school experiences.[2] In the absence of response biases that are conditional on treatment status, any statistically significant differences between students offered a voucher and those not offered a voucher may be attributed to the intervention, because average student initial abilities and family backgrounds are similar between

Additional research assistance was provided by Rachel Deyette, Jennifer Hill, and Martin West; Antonio Wendland, Tom Polseno, Shelley Weiner, Lilia Halpern, and Micki Morris provided staff assistance.

These evaluations have been supported by grants from the following foundations: Achelis Foundation, Bodman Foundation, Lynde and Harry Bradley Foundation, William Donner Foundation, Thomas B. Fordham Foundation, Milton and Rose D. Friedman Foundation, John M. Olin Foundation, David and Lucile Packard Foundation, Smith-Richardson Foundation, Spencer Foundation, and Walton Family Foundation. The methodology, analyses of data, reported findings, and interpretations of findings are the sole responsibility of the authors of this report and are not subject to the approval of SCSF, WSF, PACE, or of any foundation providing support for this research.

1. Disparate findings have emerged from these studies. For example, one analysis of the Milwaukee choice experiment found test score gains in reading and math, particularly after students had been enrolled for three or more years, whereas another study found gains only in math, and a third found gains in neither subject. See Greene, Peterson, and Du (1998); Rouse (1997); and Witte (1997). On the Cleveland program, see Greene, Howell, and Peterson (1998) and Metcalf et al. (1998). Greene, Peterson, and Du (1998) report results from analyses of experimental data; the other studies are based upon analyses of nonexperimental data.

2. Results from the Dayton evaluation after one year are reported in Howell and Peterson (2000). Second-year results for Dayton are described in West, Peterson, and Campbell (2001). First-year results for Washington are reported in Wolf, Howell, and Peterson (2000). Second-year results for Washington are reported in Wolf, Peterson, and West (2001). First-year results from the New York City evaluation are reported in Peterson et al. (1999). Second-year results from New York City are described in David Myers et al. (2000). All of the occasional papers mentioned in this note are available at [http://data.fas.harvard.edu/pepg/].

the two groups. Students and families who were evaluated entered private school in grades two through five in New York City and grades two through eight in Washington, D.C. and Dayton (and other parts of Montgomery County, Ohio).[3] This chapter reports programmatic impacts on student test scores, parents' satisfaction with their child's school, and parent reports of the characteristics of the schools the child attended.

4.1 The Three Voucher Programs

The design of the three voucher programs was similar in key respects, thereby allowing the evaluation team to combine results from the separate evaluations of these programs. All were privately funded; all were targeted at students from low-income families, most of whom lived within the central city; and all provided partial vouchers, which the family was expected to supplement from other resources. All students included in the evaluation had previously been attending public schools. The programs, however, did differ in size, timing, and certain administrative details. In this section we describe the main characteristics of the School Choice Scholarships Foundation program in New York City, the Washington Scholarship Fund program in Washington, D.C., and the Parents Advancing Choice in Education program in the Dayton metropolitan area.

4.1.1 The School Choice Scholarships Foundation Program in New York City

In February 1997, the School Choice Scholarships Foundation (SCSF) announced that it would provide 1,300 scholarships worth up to $1,400 annually for at least three years to children from low-income families then attending public schools. The scholarship could be applied toward the cost of attending a private school, either religious or secular. After announcing the program, SCSF received initial applications from over 20,000 students between February and late April 1997.

To be eligible for a scholarship, children had to be entering grades one through five, live in New York City, attend a public school at the time of application, and come from families with incomes low enough to qualify for the U.S. government's free or reduced school lunch program. To ascertain eligibility, students and an adult member of their family were asked to attend verification sessions during which family income and the child's public school attendance were documented.

Subsequent to the lottery, SCSF assisted families in identifying possible private schools their children might attend. By the end of the first year,

3. Baseline data from the D.C. and Dayton evaluations are reported in Peterson et al. (1998). Baseline data for New York City are reported in Peterson et al. (1997). Both of these reports are available at [http://data.fas.harvard.edu/pepg/].

about 82 percent of the students participating in the evaluation were using a scholarship; 79 percent of the participating students used the voucher for two full years, and 70 percent for three full years.[4]

4.1.2 The Parents Advancing Choice in Education Program in Dayton, Ohio

In the spring of 1998, Parents Advancing Choice in Education (PACE), a privately funded nonprofit corporation, offered low-income families within the Dayton metropolitan area an opportunity to win a scholarship to help defray the costs of attending the school of their choice. Eligible applicants participated in a lottery in which winners were offered a scholarship that could be used at participating private and public schools in Dayton and in other parts of Montgomery County, Ohio. Students entering kindergarten through twelfth grade qualified. For the 1998–99 school year, PACE offered scholarships to 515 students who were in public schools and 250 students who were already enrolled in private schools.

The program was announced in January 1998. Based on census data and administrative records, program operators estimated that approximately 32,000 students met the program's income and eligibility requirements. The PACE program accepted preliminary applications from over 3,000 students, of whom 1,500 attended sessions where administrators verified their eligibility for a scholarship, students took the Iowa Test of Basic Skills (ITBS), and parents completed questionnaires. These verification sessions were held in February, March, and April 1998. The lottery was then conducted on 29 April 1998.

During the first year of the program, the PACE scholarships covered 50 percent of tuition at a private school, up to a maximum award of $1,200. Support was guaranteed for eligible students for at least four years; in addition, the program expects to support students through the completion of high school, provided funds remain available. Scholarship amounts were augmented beginning in 1999 as a result of additional funds available to PACE and support for the program by the Children's Scholarship Fund, a nationwide school choice scholarship program.

Among the public school students offered a scholarship, 78 percent of the students participating in the evaluation attended a private school in the program's first year, and 60 percent were in private schools after two years.

4.1.3 The Washington Scholarship Fund Program in Washington, D.C.

The Washington Scholarship Fund (WSF), a privately funded school voucher program, was originally established in 1993. At that time, a limited number of scholarships, which could be used at a private school of the fam-

4. For a description of the kinds of private schools voucher students attended, see Howell et al. (2002).

ily's choice, were offered to students from low-income families. By the fall of 1997, WSF was serving approximately 460 children at 72 private schools. The WSF then received a large infusion of new funds from two philanthropists, and a major expansion of the program was announced in October 1997. Both general news announcements and paid advertising were used to publicize the enlarged school choice scholarship program. The WSF announced that, in the event that applications exceeded scholarship resources, winners would be chosen by lottery. The program expanded further in 1999 with support from the Children's Scholarship Fund.

To qualify, applicants had to reside in Washington, D.C. and be entering grades K-8 in the fall of 1998. The WSF awarded parents with incomes at or below the poverty line vouchers that equaled 60 percent of tuition or $1,700, whichever was less. Families with incomes above the poverty line received smaller scholarships. The maximum amount of tuition support for high school students was $2,200. The WSF has said that it will attempt to continue tuition support to the children in its program for at least three years and, if funds are available, until they complete high school. No family with income above 2.7 times the poverty line was eligible for support.

Over 7,500 telephone applications to the program were received between October 1997 and March 1998; in response to invitations sent by WSF, over 3,000 applicants attended verification and testing sessions. The lottery selecting scholarship winners was held in April 1998. The WSF awarded over 1,000 new scholarships that year, with 811 going to students not previously in a private school.

Provided they gained admission, scholarship students could attend any private school in the Washington area. During the 1998–99 school year, students participating in the evaluation attended seventy-two different private schools. Of those students offered scholarships who participated in the evaluation, 68 percent attended a private school in the first year of the program. Take-up rates declined to 47 percent in the second year and to just 29 percent at the end of the third year.

4.2 Evaluation Procedures

The evaluation procedures used in all three evaluations conform to those used in randomized field trials. The evaluation team collected baseline data prior to the lottery, administered the lottery, and then collected follow-up information one and two years later. This section details the steps taken to collect the relevant information.

4.2.1 Baseline Data Collection

During the eligibility verification sessions attended by voucher applicants, students in first grade and higher took the Iowa Test of Basic Skills (ITBS) in reading and mathematics. The sessions took place during the months of February, March, and April, immediately prior to the voucher lottery, and generally lasted about two hours. The sessions were held in private-school classrooms, where schoolteachers and administrators served as proctors under the overall supervision of the evaluation team and program sponsors. The producer of the ITBS graded the tests.[5] Students in grades four through eight also completed a short questionnaire about their school experiences.

While children were being tested, adults accompanying them filled out surveys that asked about their satisfaction with their children's schools, their involvement in their children's education, and their demographic characteristics. Parents completed these questionnaires in rooms separate from those used for testing. Administrators explained that individual responses to the questionnaire would be held in strict confidence. Respondents had considerable time to complete their surveys, and administrators were available to answer questions about the meaning of particular items. Extensive information from these surveys has been reported elsewhere.[6]

Over 5,000 public-school students participated in baseline testing in New York City. After vouchers were awarded, 960 families were selected at random from those who did not win the lottery to comprise a control group of approximately 960 families.[7]

In Dayton, 1,440 students were tested at baseline, and 1,232 parent questionnaires were completed. Of the 1,440 students, 803 were not at the time attending a private school; of the 1,232 parent questionnaires, 690 were completed by parents of students who were not attending a private school. Follow-up testing information is reported only for families whose children attended public schools at the time of application.

In Washington, D.C., 2,023 students were tested at baseline; 1,928 parent surveys asking questions about each child were completed; 938 student surveys were completed. Of the 2,023 students tested, 1,582 were not attending a private school at the time of application for a scholarship; of the 1,928 parent questionnaires, 1,446 were completed by parents whose children were not then attending a private school. Follow-up testing and survey information was obtained only from families with children then in public schools.

4.2.2 The Lottery

The evaluation team conducted the lotteries in May 1997 in New York City and April 1998 in Dayton and D.C. Program operators notified lottery

5. The assessment used in this study is Form M of the Iowa Tests of Basic Skills, Copyright (c) 1996 by The University of Iowa, published by The Riverside Publishing Company, 425 Spring Lake Drive, Itasca, Illinois 60143-2079. All rights reserved.

6. See Howell et al. (2002). Prior reports include Peterson et al. (1999); Myers et al. (2000); Howell and Peterson (2000); West, Peterson, and Campbell (2001); Wolf, Howell, and Peterson (2000); Wolf, Peterson, and West (2001). All reports available at [http://data.fas.harvard.edu/pepg/].

7. Exact procedures for the formation of the control group are described in Hill, Rubin, and Thomas (1998).

winners shortly thereafter. If a family was selected, all children in that family entering eligible grades were offered a scholarship. Separate lotteries were held in Dayton and D.C. for students then in public and private schools, ensuring random assignment to test and control groups of those families participating in the evaluation.

In New York City, Mathematica Policy Research (MPR) administered the lottery; SCSF announced the winners. The SCSF decided in advance to allocate 85 percent of the scholarships to applicants from public schools whose average test scores were less than the citywide median. Consequently, applicants from these schools, who represented about 70 percent of all applicants, were assigned a higher probability of winning a scholarship. In the information reported in the tables, results have been adjusted by weighting cases differentially so that they can be generalized to all eligible applicants who would have come to the verification sessions had they been invited, regardless of whether or not they attended a low-performing school.

Because vouchers were allocated by a lottery conducted by the evaluation team, those offered scholarships should not be expected to differ significantly from members of the control group (those who did not win a scholarship). For all three cities, baseline data confirm this expectation. For instance, in D.C., the baseline test scores of those entering grades two through eight who were offered a voucher averaged 29.6 national percentile points in reading and 23.3 in mathematics; those not offered the scholarship scored, on average, 30.6 national percentile points in reading and 23.1 points in math. As in D.C., the demographic characteristics of those offered vouchers in Dayton and New York did not differ significantly from the characteristics of those who were not offered a voucher.[8]

4.2.3 Collection of Follow-Up Information

The annual collection of follow-up information commenced in New York City in the spring of 1998 and in Dayton and D.C. in the spring of 1999. Data collection procedures were similar across cities.

In New York City, testing and questionnaire administration procedures replicated those followed at baseline. Adult members of the family completed surveys that asked a wide range of questions about the educational experiences of their oldest child within the age range eligible for a scholarship. Students completed the ITBS and short questionnaires. Both the voucher students and students in the control group were tested in locations other than the school they were then attending.

The SCSF conditioned the renewal of scholarships on participation in the evaluation. Also, non–scholarship winners selected to become members of the control group were compensated for their expenses and told that they could automatically reapply for a new lottery if they participated in these

8. For a more extended discussion on these matters, see the initial reports for each city cited in notes 2 and 3.

follow-up sessions. Detailed response rate information for the follow-up survey and testing sessions is reported in appendix A.

In Washington, D.C. and Dayton, the evaluation team began collecting follow-up information between late February and late April of 1999. As in New York, the procedures used to obtain follow-up data were essentially the same as those used to collect baseline data. Students again took the ITBS in mathematics and reading. Caretakers accompanying the children completed surveys that asked a wide range of questions about the educational experiences of the children. Students in grades four through eight also completed a questionnaire that asked about their experiences at school. Testing and questionnaire administration procedures were similar to those that had been followed at baseline.[9] The Dayton evaluation was concluded after two years; in D.C., however, a third-year follow-up collection of testing and survey information was conducted in 2001.

To obtain a high participation rate in the follow-up data collection effort, those who had declined the offer of a voucher and members of the control group were compensated for their expenses. They were also told in Washington, D.C. that if they participated in the follow-up sessions, they would be included in a new lottery. In Dayton, a second lottery was promised as a reward for participating in the first follow-up session. In the second year, however, Dayton families were only given financial rewards for participation.

Because test score results from the second and third years of the evaluation differ significantly between African American students and those from other ethnic backgrounds, the ethnic composition of the students participating in the evaluation is particularly salient. Forty-two percent of the students participating in the second year of the evaluation in New York City were African Americans. In Dayton and D.C., 74 percent and 95 percent were African American, respectively. Hispanic students participating in the second year of the evaluation constituted 51 percent of the total in New York City, 2 percent in Dayton, and 4 percent in Washington, D.C. Finally, 5 percent of the students participating in the evaluation in New York City were white. In Dayton and D.C., 24 percent and 1 percent were white, respectively. The remaining students came from a variety of other ethnic backgrounds.

4.3 Data Analysis and Reporting Procedures

The evaluation takes advantage of the fact that a lottery was used to award scholarships. As a result, it is possible to compare two groups of stu-

9. Difficulties were encountered in the administration of the first-year follow-up test at the initial pilot session in Washington, D.C. Test booklets were not available at the testing site for scholarship students in grades three through eight. Copies of the test arrived eventually, but the amount of time available for testing may have been foreshortened. Significant effects on reading scores are not apparent, but significant effects on math performance are evident, probably because the math test was the last to be administered. Statistical adjustments in the test score analysis take into account the special circumstances of the pilot session.

dents that were similar, on average, except that members of the control group were not offered a scholarship. Any statistically significant differences between the two groups may be attributed to the school experience, not the child's initial ability or family background, which were essentially the same at baseline. One possible threat to the validity of this causal inference would be differential response patterns to follow-up testing by members of the treatment and control groups based on conditions that developed after they were tested at baseline.[10] We discuss that possibility in appendix A (see also Howell et al. 2002; Howell and Peterson 2002).

This paper provides data that help answer two questions. The first is what the impact on educational outcomes was of an *offer* of a voucher to low-income families residing within a large central city. This is the intention-to-treat or ITT effect of the voucher. The ITT effect compares educational outcomes of those who were offered a voucher to the outcomes of those who were not offered a voucher. To compute program impacts on children's test scores, we estimated a statistical model that took into account students' treatment or control group status as well as baseline reading and math test scores. Baseline test scores were included to (a) adjust for minor baseline differences between the treatment and control groups on the achievement tests, and (b) to increase the precision of the estimated impacts.

Generalization from these results has the important disadvantage of assuming that usage rates of scholarships are fixed. Depending upon the size of the scholarship, the time the scholarship is offered, and the marketing of the program as a whole, however, usage rates might be highly variable. Consequently, we report ITT results for test scores in appendix B. In the text of this chapter we report answers to a second question: What was the impact on educational experiences, parental satisfaction, and test score performances of students from low-income families residing within a large central city one, two, and three years after switching from a public to a private school? This is the treatment-on-the-treated or TOT effect of the voucher. The answer to this question requires a comparison between those students who were offered vouchers and switched from a public to a private school with public-school students who would have switched to a private school had they been offered a voucher. To compute the program's impact on those who used a scholarship to attend a private school, we estimated two-stage least squares models. The instrument is the voucher lottery, which is highly correlated with attendance at a private school, but because it is randomly determined, is obviously uncorrelated with the error term in the second-stage equation. As a result, the model yields an unbiased estimate of the effects of switching to a private school.[11]

The paper reports the TOT impact on students school experiences, parental satisfaction, and test score performance of a switch from a public

10. We are indebted to Derek Neal for calling this interesting contingency to our attention.

11. This procedure is discussed in Angrist, Imbens, and Rubin (1996). The procedure, widely used by statisticians to correct for selection effects, was used to estimate the effects of actual class size reduction in Tennessee; see Krueger (1999).

to a private school for one, two, and three years. Second- and third-year results compare those in private schools for two or three years with comparable members of the control group that were not in private school for two and three years, respectively.

4.4 Test Score Findings

We compare the performance of public and private school students on the ITBS in reading and mathematics, as well as their combined performance in both subject areas. Scores range between 0 and 100 National Percentile Ranking (NPR) points, with the national median located at the 50th percentile. The results reported below represent the first student achievement information from randomized field trials on the effects of school vouchers. However, they do not so much break new ground as build upon a body of research that has explored the differences between schooling for low-income minorities in the public and private sectors.

4.4.1 Prior Research

Studies of attainment levels and test performance of students in public and private schools usually find that low-income and African American students attending private schools outperform their public school peers. According to a recent analysis of 12,000 students in the National Longitudinal Survey of Youth, for instance, even when adjustments are made for family background, students from all racial and ethnic groups are more likely to go to college if they attended a Catholic school; however, the effects are the greatest among urban minorities (Neal 1997). This study's findings are consistent with others' (Evans and Schwab 1993; Figlio and Stone 1999). After reviewing the literature on school effects on learning, University of Wisconsin Professor John Witte (1996) concludes that the empirical literature "indicate[s] a substantial private school advantage in terms of completing high school and enrolling in college, both very important events in predicting future income and well-being. Moreover, . . . the effects were most pronounced for students with achievement test scores in the bottom half of the distribution" (167).

Even the most careful of studies, however, can take into account only observed family background characteristics. They cannot be sure that they have controlled for an intangible factor—the willingness of parents to pay for their child's tuition, and all that this implies about the importance they place on education. As a result, it remains unclear whether the findings from these studies describe actual differences between public and private schools or simply differences in the kinds of students and families attending them.[12]

12. Major studies finding positive educational benefits from attending private schools include Coleman, Hoffer, and Kilgore (1982) and Chubb and Moe (1997). Critiques of these studies have been prepared by Goldberger and Cain (1982) and Wilms (1985).

The best solution to the self-selection problem is the random assignment of students to test and control groups. Until recently, evaluations of voucher programs have not utilized a random-assignment research design and therefore have not overcome the possible selection problems. Privately funded programs in Indianapolis, San Antonio, and Milwaukee admitted students on a first-come, first-served basis. In the state-funded program in Cleveland, although scholarship winners were initially selected by means of a lottery, eventually all applicants were offered a scholarship, thereby precluding the conduct of a randomized experiment. The public Milwaukee program did award vouchers by a lottery, but data collection was incomplete.[13]

As a consequence, the findings presented here on New York, D.C., and Dayton provide a unique opportunity to examine the effects of school vouchers on students from low-income families who live in central cities. In contrast to prior studies, random assignment was conducted by the evaluation team, follow-up test-score information was obtained from about one-half to four-fifths of the students who participated in the lottery, and baseline data provided information that allowed the analysts to adjust for non-response.

4.4.2 Impacts of Private-School Attendance on Test Scores

In interpreting the findings reported below, emphasis is placed on the estimated effects of attending a private school on combined test scores for all three cities, taken together. Because of minor fluctuations in data collection, average estimates from more than one city provide a better indication of programmatic effects than do the results from any one city. Also, when student performance is estimated on the basis of one-hour testing sessions, combined test score performance of students on the reading and math tests is a better indicator of student achievement than either test separately. Theoretically, the more test items used to evaluate performance, the more likely it is that one will estimate performance accurately. Empirically, performances on the two tests are highly correlated with one another (r equals about 0.7). In addition, results from the two tests, when combined together, were found to be more stable across time and from place to place, indicating that combining results from the two tests reduces what is probably idiosyncratic variations in observations of student performance.[14]

As can be seen in table 4.1, the impact of private school attendance on student test score performance differed for African Americans and members of other ethnic groups. One observes no significant differences between the test score performance of non–African American students switching from a public to a private school and the performance of their peers in the control group, after one, two, or three years. Nor were significant differ-

13. Results from these evaluations are reported in Peterson and Hassel (1998).
14. This procedure was also employed in Krueger (1999).

Table 4.1 **The Impact in Three Cities of Switching to a Private School on Test Score Performance**

Test Score Performance	Year One (Percentiles)	Year Two (Percentiles)	Year Three (Percentiles)
African Americans			
Overall	3.9*	6.3***	6.6**
	(2.0)	(2.5)	(2.8)
Math	6.1***	6.1**	4.2*
	(2.4)	(3.1)	(2.2)
Reading	2.1	5.9**	4.2
	(2.4)	(2.9)	(3.5)
All other ethnic groups			
Overall	–1.6	–1.4	–3.5
	(2.4)	(2.9)	(2.7)
Math	–2.5	–2.6	–2.7
	(3.1)	(3.9)	(3.3)
Reading	–0.7	–0.2	–4.2
	(2.5)	(3.0)	(2.9)

Notes: Figures represent the average impact of switching to a private school on test score performance scores in New York, D.C., and Dayton. Averages are based upon effects observed in the three cities weighted by the inverse of the variance of the point estimates. Standard errors reported in parentheses. For African Americans, the unweighted average effects after one year are 2.7 overall, 4.8 in math, and 0.6 in reading; after two years, the unweighted average effect sizes are 6.6 overall, 6.5 in math, and 6.8 in reading. All models control for baseline test scores and lottery indicators. Impacts expressed in national percentile rankings.

***Significant at the 1 percent level, two-tailed test.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

ences observed in these students' reading and math tests, considered separately.

The effects of switching to a private school on African American students differed markedly from the effects on students from other ethnic backgrounds. In the three cities, taken together, African American students who switched from public to private schools scored, after one year, 3.9 NPR points higher on the combined math and reading tests, and, after two and three years, 6.3 percentile and 6.6 points higher, respectively, than the African American students in the control group. Again, these are the average results for the three cities combined, weighting each city estimate in inverse proportion to its respective variance.

The findings for each city are reported in tables 4.2, 4.3, and 4.4. The largest differences after three years were observed in New York City. In this city, African American students attending private schools for three years scored 9.2 percentile points higher on the two tests combined than did students in the control group. In D.C., however, no significant differences were observed after three years, despite the fact that large two-year effects were

Table 4.2	**Impact in New York of Switching to a Private School on Test Score Performance**					
Test Score Performance	Year One (Percentiles)	*N*	Year Two (Percentiles)	*N*	Year Three (Percentiles)	*N*
African Americans						
Overall	5.4***	622	4.4**	497	9.2***	519
	(1.5)		(2.0)		(2.4)	
Math	6.9***	622	4.1*	497	11.8***	519
	(1.8)		(2.5)		(2.9)	
Reading	4.0**	622	4.5**	497	6.7**	519
	(1.8)		(2.3)		(2.9)	
All other ethnic groups						
Overall	–2.2	812	–1.5	699	–3.5	729
	(1.8)		(2.2)		(2.4)	
Math	–3.2	812	–3.2	699	–2.5	729
	(2.3)		(2.9)		(2.9)	
Reading	–1.2	812	0.2	699	–4.4*	729
	(1.9)		(2.3)		(2.6)	

Notes: Weighted two-stage least squares regressions performed; treatment status used as instrument. Standard errors reported in parentheses. All models control for baseline test scores and lottery indicators. Impacts expressed in terms of national percentile rankings.

***Significant at the 1 percent level, two-tailed test.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 4.3	**Impact in Dayton of Switching to a Private School on Test Score Performance**			
Test Score Performance	Year One (Percentiles)	*N*	Year Two (Percentiles)	*N*
African Americans				
Overall	3.3	296	6.5*	273
	(3.5)		(3.7)	
Math	0.4	296	5.3	273
	(4.0)		(4.3)	
Reading	6.1	296	7.6*	273
	(4.2)		(4.2)	
All other ethnic groups				
Overall	1.0	108	–0.2	96
	(6.4)		(9.0)	
Math	–0.8	108	0.0	96
	(7.5)		(10.7)	
Reading	2.8	108	–0.4	96
	(7.1)		(9.9)	

Notes: Weighted two-stage least squares regressions performed; treatment status used as instrument. All models control for baseline test scores. Impacts expressed in terms of national percentile rankings. Standard errors reported in parentheses.

*Significant at the 10 percent level.

Table 4.4 Impact in D.C. of Switching to a Private School on Test Score Performance

Test Score Performance	Year One (Percentiles)	N	Year Two (Percentiles)	N	Year Three (Percentiles)	N
African Americans						
Overall	–0.9	891	9.2***	668	–1.9	656
	(2.8)		(2.9)		(4.4)	
Math	7.3**	891	10.4***	668	0.9	656
	(3.3)		(3.4)		(1.9)	
Reading	–9.0**	891	8.0***	668	–4.6	656
	(3.7)		(3.4)		(5.4)	
All other ethnic groups						
Overall	7.4	39	–0.1	42	–1.8	31
	(8.7)		(9.8)		(13.3)	
Math	8.5	39	7.3	42	–9.5	31
	(10.7)		(13.4)		(15.4)	
Reading	6.3	39	–7.6	42	5.9	31
	(12.7)		(10.1)		(18.7)	

Notes: Weighted two-stage least squares regressions performed; treatment status used as instrument. Standard errors reported in parentheses. All models control for baseline test scores; in year one, models also control for initial testing session. Impacts expressed in terms of national percentile rankings.

***Significant at the 1 percent level, two-tailed test.

**Significant at the 5 percent level.

observed. In Dayton, the difference in combined test score performance was 6.5 percentile points after two years, the total duration of the evaluation.

The trend over time also varies from one city to the next. As can be seen in table 4.2, in New York City, substantial test score differences between African American students in private and public schools appear at the end of the first year (5.4 percentile points) and attenuate slightly in the second year (4.4 points) but increase to 9.2 percentile points in year three. In this city, test score gains appeared to grow over time.

In Dayton, there was a steady upward trend in the combined test score performance of African Americans between years one and two. Table 4.3 shows that African American students who switched from public to private schools performed 3.3 percentile points higher on the combined test in year one and 6.5 percentile points higher in year two. Once again, a model of accumulated gains could account for the findings.

The most uncertain results for African Americans come from Washington, D.C. As can be seen in table 4.4, no significant differences were observed in year one, a large impact was observed after two years, but no impact was observed at the end of year three. Three factors could account for such disparate findings. First, because only 29 percent of the students in the evaluation continued to use the voucher after three years (as compared to 70 percent in New York City), third-year estimations are quite imprecise. Second, the voucher experiment in D.C. was contaminated by the inauguration of a charter-school initiative that gave families more choices than

Table 4.5 **Size of the Effects of Switching to a Private School on African Americans' Overall Test Score Performances (standard deviations)**

Test Score Performance	Year One	Year Two	Year Three
Overall	0.18	0.28	0.30
Math	0.28	0.28	0.18
Reading	0.08	0.23	0.16

Note: Figures represent the unweighted average impact of switching to a private school on test scores in New York, D.C., and Dayton expressed in standard deviations.

those available in New York City; indeed, 17 percent of the treatment group and 24 percent of the control group in D.C. attended charter schools in the third year of the evaluation. Finally, the differences in the third-year results might be attributed to the more established private sector in New York City than in Washington, D.C. Catholic schools, the major provider of private education in the two cities, are better endowed and historically more rooted in the northern port city, whose Catholic, immigrant population dates back to the early nineteenth century.[15]

4.4.3 Interpreting the Magnitude of the Test Score Effects

Overall, the effects of attending a private school on student test scores are moderately large. As can be seen in table 4.5, black students who switched to private schools scored, after one year, 0.18 standard deviations higher than the students in the control group. After two and three years, the size of the effect grew to 0.28 and 0.30 standard deviations, respectively, more than a quarter of the difference in test score performances between blacks and whites nationwide (Jencks and Phillips 1999). Continuing evaluations of voucher programs may provide information on whether or not these gains can be consolidated and extended.

Another way of assessing the magnitude of these effects is to compare them to those observed in an evaluation of a class size reduction intervention conducted in Tennessee, the only other major education reform to be subjected to evaluation by means of a randomized field trial. The effects on African Americans of attendance at a private school shown here are comparable to the estimated effect of a seven-student reduction in class size. According to a recent reanalysis of data from Tennessee, the class size effect for African Americans after two years was, on average, between 7 and 8 percentile points (Krueger and Whitmore 2000).

It is also of interest to compare the size of the effects of the voucher intervention with the size of the effects reported in a RAND study entitled *Improving School Achievement,* released in August 2000 (Flanagan, Kawata, and Williamson 2000, 59). Identifying the most successful states, Texas and North Carolina, which have introduced rigorous accountability systems

15. For a fuller discussion, see Howell et al. (2002, chaps. 2, 6).

that involve statewide testing, the study finds what it says are "remarkable" one-year gains in math scores in these states of "as much as 0.06 to 0.07 standard deviation[s] per year"—or 0.18 to 0.21 over three years. The three-year effects of the school voucher intervention on black students observed here are somewhat larger.

4.4.4 Cost-Benefit Analysis

What are these test score gains for African Americans likely to mean in terms of future economic benefits? Richard Murname and his colleagues have calculated the effects of math achievement on future earnings (Murname et al. 2000). According to one estimate, a 0.30 standard deviation increase in average math achievement, if sustained, will yield a 5 percent gain in earnings seven to ten years after the student finishes high school. If an African American student in the control group was expected to earn about $30,000 a year in his late twenties, a comparable student who had switched from public to private school would be expected to earn an additional $1,500 per year.

This suggests that investments in vouchers might yield a moderate rate of return for African American families. Why, then, don't families make this investment on their own? If a control group family had simply absorbed the cost of the voucher (on average, about $1,200), even a rough calculation of the rate of return suggests that it would be an attractive investment, provided that families can borrow moneys at conventional rates. Credit constraints are a possible explanation for the decision not to utilize a private school on the part of control group families. Private lenders may be reluctant to make long-term loans at conventional lending rates to low-income borrowers, who may be high credit risks. If families can borrow the money only at rates charged to high-risk users of bank cards, then the rate of return on an investment in private schooling, although probably still positive, would be considerably less attractive—unless a family perceives nonpecuniary benefits of a private education.

Clearly, though, the lower the initial costs, the more attractive an investment in private schooling becomes. When a voucher reduces the amount that needs to be borrowed from around $2,400 a year (a rough estimate of the average cost of private school tuition, fees, books, and uniforms in these cities) to half that amount, families may decide that the benefits of an investment in private schooling now outweigh the costs. Perhaps this explains why a small voucher induced many low-income families to make the additional investment, even when members of a similarly situated control group (who did not receive the voucher offer) were less likely to do so.[16]

16. A careful analysis of this question would require a fuller examination of the probable economic benefits of test score gains, the cost of private schooling, and the interest rates faced by various classes of potential borrowers.

4.4.5 Additional Methodological Considerations

This section addresses two methodological considerations. The first involves the status of background control variables. In table 4.6 we report second- and third-year results for African Americans from statistical models that control not only for initial test scores (as do the analyses in the previous tables) but also for mother's education, mother's employment status, family size, and whether or not the family received welfare benefits. The estimated impacts on the test scores of African Americans of switching from a public to a private school in the three cities remain almost exactly the same: 6.4 percentile points in the second year and (though not shown) 6.7 percentile points in the third. Minor differences are observed when impacts within each individual city are estimated. For instance, when we estimate effects in New York City in the second year without controlling for family background characteristics, the impact is 4.3 NPR points; when family background controls are added, the impact is 4.5 NPR points. In Dayton, Ohio, when controls are introduced, the point estimate drops from 6.5 to 5.9 NPR points. In Washington, D.C., the estimated impact after two years remains 9.2 NPR points.

The second methodological consideration concerns the possibility that African Americans posted significant effects because they received a more uniform treatment. If the black students who used vouchers were dispro-

Table 4.6 Estimated Effects after Two Years of Switching from a Public to a Private School on African Americans' Combined Test Scores, With and Without Controls for Family Background Characteristics

	Private School Impact, Original Results	Private School Impact, Controlling for Family Background
Three-city average impact	6.3***	6.4**
	(2.5)	(2.5)
New York City	4.4**	4.5**
	(2.0)	(2.0)
Dayton, Ohio	6.5*	5.9
	(3.7)	(3.8)
Washington, D.C.	9.2***	9.2***
	(2.9)	(2.8)

Notes: P-values reported in parentheses. Weighted two-stage least squares regressions performed; treatment status used as instrument. All models control for baseline test scores, mother's education, employment status, whether or not the family receives welfare, and family size (missing case values for demographic variables estimated by imputation); New York model also includes lottery indicators. Impacts expressed in terms of national percentile rankings. Average three-city impact is based on effects observed in the three cities weighted by the inverse of the standard errors of the point estimates.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

portionately concentrated in a small number of good private schools, or their peers in the control group concentrated disproportionately in a few bad public schools, the error term in the estimation of private school effects would be smaller for African Americans than for other students. This would increase the probability that one would observe significant impacts on African American test scores, but not on those of other ethnic groups.

For two reasons, however, we doubt this explanation has much traction. First, the size of the standard errors is not all that differentiates the effects for African Americans and members of the other ethnic groups. For Latinos in New York and whites in Dayton, the point estimates consistently hover around zero, whereas for African Americans in both cities the point estimates are quite large. Second, when surveying the private school attendance patterns of students from different ethnic groups, one finds little evidence that treatment effects were more uniform for some groups than others. African Americans, for the most part, did not attend a relatively smaller number of public or private schools than did members of other ethnic groups.[17]

4.5 Parent Satisfaction

Most studies have found that families who use vouchers to attend an area private school are much more satisfied with their schooling than are families who remain in public schools.[18] The results presented in table 4.7 confirm these earlier findings. A significantly higher proportion of private school parents were "very satisfied" with the following aspects of their schools: school safety, teaching, parental involvement, class size, school

17. Since information about the distribution of students among schools is available from the first year of the Dayton evaluation, we were able to estimate the extent to which African Americans and non–African Americans were subject to uniform treatment simply by dividing the number of students in a category by the number of schools they attended. On the whole, we found fairly low uniformity of treatment and not much difference between racial groups. For both African American and non–African American students receiving treatment, the degree of concentration among schools was, on average, just three students. Among students in the control group, the degree of concentration was 3 students per school for African Americans and 1.3 students for non–African Americans. According to this estimation, then, some difference in the degree of concentration between African American and non–African American students is evident, but the difference is not large.

One may also estimate the degree of uniformity of treatment by examining the percentage of students in the three schools serving the largest number of students. When one estimates in this way, one again finds some difference between treatment and control groups. However, in this case it is the non–African Americans who appear the most concentrated. A total of 37 percent of the African American treatment students enrolled in just three schools, as compared to 15 percent of the African American members of the control group. For non–African Americans, these figures were 56 percent and 16 percent, respectively. According to this estimate, the nonblack members of the Dayton experiment experienced a more uniform dose of treatment than did the black students in the study.

18. A summary of findings from earlier studies is available in Peterson (1998, 18). Schneider et al. (1998) finds higher levels of parental satisfaction within New York City public schools when parents are given a choice of school.

Table 4.7 **Parent Satisfaction with School, Two Years After Beginning of Voucher Programs (% "very satisfied")**

Parent Satisfaction Category	Switched to Private School (1)	Public School Control Group (2)	Year Two Programmatic Impact (3)
What taught in school	44	15	29***
Ability to observe religious traditions	37	8	29***
School safety	44	16	27***
Teacher skills	43	17	26***
Teacher-parent relations	43	18	25***
Student respect for teachers	40	16	24***
Academic quality	38	15	23***
Teaching values	36	14	23***
Discipline	35	14	22***
Staff teamwork	34	13	21***
Class size	32	12	20***
Clarity of school goals	34	14	20***
Parental involvement	30	15	15**
Location	40	33	7

Notes: These figures represent the average results for New York City, Dayton, and D.C. Observations from each city are weighted by the inverse of their variance. Column (1) presents those who were offered a scholarship and subsequently used it to attend a private school. Column (2) presents those in the control group who would have used a scholarship had they been offered one. Column (3) presents estimated impact of participation in the program, using a two-stage least squares model.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

facility, student respect for teachers, teacher communication with parents with respect to their child's progress, extent to which child can observe religious traditions, parental support for the school, discipline, clarity of school goals, staff teamwork, academic quality, the sports program, and what is taught in school. Thirty-eight percent of private school parents were very satisfied with the academic quality of the school after two years, as contrasted with just 15 percent of the control group. Similarly, 44 percent of the private school parents expressed the highest satisfaction with "what's taught in school," compared with 15 percent of the control group.

To see whether satisfaction levels are the result of a Hawthorne effect, the propensity of individuals to welcome change for its own sake, an index of satisfaction was constructed from the items reported above. The positive impact on satisfaction levels in the three cities, as measured by this index, was 0.97 standard deviations in the first year, 0.89 in the second, and 0.85 in the third year.[19] In other words, although overall satisfaction levels attenu-

19. Procedures for constructing the index as well as additional information on satisfaction levels are reported in Howell et al. (2002, chap. 7).

ated slightly from the first to the second and third years of the evaluation, Hawthorne effects appear minimal.

4.6 Other Voucher Impacts

Although test score performance and parental satisfaction are the outcomes of greatest interest to most observers, parental surveys provided additional information about the impacts of voucher opportunities on selected characteristics of the schools attended by students. Significant differences were identified in the school facilities available to students, school size, class size, school climate, homework assignment practices, and school communication with families.

4.6.1 School Facilities

Public school expenditures eclipse private school expenditures. Nationwide, the average private school expenditures per pupil in 1993–94 were estimated at $3,116, considerably less than the $6,653 spent, on average, on public school pupils (Coulson 1999, 277). In part, this disparity is due to the wider array of services that public schools provide their students. Nonetheless, even when adjustments are made for the kinds of services rendered, public schools in New York City, Dayton, and D.C. spend roughly twice as much as private schools.

Per-pupil expenditures for both Catholic and public schools were available for schools in three boroughs of New York City.[20] In comparing expenditures, the amount spent by New York public schools for all items that did not clearly have a private school counterpart was deducted. Among other things, deductions were taken for all monies spent on transportation, special education, school lunch, other ancillary services, and the cost of financing the far-flung bureaucracy that runs the citywide, boroughwide, and districtwide operations of the New York City public schools.

All these deductions from public school expenditures amounted to no less than 40 percent of the cost of running the New York City public schools. However, even after all these and other deductions were taken, public schools were still spending over $5,000 per pupil each year, more than twice the $2,400 spent on similar services in New York City's Catholic schools.

In Washington, D.C., the median tuition at the private schools attended by the scholarship students included in the evaluation was $3,113 in the year 1998–99.[21] The average is substantially higher than the median because of the high tuition charged by a few independent schools, such as Sidwell

20. Estimates are based on information about Catholic schools in Manhattan, the Bronx, and Brooklyn from an unpublished memorandum submitted to the Program on Education Policy and Governance from the New York archdiocese in August 1999. Public school expenditure by school for the city of New York is available on the Board of Education website.

21. Private school tuition rates were estimated in part from information provided in Coerper and Mersereau (1998). For schools not listed in this volume, information was obtained in

Table 4.8 **Size and Quality of School Facilities, One Year After Beginning of Voucher Programs (%)**

Parental Reports	Switched to Private School (1)	Public School Control Group (2)	Programmatic Impact (3)
Average school size	278	450	–172***
Average class size	20	23	–3***
Percentage satisfied with school facilities	28	9	19***
Percentage with the following resources:			
Special programs for non-English speakers	43	71	–28***
Nurses' office	75	94	–19***
Special programs for learning disabled	67	81	–14***
Cafeteria	86	96	–10***
Child counselor	77	85	–8***
Library	92	96	–5**
Gym	88	88	0
Special programs for advanced learners	59	58	1
Arts program	82	81	1
Computer lab	86	84	2
Music program	88	84	4
After-school program	91	86	6**
Individual tutors	70	54	16***

Notes: See notes to table 4.7.

Friends, which charged the Clintons over $15,000 per year for their daughter's tuition. Assuming that the ratio of tuition to total educational expenditure in Washington, D.C. is the same as in the three boroughs in New York City discussed previously, private educational expenditures, on average, totaled roughly $4,000. Again, considering only those services and programs that both public and private schools cover, adjusted per-pupil expenditures in Washington public schools reached $8,185, as estimated from data for the 1995–96 school year.[22]

Much the same patterns emerge in Dayton. In 1998–99, students in the Dayton voucher program paid, on average, $2,600 in tuition, whereas the Dayton public school system spend an adjusted average of $5,528 per pupil.

Parental reports help explain these expenditure data. According to the parental surveys, private schools were less likely to have a library, a nurse's office, a cafeteria, child counselors, and special programs for non–English speakers and students with learning problems (see table 4.8). The greatest difference was for programs for non-English-speaking students. Forty-

telephone conversations with school staff. Some schools have a range of tuition charges, depending on the number of students from the family attending the school and other factors. The tuition used for this calculation is the maximum charged by the school. The tuition also includes all fees, except for the registration fee, which is ordinarily treated as partial payment toward tuition. Figures are weighted proportionate to the number of students in the evaluation attending a particular school. Public school expenditure includes the costs of transportation and special education, which may not be provided by private schools.

22. Data taken from the U.S. Department of Education (2000).

three percent of the private school parents reported such a program in their school, compared with 71 percent of the control group parents. Similarly, 75 percent of the private school parents reported their school had a nurse's office, as compared to 94 percent of public school parents. Public schools are also larger. In some instances, either no significant differences were detected or private school parents reported more services. The two groups of parents did not differ in their reports of the availability of a gym, a computer laboratory, art and music programs, and special programs for advanced learners. Private school parents, meanwhile, were more likely to say their school had individual tutors and an after-school program.

Despite the more limited financial resources of the private school, parents reported that their children attended classes with an average of twenty students, as compared to twenty-three in public schools (Peterson, Myers, and Howell 1998, table 5). However, the reduction in class size was only three students, considerably less than the amount generally thought to be necessary to achieve significant gains from class size reduction.[23] As estimated by parents, the effect of choosing the private sector was to reduce the average size of the school by 172 students or nearly 40 percent—from an average of 450 students to 278 students.

4.6.2 School Climate

In their study of public and private schools, John Chubb and Terry Moe (1990) found that the educational environment of private schools was more conducive to learning than that of public schools. They pointed out that public schools are governed by state laws, federal regulations, school board requirements, and union-contract rules that impose multiple and not always consistent obligations on teachers and principals. Because they must respond to numerous legal and contractual requirements, school administrators and teachers focus more on rule compliance than on educational mission, undermining the morale of educators whose original objective was to help children learn.

The problem, Chubb and Moe say, is particularly prevalent in big-city schools. Urban private schools operate with greater autonomy, focus more directly on their educational mission, and, as a result, achieve a higher degree of internal cohesion. To do otherwise would jeopardize their survival as fragile institutions dependent upon the annual recruitment of new students. As a consequence, principals and teachers in the private sector enjoy higher morale. Their interactions with one another and with their students are more positive, fostering a more effective learning environment.

Our findings confirm Chubb and Moe's. If parent reports are accurate, the

23. The reduction in class size in the Tennessee experiment was an average of approximately seven to eight students (Krueger 1999). For further discussion of this point, see Howell et al. (2002, 158–64).

Table 4.9 **Parents' Perceptions of School Climate, One Year After Beginning of Voucher Programs (%)**

Reported by Parents as Serious Problem	Switched to Private School (1)	Public School Control Group (2)	Programmatic Impact (3)
Fighting	32	63	–31***
Kids missing class	26	48	–22***
Tardiness	33	54	–21***
Kids destroying property	22	42	–20***
Cheating	26	39	–13***

Notes: See notes to table 4.7.

Table 4.10 **Homework, One Year After Beginning of Voucher Programs (%)**

Reported by Parents	Switched to Private School (1)	Public School Control Group (2)	Programmatic Impact (3)
Child has more than one hour of homework	72	56	16***
Difficulty of homework appropriate for child	90	72	18***

Notes: See notes to table 4.7.

scholarship programs in New York, D.C., and Dayton had a major impact on the daily life of students at school. As table 4.9 shows, public school parents were more likely to report that the following were serious problems at their school: students destroying property, tardiness, missing classes, fighting, cheating, and racial conflict. For example, 32 percent of the private school parents thought that fighting was a serious problem at their school versus 63 percent of the control group. Thirty-three percent of parents perceived tardiness as a problem, as compared to 54 percent of the control group. No more than 22 percent of private school parents, but 42 percent of the control group, said that destruction of property was a serious problem at their school.

4.6.3 Homework and Parental Communication

Thomas Hoffer, Andrew Greeley, and James Coleman (1985) have attributed the higher level of student performance in private schools to the amount of homework expected of students and to the frequency of communication between schools and parents. The reports by parents are consistent with their interpretation.[24] Table 4.10 shows that 72 percent of private school parents reported that their child had at least an hour of

24. For very similar first-year results, see Peterson, Myers, and Howell (1998), table 9.

Table 4.11 **School Communication with Parents, One Year After Beginning of Voucher Programs (%)**

Reported by Parents	Switched to Private School (1)	Public School Control Group (2)	Programmatic Impact (3)
Parents receive newsletter	88	68	20***
Parents participate in instrument	68	50	18***
Parents notified of disruptive behavior	91	77	14***
Parents receive notes from teacher	93	78	14***
Parents speak to classes about jobs	44	33	11**
Parents regularly informed about student grades	93	84	9**
Parent open houses held at school	95	90	5**
Regular parent-teacher conferences	95	90	5**

Notes: See notes to table 4.7.

homework a day, whereas only 56 percent of the control group parents reported a similar amount of homework. Private school parents were also more likely to say the homework was appropriate for their child. Seventy-two percent of the control group parents gave this response, as compared to 90 percent of private-school parents.

Compared with control-group parents, parents of students in private schools also said that they received more communication from their school about their child. The results presented in table 4.11 indicate that a higher percent of private school parents versus control group parents reported the following: being more informed about student grades halfway through the grading period; being notified when their child is sent to the office the first time for disruptive behavior; speaking to classes about their jobs; participating in instruction; receiving notes about their child from the teacher; receiving a newsletter about what is going on in school; and regular parent-teacher conferences.

4.7 Conclusions

Because random assignment to test and control groups assures that all significant effects may be attributed to the intervention, and not to the students' initial abilities or their family backgrounds, randomized field trials are the best available tool for detecting the effects of an educational intervention. Nonetheless, when one interprets the findings from the evaluation of any one program in a particular city, generalizations to a larger universe are problematic. Conditions specific to that place or minor fluctuations in testing conditions might skew results in one direction or another.

Still, when similar results emerge from evaluations of school voucher programs in three sites in different parts of the United States, they provide a stronger basis for drawing conclusions and generalizing to larger populations. Thus, the average impact across the three sites may provide a reasonable estimate of the likely initial impact of a school voucher initiative elsewhere.

In the three cases, taken together, we found effects of school vouchers only on the average test performance of African American students. African American students who switched from public to private schools in the three cities scored after two years, on average, approximately 6.3 percentile points higher on the ITBS than comparable African Americans who remained in public schools. After three years, private school attendance in two cities had an impact of 6.6 percentile points, an effect of 0.30 standard deviations.

At this point we do not know why the gains from switching to a private school are evident for African American students after two and three years, but not for students from other ethnic backgrounds. However, parents reported that private schools are smaller in size, maintain a better disciplinary climate, ask students to do more homework, maintain closer communication with families, and have somewhat smaller classes (about 3 fewer pupils). These school characteristics may be particularly helpful to students who are African American.

One must qualify any generalizations from the results of this pilot program to a large-scale voucher program that would involve all children in a large urban school system. Only a small fraction of low-income students in these three cities' schools were offered vouchers, and these voucher students constituted only a small proportion of the students attending private schools in these cities. A much larger program could conceivably have quite different program outcomes.

Still, slightly larger voucher programs initially directed at low-income families would attract those families with the greatest interest in exploring an educational alternative, exactly the group that applied for a voucher in these three cities. Thus, positive consequences of school choice reported herein may prove encouraging to those who seek to extend and expand school choices for low-income, inner-city families, and negative findings may indicate problems that need to be addressed. It is hoped that additional careful research will accompany larger programs established by private philanthropists and public authorities.

Appendix A

Response Rates

To promote high response rates, voucher program operators either required or strongly urged recipients to participate in testing sessions if they wished to have their voucher renewed for the next school year. In addition, evaluation teams offered financial incentives and new opportunities to win a voucher to encourage members of the control group and members of the treatment group who remained in public schools to return for follow-up testing.[25] Still, substantial numbers of students were not tested at the end of one, two, and three years.

Response rates were 100 percent at baseline, because families and students were not entered into the lottery unless they provided baseline information. Response rates after one year were 82 percent in New York, 56 percent in Dayton, and 63 percent in Washington, D.C. After two years, response rates in the three cities were 66 percent, 49 percent, and 50 percent, respectively. After three years, response rates were 67 percent in New York and 60 percent in Washington, D.C. Response rates were similar for treatment and control groups in all three cities.[26] The largest difference was in New York City in the second year, where the treatment group's response rate was 7 points higher than the control group rate.[27]

Comparisons of baseline test scores and background characteristics reveal only minor differences between respondents and nonrespondents in all three cities. Table 4A.1 presents, for example, baseline data on respondents and nonrespondents in the treatment and control groups after two years in the three cities; separate comparisons for African Americans are included in table 4A.2. Some differences in race, welfare, and religious orientation were detected, but they point in different directions in different cities and do not appear to systematically produce a more advantaged group of respondents in the treatment group or a particularly disadvantaged control group. In all three cities, inter-group differences in test scores, religious identification, residential mobility rates, church attendance, and family size were essentially nonexistent.

To adjust for the bias associated with nonresponse, in each year and city we generated weights for parents and students in the treatment and control

25. In New York City and Washington, D.C., families in the control group were entered in a new lottery if they attended follow-up testing in years one and two. In Dayton, control group families were entered in a new lottery after the first year of the program; in year two, they were offered higher compensation instead. Families that began the study as members of a control group were dropped from the evaluation if they subsequently won a follow-up lottery. Although this was necessary to preserve the random design of the evaluation, excluding such families had the effect of reducing the size of the control groups slightly. Although families that did complete the surveys may be systematically different from those that did not, dropping the randomly selected subset of survey respondents should only decrease the efficiency of the estimates, not bias the findings. In D.C. in year three, all control group families and all those that did not use the initial vouchers offered them were offered a voucher.

26. The one exception here concerns the year-two evaluation in New York City, in which the treatment group's response rate was 7 points higher than the control group's rate.

27. These response rates are similar to those in other randomized field trials that follow students over time. In his reanalysis of data from the Tennessee class size study, for example, Krueger (1999), while not providing annual attrition rates, reports that "only half the students who entered the project in kindergarten were present for all grades K-3" (506).

Table 4A.1 **Baseline Characteristics of Respondents and Nonrespondents in Treatment and Control Groups in Year Two**

	Treatment		Control	
	Attended Year Two	Didn't Attend Year Two	Attended Year Two	Didn't Attend Year Two
New York City				
% African American	42.4	48.3	41.4	47.2
% welfare recipients	53.2	64.5	59.4	62.3
% Catholic	54.7	46.4	53.7	43.2
% protestant	34.3	39.4	35.0	38.8
Average overall test scores	20.1	19.5	22.8	22.6
Average family size	2.6	2.6	2.4	2.9
Average residential mobility	3.7	3.6	3.7	3.7
Average church attendance	3.6	3.3	3.4	3.5
Average mother's education	2.4	2.4	2.4	2.5
Dayton				
% African American	74.0	65.2	71.9	69.3
% welfare recipients	16.7	13.8	16.2	16.7
% Catholic	5.8	14.0	13.4	18.1
% protestant	65.2	58.1	64.6	56.9
Average overall test scores	26.3	26.3	27.2	26.2
Average family size	3.0	3.0	3.0	3.1
Average residential mobility	3.4	3.3	3.3	3.6
Average church attendance	3.4	3.3	3.6	3.7
Average mother's education	5.6	5.4	5.3	5.6
D.C.				
% African American	90.4	92.1	90.9	92.1
% welfare recipients	38.0	34.1	32.1	30.3
% Catholic	15.5	12.6	16.0	13.8
% protestant	72.7	69.9	65.6	70.6
Average overall test scores	26.5	26.4	26.9	26.7
Average family size	3.1	3.1	3.3	3.0
Average residential mobility	3.4	3.5	3.5	3.4
Average church attendance	3.7	3.5	3.7	3.7
Average mother's education	5.4	5.0	5.3	5.2

Notes: Averages refer to the unweighted mean scores of responses on the parent surveys. Mother's education was scaled slightly differently in New York City than in Dayton and Washington, D.C., making intercity comparisons on that item inappropriate.

groups. Because those invited to participate in the follow-up studies had provided information about their family characteristics at baseline, it was possible to calculate the probability that each participant in the baseline survey would attend a follow-up session. To do so, we estimated simple logit regressions that used a set of variables assembled from baseline surveys to predict the likelihood that each student would attend a follow-up session. Covariates included mother's education, employment status, marital status, and religious affiliation; family size; whether the family received welfare

Table 4A.2 **Baseline Characteristics of African American Respondents and Nonrespondents in Treatment and Control Groups in Year Two**

	Treatment		Control	
	Attended Year Two	Didn't Attend Year Two	Attended Year Two	Didn't Attend Year Two
New York City				
% welfare recipients	55.3	63.5	65.8	65.8
% Catholic	17.8	18.0	19.5	8.6
% protestant	66.9	67.3	66.8	65.2
Average overall test scores	20.6	19.1	21.2	23.8
Average family size	2.6	2.8	2.5	3.1
Average residential mobility	3.8	3.7	3.7	3.7
Average church attendance	3.4	3.4	3.2	3.3
Average mother's education	2.5	2.5	2.5	2.6
Dayton				
% welfare recipients	15.9	15.0	17.9	20.7
% Catholic	7.6	12.4	8.7	4.7
% protestant	66.1	61.9	76.5	69.8
Average overall test scores	24.3	21.6	23.2	22.0
Average family size	2.7	2.6	2.8	2.9
Average residential mobility	3.4	3.2	3.3	3.6
Average church attendance	3.8	3.7	3.9	4.0
Average mother's education	6.1	5.8	5.5	5.4
D.C.				
% welfare recipients	38.8	31.6	34.7	27.9
% Catholic	13.0	15.1	13.7	16.0
% protestant	76.2	69.4	67.8	67.7
Average overall test scores	26.2	28.3	26.1	28.4
Average family size	3.1	3.0	3.3	2.9
Average residential mobility	3.5	3.5	3.5	3.5
Average church attendance	3.7	3.6	3.7	3.8
Average mother's education	5.4	5.2	5.3	5.4

Notes: See notes to table 4A.1.

benefits; whether the student was African American; the student's baseline math score; whether the student had a learning disability; and whether the student had experienced disciplinary problems.[28]

To allow for as much flexibility as possible, separate logit models were estimated for treatment and control group members. For illustrative purposes, table 4A.3 reports the results in Washington, D.C. after two years. Similar results were obtained for other cities and other years.[29] For the most part, the family and student characteristics had a similar impact on re-

28. When baseline information was missing, means were imputed.

29. In each study (in New York City and Washington and in Dayton after Year I) the models include slightly different independent variables.

Table 4A.3 Logit Estimates Used to Construct Weights For Treatment and Control Groups in Washington, D.C. in Year Two

	Treatment Group	Control Group
Family characteristics		
Catholic	–0.5*	–0.8***
Family size	0.2**	0.2**
Employment status	–0.6**	–0.1
Married	–0.6***	–0.3
Mother's education	0.0	–0.1**
Welfare	–0.3	0.2
African American	0.8***	0.6***
Student characteristics		
Learning disabled	0.7**	–1.0**
Disciplinary problems	0.7**	0.7**
Math	–0.0	–0.0**
Constant	–1.1**	–0.6
Pseudo R^2	0.07	0.07
Log-likelihood	–353.11	–479.83
N	580	866

Notes: The dependent variable is coded 1 if the student attended the year-two follow-up session in Washington, D.C., and 0 otherwise. The treatment group consists of all students who were offered a voucher and participated in the baseline study; the control group consists of all students who were not offered a voucher.

***Significant at the 1 percent level, two-tailed test.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

sponse rates for both treatment and control group members. Catholics were less likely to attend follow-up sessions, as were mothers who were employed full time or were married. Larger families were more likely to attend follow-up sessions, as were African American families and families of students with disciplinary problems. Mother's education, welfare benefits, and math scores had a small or insignificant impact for both treatment and control group members. The most striking difference between the two models concerned students with learning disabilities. Although learning disabled students in the treatment group were significantly more likely to attend follow-up sessions, such students in the control group were significantly less likely to attend follow-up sessions.

The models generated a set of predicted values that represent the probability that individuals, given their baseline characteristics, would attend the follow-up session. The weights are simply the inverse of these predicted values, that is,

$$W_j = \frac{1}{F(X\beta)},$$

where $F(\,)$ is the model's logistic distribution function. The range of possible values for W_j was then capped so that the highest weight was four times the value of the lowest. (This restriction affected only a handful of observations.) The weights were then rescaled so that the sum of the weights equaled the sum of the total number of actual observations.[30]

To generate the weights, we could use only observable characteristics as recorded in parent surveys. To the extent that there were unmeasured or unobservable characteristics that encouraged some families, but not others, to attend follow-up sessions, the weights may not have eliminated the bias associated with nonresponse. However, in order for response bias to explain our findings, three conditions would have to hold. First, respondents would need to differ from nonrespondents on an unmeasured factor that influenced test performance. Second, the difference would have to be larger for one group (treatment or control) than for the other. Third, the difference would have to hold for black students but not for students of other ethnic groups. Although we cannot rule out the possibility that all three conditions existed in our study, we find it unlikely enough to be reasonably confident that response bias did not artificially generate the results we report.

It is possible that change in academic performance over time rather than baseline characteristics affected the likelihood that different subgroups within the treatment and control groups would attend subsequent testing sessions. If treatment group families that did not benefit from vouchers dropped out of the study while control group families that were suffering most in public schools continued to attend follow-up sessions consistently, then observed impacts may be somewhat inflated.

Three questions deserve consideration. Did gains in test scores from baseline to year one (two) decrease the probability that members of the control group would attend the year-two (-three) testing session? Did gains increase the probability that members of the treatment group would attend the year-two (-three) testing session? And were the differences in observed impacts on response rates for the treatment and control groups statistically significant?

Table 4A.4 estimates a series of logistic regressions that answer these questions. The dependent variable identifies whether a student attended the year-two (-three) follow-up session. The covariates include baseline math and reading test scores, the change in the total test score from baseline to year one (two), and the change interacted with treatment status. Separate models were run for African Americans and members of other ethnic groups. At the bottom of each column we report the probability that we can reject the following three null hypotheses: (a) Changes in test scores have a statistically insignificant effect on attendance at subsequent testing sessions for the control group; (b) the effect for the treatment group is statistically insignificant; and (c) the differences in observed effects for the two groups are not statistically significant.

30. Weights for New York City, as provided by Mathematica Policy Research.

Table 4A.4 **Effect of Change in Test Scores from Baseline to Year One and Two on the Likelihood That Students Attend Subsequent Testing Sessions**

	New York City				Washington, D.C.	Dayton	
	Year Two Attendance		Year Three Attendance		Year Two Attendance	Year Two Attendance	
	Blacks	Latinos	Blacks	Latinos	Blacks	Blacks	Whites
Y1-baseline	–0.005	–0.001	. . .	0.002	–0.012	0.001	
Y2-baseline	. . .	. . .	–0.023*	–0.004	. . .	. . .	. . .
(Y1-B)*treat	0.004	0.006	. . .	. . .	0.002	0.023*	–0.004
(Y2-B)*treat	. . .	. . .	0.031*	0.004	. . .	. . .	. . .
Baseline math	–0.003	0.012*	0.007	0.002	–0.001	0.001	–0.003
Baseline reading	0.007	–0.004	0.006	–0.008	–0.001	0.003	0.004
N	623	709	497	612	891	298	108
Log-likelihood	–355.81	–355.68	–212.40	–191.13	–580.55	–189.52	–70.39
Pseudo-R^2	0.00	0.01	0.02	0.00	0.00	0.01	0.00
P for H_0: $B_1 = 0$	0.63	0.91	0.09	0.76	0.79	0.21	0.94
P for H_0: $B_1 + B_2 = 0$	0.90	0.57	0.50	0.96	0.50	0.28	0.79
P for H_0: $B_2 = 0$	0.77	0.46	0.07	0.78	0.79	0.07	0.79

Notes: Logit regression models performed on unweighted data. Y1-baseline refers to the change in the total math and reading test scores from baseline to year one; Y2-baseline refers to change from baseline to year two. (Y1-B)* offered voucher is an interaction term between one variable that is the difference between year one and baseline test scores and another variable that indicates whether a student was offered a voucher. The dependent variable is coded 1 if the student attended either the second- or third-year follow-up session.

***Significant at the 1 percent level, two-tailed test.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

On the whole, the signs of the coefficients are in the expected direction. Gains in test scores from baseline to years one and two increased the probability that members of the treatment group attended the subsequent testing session and decreased the probability for members of the control group. The only models that generated statistically significant impacts, however, were for African Americans in New York after three years and for African Americans in Dayton after two years. None of the observed impacts for Hispanics were statistically significant in any year or city.

The model that predicts year-three attendance for African Americans in New York City generated the largest effects. Holding all variables at their means, the model predicted that 83 percent of the students who attended the year-two session would attend the year-three session. An increase of 10 NPR points from baseline to year two translated into a 3 percentage point drop in the probability that a control group member would attend the year-three testing session and a 1 percentage point increase in the probability that a member of the treatment group would attend the year-three follow-up session. Unless weighting adjusted for these differences, this response pattern may have marginally contributed to the positive estimate of voucher impacts on test scores.

In New York City, eighty-two African American students who had attended the year-two testing session failed to show up in year three. The data presented above suggest that those individuals consisted disproportionately of control group members whose scores decreased from baseline to year two and treatment group members whose scores increased, possibly inflating the estimated impact of attending a private school. To further explore their influence on estimated impacts, we imputed year-three test scores for those individuals based on their treatment status, baseline test scores, test score changes between baseline and year two, and the year-three weights. Although the observed impacts do drop in magnitude, they remain statistically significant. When we examined only those African American students who attended the year-three follow-up session, the estimated impact of switching from a public to a private school at year three was 5.4 NPR points, with 515 observations and a t-statistic of 3.7. When we looked at the same population but then imputed year-three test scores for those students who showed up in year two but not in year three, the size of the estimated impact of attending a private school dropped to 8.0 NPR points, with a t-statistic of 3.0.[31]

Another way of estimating the effects of response rates on outcomes is to distinguish between earlier and later respondents. Not all participants came

31. King et al. (2001). The point estimates reported come from weighted regressions: Imputed weights for missing observations in year three were constrained to have positive values. Impacts generated from unweighted regressions using imputed values and observables also are comparable.

to the first testing session to which they were invited. Given that we know the dates when students came in for testing, we can generate exact estimates of the impact of attending a private school for smaller response rates. In year one in New York City, for instance, we had an 82 percent response rate. By successively dropping the portion of students who attended later testing sessions, we can readily calculate the impacts for lower response rates.

If observed positive impacts derive from imperfect response rates, we should expect the estimated impact of attending a private school to increase as response rates decline. Presumably, those students who benefit most from treatment should come earlier to the testing sessions, along with those students in the control group who were performing most poorly in public schools. Impacts of attending a private school, then, should be quite large for lower response rates. The differences between the two groups, however, should attenuate (and may actually switch signs) as response rates increase.

Table 4A.5 reports the estimated impact of attending a private school for African American students for variable response rates. In each row, the first column represents the estimated impact for the full sample of African American students who attended testing sessions. Subsequent columns provide estimates of impacts for lower response rates, based on when students came in for testing.

As can be seen in table 4A.5, the New York City estimates remained remarkably stable for different response rates. Had we stopped testing students in year one after the first 30 percent of the sample showed up, we would have recovered almost exactly the same findings that we did after another 52 percent participated: The point estimate for the first 30 percent of students to be tested was 5.7 percentile points, and it was 5.4 for the full sample. In New York City in years two and three, rather than decreasing as response rates improved, the estimated impacts actually became larger.

Table 4A.5 **Estimated Impacts of Attending a Private School for African Americans in New York City for Variable Response Rates (% of Respondents Attending Follow-Up Sessions)**

	30	40	56	60	66	70	82
Year one impact	5.7***	4.4***	4.2***	4.8***	5.0***	5.3***	5.4***
Year two impact	3.6	2.7	4.4**	3.2*	4.3**	. . .	. . .
Year three impact	4.2	6.9***	7.1***	8.3***	9.2***[a]	. . .	. . .

Notes: Weighted two-stage least squares regressions performed; treatment status used as instrument. Differential response rates calculated by including in the analysis only the relevant percentage of students to initially attend testing sessions.

[a]This is the test score impact for the full sample, at a 67 percent response rate.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level, two-tailed test.

Moving from a 30 percent response rate to a 66 percent rate, the estimated test score impact of attending a private school increased by roughly 1 NPR point and became statistically significant. From these findings, at least, there is little to suggest that we would have observed significantly different impacts had we managed to test a greater number of students in the treatment and control groups. If anything, these results suggest that we may have underestimated the true effects of switching from a public to a private school.

Appendix B
The Effects of the Offer of a Voucher

Tables 4B.1, 4B.2, and 4B.3 report estimated effects of a voucher offer on student test score performance in each city. These ITT effects are smaller than actual treatment effects, because many students who were offered vouchers did not make use of them, and others who were not offered vouchers found alternative ways of financing a private education. The percentages using a voucher in each city are reported in the text of this chapter.

Table 4B.1 Impact in New York of Being Offered a Voucher on Test Score Performance

Test Score Performance	Year One (Percentiles)	*N*	Year Two (Percentiles)	*N*	Year Three (Percentiles)	*N*
African Americans						
Overall	4.5**	622	3.3**	497	5.5***	519
	(1.2)		(1.5)		(1.4)	
Math	5.7***	622	3.1*	497	7.0***	519
	(1.5)		(1.9)		(1.7)	
Reading	3.3**	622	3.4**	497	4.0**	519
	(1.5)		(1.7)		(1.7)	
All other ethnic groups						
Overall	–1.3	812	–1.0	699	–2.3	729
	(1.3)		(1.5)		(1.5)	
Math	–2.4	812	–2.2	699	–1.7	729
	(1.7)		(2.0)		(1.9)	
Reading	–0.9	812	0.1	699	–2.9*	729
	(1.4)		(1.6)		(1.7)	

Notes: Weighted ordinary least squares regressions performed. All models control for baseline test scores and lottery indicators. Impacts expressed in terms of national percentile rankings. Standard errors reported in parentheses.

***Significant at the 1 percent level, two-tailed test.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 4B.2 Impact in Washington, D.C. of Being Offered a Voucher on Test Score Performance

Test Score Performance	Year One (Percentiles)	*N*	Year Two (Percentiles)	*N*	Year Three (Percentiles)	*N*
African Americans						
Overall	–0.3	891	3.8***	668	–0.5	656
	(1.1)		(1.2)		(1.2)	
Math	2.9**	891	4.3***	668	0.3	656
	(1.3)		(1.4)		(1.4)	
Reading	–3.6**	891	3.3**	668	–1.3	656
	(1.5)		(1.4)		(1.5)	
All other ethnic groups						
Overall	4.7	39	–0.1	42	–0.9	31
	(5.6)		(5.6)		(6.6)	
Math	5.5	39	4.1	42	–4.7	31
	(7.2)		(7.4)		(7.7)	
Reading	4.0	39	–4.3	42	2.9	31
	(8.0)	(5.7)	(9.1)			

Notes: Weighted ordinary least squares regressions performed. Standard errors reported in parentheses. All models control for baseline test scores; in year one models also control for initial testing session. Impacts expressed in terms of national percentile rankings.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 4B.3 Impact in Dayton of Being Offered a Voucher on Test Score Performance

Test Score Performance	Year One (Percentiles)	*N*	Year Two (Percentiles)	*N*
African Americans				
Overall	1.9	296	3.5*	273
	(2.0)		(2.0)	
Math	0.2	296	2.8	273
	(2.3)		(2.3)	
Reading	3.5	296	4.1*	273
	(2.4)		(2.3)	
All other ethnic groups				
Overall	0.7	108	–0.1	96
	(4.1)		(4.0)	
Math	–0.5	108	0.0	96
	(4.8)		(4.7)	
Reading	1.8	108	–0.2	96
	(4.5)		(4.4)	

Notes: Weighted ordinary least squares regressions performed. Standard errors reported in parentheses. All models control for baseline test scores. Impacts expressed in terms of national percentile rankings.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

References

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91:444–62.

Bridge, R. J., and J. Blackman. 1978. *Family choice in education.* Vol. 4 of *A study of alternatives in American education.* Santa Monica, Calif.: RAND Corporation.

Chaplin, Duncan D. 1998. Raising standards: The effects of high school math and science courses on future earnings. *Virginia Journal of Social Policy and Law* 6 (1): 111–26.

Chubb, John E., and Terry M. Moe. 1990. *Politics, markets, and America's schools.* Washington, D.C.: Brookings Institution.

Coerper, Lois H., and Shirley W. Mersereau. 1998. *Independent school guide for Washington, D.C. and surrounding area.* 11th ed. Chevy Chase, Md.: Independent School Guides.

Coleman, James S., Thomas Hoffer, and Sally Kilgore. 1982. *High school achievement.* New York: Basic Books.

Coulson, Andrew J. 1999. *Market education: The unknown history.* New Brunswick, N.J.: Social Philosophy and Policy Center and Transaction Publishers.

Elmore, Richard. 1990. Choice as an instrument of public policy: Evidence from education and health care. In *Choice and control in American education,* Vol. 1: *The theory of choice and control in American education,* ed. W. Clune and J. Witte, 285–318. New York: Falmer.

Evans, William N., and Robert M. Schwab. 1993. Who benefits from private education? Evidence from quantile regressions. University of Maryland, Department of Economics. Working Paper.

Figlio, David, and Joe Stone. 1999. Are private schools really better? In *Research in labor economics,* vol. 16, ed. Soloman W. Polachek and John Robst, 115–40. Stamford, Conn.: JAI Press.

Flanagan, Ann, Jennifer Kawata, and Stephanie Williamson. 2000. *Improving student achievement: What NAEP test scores tell us.* Santa Monica, Calif.: RAND Corporation.

Goldberger, Arthur S., and Glen G. Cain. 1982. The causal analysis of cognitive outcomes in the Coleman, Hoffer, and Kilgore report. *Sociology of Education* 55 (April-July): 103–22.

Greene, Jay P., William G. Howell, and Paul E. Peterson. 1998. Lessons from the Cleveland scholarship program. In *Learning from school choice,* ed. Paul E. Peterson and Bryan C. Hassel, 357–92, Washington, D.C.: Brookings Institution.

Greene, Jay P., Paul E. Peterson, and Jiangtao Du. 1998. School choice in Milwaukee: A randomized experiment. In *Learning from school choice,* ed. Paul E. Peterson and Bryan C. Hassel, 335–56. Washington, D.C.: Brookings Institution.

Hill, Jennifer, Donald B. Rubin, and Neal Thomas. 1998. The design of the New York school choice scholarship program evaluation. Paper presented at the American Political Science Association annual meeting. 31 August, Boston, Massachusetts.

Hoffer, Thomas, Andrew Greeley, and James Coleman. 1985. Achievement growth in public and Catholic schools. *Sociology of Education* 58 (April): 74–97.

Howell, William G., and Paul E. Peterson. 2000. School choice in Dayton, Ohio: An evaluation after one year. Harvard University, Program on Education Policy and Governance. Occasional paper, February.

Howell, William G., Paul E. Peterson, Patrick J. Wolf, and David Campbell. 2002.

The education gap: Vouchers and urban schools. Washington, D.C.: Brookings Institution.

Jencks, Christopher, and Meridith Phillips, ed. 1999. *The black-white test score gap.* Washington, D.C.: Brookings Institution.

King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 2001. Analyzing incomplete political science data: An alternative algorithm for multiple imputation. *American Political Science Review* 95 (1): 46–69.

Krueger, Alan. 1999. Experimental estimates of education production functions. *Quarterly Journal of Economics* 114:497–533.

Krueger, Alan, and Diane Whitmore. 2000. Would smaller classes help close the black-white achievement gap? Paper presented at conference entitled Closing the Gap: Promising Strategies for Reducing the Achievement Gap. February, Washington, D.C.

Metcalfe, Kim K., William J. Boone, Frances K. Stage, Todd L. Chilton, Patty Muller, and Polly Tait. 1998. A comparative evaluation of the Cleveland scholarship and tutoring grant program: Year one: 1996–97. Bloomington, Ind.: Indiana University School of Education, Smith Research Center.

Murname, Richard, John B. Willet, Yves Duhaldeborde, and John H. Tyler. 2000. How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management* 19 (4): 547–68.

Myers, David, Paul E. Peterson, Daniel Mayer, Julia Chou, and William P. Howell. 2000. School choice in New York City after two years: An evaluation of the school choice scholarships program. Harvard University, Program on Education Policy and Governance. Occasional paper, September.

Neal, Derek. 1997. The effects of Catholic secondary schooling on educational achievement. *Journal of Labor Economics* 15 (1): 98–123.

Peterson, Paul E. 1998. School choice: A report card. In *Learning from school choice,* ed. Paul E. Peterson and Bryan C. Hassel, 3–32. Washington, D.C.: Brookings Institution.

Peterson, Paul E., Jay P. Greene, William G. Howell, and William McCready. 1998. Initial findings from an evaluation of school choice programs in Dayton, Ohio and Washington, D.C. Harvard University, Program on Education Policy and Governance. Occasional paper, October.

Peterson, Paul E., and Bryan C. Hassel, ed. 1998. *Learning from school choice.* Washington, D.C.: Brookings Institution.

Peterson, Paul E., David E. Myers, Josh Haimson, and William G. Howell. 1997. Initial findings from the evaluation of the New York school choice scholarships program. Harvard University, Program on Education Policy and Governance. Occasional paper, November.

Peterson, Paul E., David E. Myers, and William G. Howell. 1998. An evaluation of the New York City school choice scholarships program: The first year. Harvard University, Program on Education Policy and Governance. Occasional Paper, October.

Peterson, Paul E., David E. Myers, William G. Howell, and Daniel P. Mayer. 1999. The effects of school choice in New York City. Chap. 12 in *Earning and learning: How schools matter,* ed. Susan B. Mayer and Paul E. Peterson. Washington, D.C.: Brookings Institution.

Rouse, Cecilia. 1997. Private school vouchers and student achievement: An evaluation of the Milwaukee parental choice program. Princeton University, Department of Economics. Working Paper.

Schneider, Mark, Paul Teske, Melissa Marschall, and Christine Roch. 1998. Tiebout, school choice, allocative and productive efficiency. Paper presented at the annual meeting of the American Political Science Association. 3–6 September, Boston, Mass.

U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics. 2000. *Common core of data, school years 1993–94 through 1997–98.* Washington, D.C.: U.S. Department of Education.

West, Martin R., Paul E. Peterson, and David E. Campbell. 2001. School choice in Dayton, Ohio after two years. Harvard University, Program on Education Policy and Governance. Occasional paper, August.

Wilms, Douglas J. 1985. Catholic school effects on academic achievement: New evidence from high school and beyond follow-up study. *Sociology of Education* 58: 98–114.

Witte, John F. 1996. School choice and student performance. In *Holding schools accountable: Performance-based reform in education,* ed. Helen F. Ladd, 149–76. Washington, D.C.: Brookings Institution.

———. 1997. Achievement effects of the Milwaukee voucher program. Paper presented at the annual meeting of the American Economics Association. January, New Orleans, La.

Wolf, Patrick J., William G. Howell, and Paul E. Peterson. 2000. School choice in Washington, D.C.: An evaluation after one year. Harvard University, Program on Education Policy and Governance. Occasional paper, February.

Wolf, Patrick J., Paul E. Peterson, and Martin R. West. 2001. Results of a school voucher experiment: The case of Washington, D.C. after two years. Harvard University, Program on Education Policy and Governance. Occasional paper, August.