Test-Score Effects of School Vouchers in Dayton, Ohio, New York City, and Washington, D. C.: Evidence from Randomized Field Trials

by

William G. Howell
Assistant Professor, Department of Political Science, University of Wisconsin

Patrick J. Wolf
Assistant Professor, Public Policy Institute, Georgetown University
Guest Scholar, The Brookings Institution

Paul E. Peterson
Director, Program on Education Policy and Governance, Harvard University
Senior Fellow, Hoover Institution, Stanford University

and

David E. Campbell
Research Associate, Program on Education Policy and Governance, Harvard University

August 2000

Paper prepared for the annual meetings of the American Political Science Association, Washington, D. C., September 2000.

The Program on Education Policy and Governance (PEPG) is located within the Taubman Center on State and Local Government, John F. Kennedy School of Government and within the Center for American Political Studies, Department of Government, Harvard University. Mailing address: Taubman 306, Kennedy School of Government, 79 J. F. Kennedy St., Harvard University, Cambridge, MA 02138 Phone: 617-495-7976/495-8312; Fax: 617-496-4428;
Website: http://data.fas.harvard.edu/pepg/

Test-Score Effects of School Vouchers in Dayton, Ohio, New York City, and Washington, D. C.: Evidence from Randomized Field Trials

(Executive Summary)

In the late 1990s, three privately-funded school voucher programs for students from low-income families were established in the Dayton, Ohio metropolitan area, New York City, and Washington, D. C. The New York City program, sponsored by the School Choice Scholarships Foundation (SCSF), was announced in the fall of 1996; students receiving vouchers entered private schools in the fall of 1997. Two additional programs were created one year later, one in the Dayton metropolitan area, sponsored by Parents Advancing Choice in Education (PACE), and one in D. C., sponsored by the Washington Scholarship Fund (WSF). WSF expanded a previously established program, originally created in 1993. In 1999, the Children s Scholarship Fund, a nationwide school-choice scholarship program, provided additional support to these programs.

The main findings from two-year evaluations of the three programs are as follows:

- In the three cities taken together, the average, overall test-score performance of African American students who switched from public to private schools was, after one year, 3.3 NPR points higher, and, after two years, 6.3 NPR points higher than the performance of the control group remaining in public schools. In each city, the difference after two years was statistically significant.

- No statistically significant effects, either positive or negative, were observed for students from other ethnic groups who switched from public to private schools.

- A difference of 6.3 NPR points in overall test performance is 0.33 standard deviations, generally thought to be a moderately large effect. Nationwide, differences between black and white test scores are, on average, approximately one standard deviation. The school voucher intervention, after two years, erases, on average, about one-third of that difference. If the trend line observed over the first two years continues in subsequent years, the black-white test gap could be eliminated in subsequent years of education for black students who use a voucher to switch from public to private school. But it remains to be seen whether the gains black students experienced after two years continue to increase over time.

- By comparison, the effect of two years of participation by African Americans in a class- size reduction randomized field trial in Tennessee, which reduced class size by seven students, was to improve test scores by 4.9 NPR points, or approximately 0.21 standard deviations. As another point of comparison, the RAND study of *Improving School Achievement* reports what are said to be

remarkable one-year gains in some states that have rigorous statewide testing programs (e. g., Texas and North Carolina) that are as much as 0.06 to 0.07 standard deviation[s] per year, or 0.12 to 0.14 standard deviations over two years. The effects of vouchers after two years, as observed here, are over twice as large.

- These results are from randomized field trials. Students initial abilities and family background generally do not influence the results, because students were randomly assigned to test and control groups. Furthermore, all results take into account initial ability levels.

- 42 percent of the students participating in the second year of the evaluation in New York City were African Americans. The percentages in Dayton and D. C. were 74 percent and 94 percent, 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. The percentages of whites in Dayton and D. C. were 24 percent and 1 percent, respectively. The remaining students came from a variety of other ethnic backgrounds.

- Results for African Americans did not vary significantly by subject matter. Average differences, as observed in the three cities together, between those attending private schools and the control group in public school were 6.2 NPR points in math, and 6.3 percentile points in reading.

- Results varied somewhat by city. Overall test score performance after two years by African American students switching to a private school, as compared to the control group, was, on average, 4.3 NPR points higher in New York City, 6.5 points higher in Dayton, Ohio, and 9.0 points higher in Washington, D. C.

- In D. C., older students switching to private schools had trouble adapting to their school in the first year, but recovered lost ground and gained substantially by the end of the second year. After one year, older African American students attending private schools trailed their public school peers in overall test performance by 9.0 points. But by the end of two years, this older group of African American students had combined test score performances that were 8.1 percentile points higher than those of the control group.

The vouchers could be used to attend any private school within the metropolitan area that the family chose. In Dayton, the vouchers could also be used to attend a public school outside the school district, but the few students who made this choice were excluded from the evaluation.

Over 20,000 students filled out initial applications for school vouchers in New York City, over 7,500 applied in Washington, D. C., and over 3,000 applied in Dayton,

Ohio. Because the demand exceeded the supply of vouchers available, vouchers in all three cities were awarded by lotteries that gave each family an equal chance of winning a voucher.

The voucher programs offered lottery winners annual scholarships of up to $1,700 to help pay tuition at a private elementary school for at least four years. Telephone applications were received in the fall and winter of the year prior to the first year of the voucher program. In response to invitations sent by the program operators, applicants attended verification sessions where eligibility was determined, students were given baseline tests, older students filled out short questionnaires, and adult family members completed longer questionnaires. The lotteries were held in April or May prior to the beginning of the next school year. The data reported in this paper are taken from student performances on tests administered at follow-up sessions one and two years after the beginning of the program.

Since scholarships were awarded by means of a lottery in each city, the evaluations of these three programs were all designed as randomized field trials, a research method characteristically used in medical research to determine the effectiveness of drugs or other interventions. When an evaluation takes the form of a randomized field trial, the group receiving the offer of a school voucher is, on average, essentially identical to the control group with which it is compared, the only difference between the two groups being the luck of the lottery draw. Any differences observed during the randomized field trial, therefore, may be attributed to the school the child attended, not to the child s initial ability and family background characteristics, which generally do not differ between the two groups.

Students included in the evaluation were entering grades 2-5 in New York City and grades 2-8 in Washington D. C. and Dayton. Only those students who had previously been attending public school were included in the evaluation. Students were tested on the Iowa Test of Basic Skills (ITBS). Each student was given a National Percentile Ranking (NPR) score in math and reading which may vary between 0 and 100. Nationwide, median student performance is 50. Results are reported for math, reading, and a combined score that is the average of the math and reading scores.

At this time the evaluation team is unable to explain why school vouchers have positive effects on African American students but no detectable effects on others. However, the evaluation team plans to explore this question by detailed examination of parental and student reports on school life collected at the time students were tested.

The evaluation of the voucher programs in the three cities is an activity of the Harvard Program on Education Policy and Governance (PEPG), which is jointly sponsored by the Taubman Center on State and Local Government, Kennedy School of Government, Harvard University and the Center for American Political Studies in the Faculty of Arts and Sciences, Harvard University. Paul E. Peterson, Henry Lee Shattuck Professor of Government and Director of PEPG at Harvard University and a senior fellow at the Hoover Institution, Stanford University, is the director of the evaluations of the

Dayton and Washington, D. C. programs. William Howell is Assistant Professor, Department of Political Science, University of Wisconsin. Patrick Wolf is Assistant Professor, Public Policy Institute, Georgetown University and Guest Scholar, The Brookings Institution. David Campbell is a PEPG research associate. The evaluation of the SCSF program in New York City is a collaborative effort jointly conducted by Mathematica Policy Research (MPR) and PEPG, Paul E. Peterson and David Myers, Senior Fellow, MPR, serving as co-principal investigators.

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. Findings and interpretation are those of the authors of the study and not necessarily those of either the sponsoring foundations or program operators.

Test-Score Effects of School Vouchers in Dayton, Ohio, New York City, and Washington, D. C.: Evidence from Randomized Field Trials

In the past decade considerable data have been collected on how school vouchers impact low-income families and their children.[1] Ten years ago, the only information available about this widely debated question came primarily from an experimental public-school choice program attempted in Alum Rock, California during the 1960s.[2] But in the early and mid-1990s, new voucher programs sprouted across the country in such cities as

---

[1] 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 co-operating 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 co-operation. 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 Mathematica Policy Research, who is 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 and Matthew Charles for their key roles in coordinating data collection efforts.

We received helpful advice from Paul Hill, Christopher Jencks, 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. Additional research assistance was provided by Rachel Deyette, Jennifer Hill, and Martin West; 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.

[2] R. J. Bridge and J. Blackman, *A Study of Alternatives in American Education: Vol. 4. Family Choice in Education (*Santa Monica, CA: Rand Corporation, 1978); Richard Elmore, "Choice as an Instrument of Public Policy: Evidence from Education and Health Care," in W. Clune & J. Witte, eds., *Choice and Control in American Education: Vol. 1. The Theory of Choice and Control in American Education* (New York: Falmer, 1990), pp. 285-318.

Milwaukee, Dayton, Cleveland, Indianapolis, San Antonio, Washington, D. C. and New York City. Initially, many of 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 underway, making 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.[3]

Despite their limitations, these early evaluations provided program operators and evaluation teams with valuable opportunities to learn the problems and pitfalls accompanying the study of school vouchers. Subsequent evaluations of voucher programs in Dayton, New York and Washington, D. C. 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 and two years later, the evaluation team

[3] 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, while another study found gains only in math, and a third found gains in neither subject. Jay P. Greene, Paul E. Peterson, and Jiangtao Du, School Choice in Milwaukee: A Randomized Experiment, in Paul E. Peterson and Bryan C. Hassel, eds., *Learning from School Choice* (Washington, D. C.: Brookings, 1998), pp. 335-56; Cecilia Rouse, Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program, Department of Economics, Princeton University, 1997; John F. Witte, Achievement Effects of the Milwaukee Voucher Program, paper presented at the 1997 annual meeting of the American Economics Association. On the Cleveland program, see Jay P. Greene, William G. Howell, and Paul E. Peterson, Lessons from the Cleveland Scholarship Program, in Paul E. Peterson and Bryan C. Hassel, eds., *Learning from School Choice* (Washington, D. C.: Brookings, 1998), pp. 357-92; Kim K. Metcalf, William J. Boone, Frances K. Stage, Todd L. Chilton, Patty Muller, and Polly Tait, A Comparative Evaluation of the Cleveland Scholarship and Tutoring Grant Program: Year One: 1996-97, School of Education, Smith Research Center, Indiana University, March 1998. Greene, Peterson, and Du, 1998 report results from analyses of experimental data; the other studies are based upon analyses of non-experimental data.

again tested the students and asked parents about their children s school experiences.[4] Any statistically significant differences between students offered a voucher and those not offered a voucher may be attributed to experiences at school, because average student initial abilities and family backgrounds are similar.

This paper reports the estimated effects of switching from a public to a private school on the test score performances of students after one and two years. Students who were evaluated entered private school in grades 2-5 in New York City and grades 2-8 in Dayton (and other parts of Montgomery County, Ohio) and Washington, D. C.[5] Specifically, the evaluation estimates the impact of the program on student test scores on the Iowa Test of Basic Skills (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; nationally, the median student s performance is at the 50th percentile.

---

[4] Results from the Dayton evaluation after one year are reported in William G. Howell and Paul E. Peterson, "School Choice in Dayton, Ohio: An Evaluation After One Year," Paper prepared for the Conference on Charters, Vouchers and Public Education, 2000, (Program on Education Policy and Governance, Kennedy School of Government, Harvard University, Cambridge). Website address: http://data.fas.harvard.edu/pepg/. First-year results for Washington are reported in Patrick J. Wolf, William G. Howell and Paul E. Peterson, " School Choice in Washington, DC: An Evaluation after One Year," (Paper prepared for the Conference on Charters, Vouchers and Public Education, 2000, sponsored by the Program on Education Policy and Governance, Kennedy School of Government, Harvard University, Cambridge, MA; Website address: http://data.fas.harvard.edu/pepg/. First-year results from the New York City evaluation are reported in Paul E. Peterson, David E. Myers, William G. Howell, and Daniel P. Mayer, The Effects of School Choice in New York City, in Susan B. Mayer and Paul E. Peterson, eds., *Earning and Learning: How Schools Matter* (Washington, D.C.: Brookings, 1999), Ch. 12.

[5] Baseline data from the D. C. and Dayton evaluations are reported in Paul E. Peterson, Jay P. Greene, William G. Howell and William McCready, "Initial Findings from an Evaluation of School Choice Programs in Dayton, Ohio and Washington, D. C." Paper prepared under the auspices of the Program on Education Policy and Governance, Harvard University, for presentation before the annual meetings of the Association of Public Policy and Management, New York City, NY October, 1998. The paper is available at http://data.fas.harvard.edu/pepg/. Baseline data for New York City are reported in Paul E. Peterson, David Myers, Josh Haimson, and William G. Howell, Initial Findings from the Evaluation of the New York School Choice Scholarships Program, Program on Education Policy and Governance, Taubman

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; 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. However, the programs differed in size, timing and certain administrative details. In this section we describe the main characteristics of the School Choice Scholarships Foundation (SCSF) program in New York City, the Parents Advancing Choice in Education (PACE) program in the Dayton metropolitan area, and the Washington Scholarship Fund (WSF) program in Washington, D. C.

SCSF Program in New York City

In February 1997 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 currently 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 twenty thousand 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 school

---

Center on State and Local Government, Kennedy School of Government, Harvard University, 1997. Report is available at http://data.fas.harvard.edu/pepg/.

lunch program. To ascertain eligibility, students and an adult member of their family were asked to attend verification sessions during which family income and their 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 second year, about 64 percent of these children were using a scholarship: 62 percent of the children had used a scholarship for two full years, 12 percent used one just in the first year, and 2 percent used a scholarship only in the second year.

PACE Program in Dayton, Ohio

In the spring of 1998, Parents Advancing Choice in Education (PACE), a privately-funded non-profit 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. Interested families were asked to call PACE, which took preliminary applications from over 3,000 students. PACE asked applicants to attend sessions where administrators verified their eligibility for a

scholarship, students took the Iowa Test of Basic Skills (ITBS), and parents completed questionnaires. Over 1,500 applicants attended these verification sessions in February, March and April 1998. The lottery was then held on April 29, 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 increased beginning in 1999 as a result of increased 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, 54 percent used the scholarship to attend a private school in the program's first year. Thirty-three schools accepted students who had not previously been attending a private school; 201 of these students attended twelve Roman Catholic schools, 14 attended a Lutheran school, 34 attended three other Christian schools, and 14 attended four secular, non-public schools.[6]

WSF 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 family'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. 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. 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. WSF awarded recipients with incomes at or below the poverty line vouchers that equaled 60 percent of tuition or $1,700, whichever was less. Families with income above the poverty line received smaller scholarships. The maximum amount of tuition support for high school students was $2,200. WSF has said that it will attempt to continue tuition support to the children in its program for at least three years and hopefully, if funds are available, until they complete high school. No family with income more than two-and-a-half 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 on April 29, 1998. WSF announced that it expected to award over one thousand scholarships, with a majority 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. In order to assist families in finding a school, WSF made extensive efforts during the summer months of 1998 to inform scholarship recipients of private school options.

---

[6] Information provided to the evaluation team by Parents Advancing Choice in Education, January 2000.

Of those students offered scholarships, 53 percent made use of them to attend a private school in the first year of the program. Of those who participated in the second year of the evaluation, 68 percent of the scholarship recipients attended Catholic schools, 20 percent attended other religious schools, 9 percent attended secular schools, and for 2 percent it was not possible to determine the school's affiliation.

At the end of the second year, 70 percent of the students in the control group attended a District of Columbia public school, 17 percent attended a charter school, 9 percent a religious school, one percent a private secular school, one percent a public school outside the District of Columbia, one percent a magnet school and one percent were home-schooled. The type of school attended by the remaining one percent could not be determined. Of those who declined the scholarship offered to them, 73 percent attended a District of Columbia school, 24 percent attended a charter school, one percent a public school outside the District of Columbia, and one percent attended a magnet school. For one percent of the decliners it was not possible to determine the type of school, and the remainder attended a variety of alternative schools.

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. The following section details the steps taken to collect the relevant information.

Baseline Data Collection

During the eligibility verification sessions attended by voucher applicants, students took the Iowa Test of Basic Skills (ITBS) in reading and mathematics in order to provide

baseline information on student performance prior to the beginning of the program. Students in kindergarten applying for a scholarship for first grade were not tested at baseline, however. The sessions took place during the months of February, March and April immediately prior to the voucher lottery. These sessions generally lasted about two hours. The sessions were held in private-school classrooms, where school teachers and administrators served as proctors under the overall supervision of the evaluation team and program sponsors. The producer of the ITBS graded the tests.[7] Students in grades four through eight also completed a short questionnaire inquiring 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 responses to the questionnaire would be held in strict confidence and would be used for statistical purposes only. Respondents had considerable time to complete their surveys, and administrators were available to answer questions about the meaning of particular items. Information from these surveys has been reported elsewhere.[8]

---

[7] 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.

[8] Howell and Peterson, 2000; Wolf, Howell and Peterson, 2000. Paul E. Peterson, David E. Myers, William G. Howell, and Daniel P. Mayer, The Effects of School Choice in New York City, in Susan B. Mayer and Paul E. Peterson, eds., *Earning and Learning: How Schools Matter* (Washington, D.C.: Brookings, 1999), Ch. 12. For detailed results from the second-year evaluation of New York City s voucher program, see David Myers, Paul E. Peterson, Daniel Mayer, Julia Chou, and William P. Howell, School Choice in New York City after Two Years: An Evaluation of the School Choice Scholarships Program, September 2000. Occasional Paper, Program on Education Policy and Governance, Taubman Center on State and Local Government, Kennedy School of Government, Harvard University). Report available at http://data.fas.harvard.edu/pepg/

Over 5,000 students participated in baseline testing in New York City. After vouchers were awarded, approximately 1000 families were selected at random from those who did not win the lottery to comprise a control group of approximately 960 families.[9]

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 students who were in 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 of students who were not then attending a private school. Follow-up testing and survey information was obtained only from families with children not in private schools at the time of application.

The Lottery

The evaluation team conducted the lottery in May 1997 in New York City and April, 1998 in Dayton and D. C. Program operators notified lottery winners in May of the year in which the lottery was conducted. If a family was selected, all children in that family entering eligible grades were offered a scholarship. In order to ensure that an adequate number of scholarships were given to students not currently attending a private

[9] Exact procedures for the formation of the control group are described in Jennifer Hill, Donald B. Rubin and Neal Thomas, The Design of the New York School Choice Scholarship Program Evaluation. Paper presented before the American Political Science Association annual meeting in Boston, MA, August 31, 1998.

school, separate lotteries were held in Dayton and D. C. for students currently in public and private schools. This procedure also assured random assignment to test and control groups of those families participating in the evaluation. Only those students who were in public schools at the time of the lottery are included in these evaluations.

Because many more families applied for scholarships in New York City than originally had been anticipated, the evaluation team randomly selected families for vouchers through a two-stage procedure. As families applied for vouchers, they were formed into groups on the basis of their application date. During the early stages, all families were invited to eligibility assessment and data collection sessions. However, after it became clear that more families would be attending these sessions than could be accommodated, the evaluation team began randomly selecting applicants, inviting only those selected to attend the sessions. After the first stage was completed, families who attended these sessions and met the eligibility requirements were then randomly selected for the scholarship group or the control group. To ensure that all families from the different groups had the same chance of being selected for a voucher, the evaluation team adjusted the second-stage selection probabilities to reflect the differential chances of being invited to the verification sessions.

In New York City, the final lottery was held in mid-May 1997. Mathematica Policy Research (MPR) administered the lottery; SCSF announced the winners. Within the parameters established by SCSF, all applicants had an equal chance of winning the lottery. 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 are not expected to differ significantly from members of the control group (those who did not win a scholarship). Baseline data confirm this expectation.[10] In D. C., there were no significant differences in baseline demographic characteristics or initial test scores. The baseline test scores of those entering grades two through eight who were offered a voucher averaged 30.4 national percentile points in reading and 23.8 in mathematics. Those not offered the scholarship scored, on average, 30.3 national percentile points in reading and 22.8 points in math. As in D. C., the demographic characteristics of those offered vouchers in Dayton did not differ significantly from the characteristics of those who were not offered a voucher.[11] However, those offered a voucher scored 6.5 percentile points lower in math and 3.1 points lower in reading than those not offered a scholarship, a statistically significant

[10] For additional baseline information on Washington, D. C. and Dayton, Ohio, see Paul E. Peterson, Jay P. Greene, William G. Howell, and William McCready, "Initial Findings from an Evaluation of School Choice Programs in Washington, D. C. and Dayton, Ohio," Occasional Paper, Program on Education Policy and Governance, Kennedy School of Government, Harvard University, October 24, 1998, Appendix. Paper prepared under the auspices of the Program on Education Policy and Governance, Harvard University, for presentation before the annual meetings of the Association of Public Policy and Management, New York City, NY October, 1998. Available at http://data.fas.harvard.edu/pepg/; for New York City, see Paul E. Peterson, David Myers, Josh Haimson, and William G. Howell, Initial Findings from the Evaluation of the New York School Choice Scholarships Program, Program on Education Policy and Governance, Taubman Center on State and Local Government, Kennedy School of Government, Harvard University, 1997. Report is available at http://data.fas.harvard.edu/pepg/.

[11] For a more extended discussion of the characteristics of applicants for the Dayton scholarship program, see Peterson, Greene, Howell and McCready, 1998.

difference.[12] Estimated effects of the program on subsequent test scores adjust for baseline test scores.

Collection of Follow-up Information

The second-year follow-up information was collected in New York City in the spring of 1999 and in Dayton and D. C. in the spring of 2000. Data collection procedures were similar across cities.

Because test-score results from the second-year 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 per cent of the students participating in the second year of the evaluation in New York City were African Americans. The percentages in Dayton and D. C. were 74 percent and 94 percent, 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. The percentages of whites in Dayton and D. C. were 24 percent and 1 percent, respectively. The remaining students came from a variety of other ethnic backgrounds.

New York City. To evaluate the effects of the SCSF program in New York City, Mathematica Policy Research (MPR) assembled two statistically equivalent groups of families: (1) a voucher group with 1,000 families and (2) a control group with 960 families. Procedures used to construct the two groups and to collect first-year follow-up information and the results from the evaluation of the first year of the program are

[12] Baseline test score information for New York City is contained in Myers, Peterson, Mayer, Chou, and Howell, 2000.

described elsewhere.[13] For the second-year follow-up, families were invited in April, May and June of 1999 to attend sessions during which students again took the ITBS in mathematics and reading. Adult members of their 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 in grades three through six were also asked to complete short questionnaires.

Testing and questionnaire administration procedures were similar to those that had been followed during the baseline and first year follow-up sessions. Both the voucher students and students in the control group were tested in locations other than the school they were currently attending.

Sixty-six percent of the students included in the evaluation attended the second-year testing sessions in New York City. Sixty-nine percent of those offered vouchers, as compared to 62 percent of the students in the control group, participated in these sessions. This fairly high response rate was achieved in part because 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 follow-up sessions. Detailed response rate information for the second follow-up survey and test, along with response rates for the baseline and the first follow-up surveys and test administrations are reported elsewhere.[14]

---

[13] Myers, Peterson, Mayer, Chou, and Howell, 2000. Also, see Peterson, Myers, Haimson, and Howell, 1997; Peterson, Myers, Howell and Mayer, 1999.

[14] Myers, Peterson, Mayer, Chou, and Howell, 2000. Although the background characteristics of participants and non-participants in the second year follow-up, as observed in the baseline survey conducted in 1997, resembled one another in most respects, they differed significantly in some. As compared to non-participants, participants were more likely to be non-Puerto Rican Hispanic. Mothers were more likely to be born outside the United States, more likely to have lived in the same residence, less likely to be working,

To adjust for survey non-response in our statistical analyses, we use an analytic model to predict non-response based on a variety of background characteristics. The predicted probability of not responding is then used to adjust the sample weights.

Dayton and Washington, D. C. In D. C. and Dayton, to estimate the impact of a private school on student test performance after one and two years of attendance at these schools, the evaluation team collected follow-up information between late February and late April, 2000.

The procedures used to obtain follow-up data were essentially the same as those used to collect baseline data, except that data were collected only from students who had not been in private school at the time of the initial scholarship application. Students again took the ITBS in mathematics and reading. Caretakers accompanying the child completed surveys that asked a wide range of questions about the educational experiences of each of their children. Students in grades four through eight also completed a questionnaire that asked them about their experiences at school. Testing and questionnaire administration procedures were similar to those that had been followed at baseline.[15]

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

---

more likely to state their religious affiliation as Catholic, less likely to use food stamps or welfare. They originally reported an average income of around $9,900, as compared to $8,500 for the non-participants. They were less likely to speak English at home.

Members of the control group who participated in the second-year follow up were less likely than non-participants to be black and more likely to be non-Puerto Rican Hispanic. They were more likely to report that their child had received help for a disability. They were more likely to have a Catholic religious affiliation. They were more likely to be receiving supplemental security income. They were less likely to speak English at home.

[15] Difficulties were encountered in the administration of the first-year test at the initial pilot session in Washington, D. C. Test booklets were not available at the testing site for scholarship students in grades 3-8. 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.

compensated for their expenses. They were also told in Washington, D. C. that they would be included in a new lottery, if they participated in the follow-up sessions. In Dayton, a second lottery was promised as a reward for participating in the first follow-up session but not for the second. Instead, families were given a higher level of compensation for participating in the follow-up session.

In D. C., follow-up survey information was obtained for 1,052 students from parents and caretakers. First-year follow-up test information was obtained from 995 students who had also been tested at baseline, a response rate of 63 percent. Of these students, 486 were members of the control group and 509 were members of the treatment group. In the second-year follow up, the overall response rate was 50.3 percent. The response rate for those offered scholarships was 50.5 percent; the response rate for the control group was 50 percent.

In Dayton, in the first-year follow-up, 57 percent of the students in the control group and 56 percent of those offered scholarships returned to take the reading and math tests. In the second year, 49.5 percent of the students in the control group and 47.8 percent of those offered scholarships returned to take the reading and math tests. The Appendix compares the characteristics of participants and non-participants in the second-year follow-up sessions.

In Dayton and D. C., as in New York, baseline demographic and test score information was used to adjust for non-response to requests for participation in follow-up sessions.[16]

[16] The Appendix to this report compares the characteristics of participants and non-participants in the second-year follow-up sessions in Dayton and Washington, D. C. For a discussion of the weighting procedures used in these evaluations, see Howell and Peterson, 2000; Wolf, Howell and Peterson, 2000. For New York City, see Myers, Peterson, Mayer, Chou, and Howell, 2000.

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 students 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.

This report provides data that help answer two questions. The first question is as follows:

> What was the impact on test-score performances of students from low-income families residing within a large central city one and two years after an *offer* of a voucher?

This question can be answered straightforwardly by comparing the test scores of those who were offered a scholarship with the test scores of the control group. Because scholarships were awarded at random, the two groups may be assumed to be, on average, equivalent statistically, save the offer of a scholarship. Any differences between the two groups can, then, be attributed to the scholarship offer.

To compute program impacts on children s test scores, we estimated a statistical model that took into account students scholarship or control-group status as well as baseline reading and math test scores. Baseline test scores were included to: 1) adjust for minor baseline differences between the treatment and control groups on the achievement tests; and 2) to increase the precision of the estimated impacts.

For some policy analysts, this first question is the crucial policy question: What happens when a school choice program is put into effect? How does the program impact

the population of low-income families who were offered a school-choice scholarship? This query is similar to a question often asked in medical research: What will happen if a particular pill is marketed? How will the health of potential users be altered, whether or not all patients use the pill as prescribed?

This analytic strategy has certain methodological advantages because calculation of the impact of the scholarship offer is quite straightforward. However, generalization from these results alone has the important disadvantage of assuming that usage rates of scholarships are fixed when in fact they might be highly variable, depending upon the size of the scholarship, the time the scholarship is offered, and the marketing of the program as a whole. Also, if programmatic impacts are substantial, participation rates may increase with the passage of time.

For these reasons, most analysts want an answer to the second question as well:

> What was the impact on test-score performances of students from low-income families residing within a large central city one and two years after switching from a public to a private school?

In medical research, the parallel question is: What are the consequences of actually taking a pill, as prescribed? In the case of the education intervention evaluated here, the answer to this second question requires a comparison between those switching from a public to a private school and a comparable control group who remained in a public school.[17]

---

[17] To compute the program s impact on those who used a scholarship to attend a private school, we used an instrumental variable estimator, which provides an unbiased estimate of the effects of switching to a private school. This procedure is discussed in Joshua D. Angrist, Guido W. Imbens, and Donald B. Rubin, "Identification of Causal Effects using Instrumental Variables," *Journal of the American Statistical Association*, 91 (1996), 444-462. 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 Alan Krueger, Experimental Estimates of Education Production Functions. *Quarterly Journal of Economics* , 114 (1999), 497-533.

Although we present information describing the effects of an *offer* of a scholarship, the text of this report will discuss, for the most part, the impact on students in the first and second years of a switch from a public to a private school. Second-year results compare those in private schools for two years with a control group that was not in private school for two years. Although most of the control-group students had never attended a private school, some had attended for one year.

Generalizations from the Findings

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 directed at low-income families initially 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, while 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/or public authorities.

Test Score Findings

The results reported below provide the first information from randomized field trials on the effects of school vouchers over a two-year period of time from three sites. However, they build upon a body of research that has explored the differences between schooling for low-income minorities in the public and private sectors.

Prior Research

Several studies have compared the test performance of students in public and private schools, and they usually find that low-income and African American students attending private schools outperform their public-school peers. Information on the effects of attendance at a Catholic high school are contained in a recent University of Chicago analysis of 12,000 students in the National Longitudinal Survey of Youth, conducted by the Department of Education. This report finds that, 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, but the effects are the greatest among urban minorities. The probability of graduating from college rises from 11 to 27 percent, if such a student attends a Catholic high school.[18] This study s findings are consistent with other studies.[19] After reviewing the literature on school effects on learning, University of Wisconsin Professor John Witte concludes that studies of private schools "indicate a substantial private school advantage in terms of completing high school and enrolling in college, both very important events in

[18] Derek Neal, "The Effects of Catholic Secondary Schooling on Educational Achievement," (Harris School of Public Policy, University of Chicago and National Bureau for Economic Research, 1996), p. 26.
[19] William N. Evans and Robert M. Schwab, "Who Benefits from Private Education? Evidence from Quantile Regressions," (Department of Economics, University of Maryland, 1993); David Figlio and Joe Stone, "School Choice and Student Performance: Are Private Schools Really Better?" (University of Wisconsin Institute for Research on Poverty, 1977).

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."[20]

Even the most careful of these studies, however, can take into account only observed family background characteristics. They cannot be sure that they have taken into account an intangible factor the willingness of a family 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.[21] In the jargon of the research community, this is called the self-selection problem, a problem for researchers that arises when a population differentiates itself by freely selecting a particular situation, in this case, private school. How can one be sure that the findings are not due to the self-selected character of the population, not the education intervention?

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. And in the state-funded program in Cleveland, though scholarship winners were initially selected by means of a

---

[20] John F. Witte, "School Choice and Student Performance," in Helen F. Ladd, ed., *Holding Schools Accountable: Performance-Based Reform in Education* (Washington, D. C.: Brookings, 1996), p. 167.

[21] Major studies finding positive educational benefits from attending private schools include James S. Coleman, Thomas Hoffer, and Sally Kilgore, *High School Achievement* (New York: Basic Books, 1982); John E. Chubb and Terry M. Moe, *Politics, Markets, and America s Schools* (Washington: Brookings 1990); Derek Neal, The Effects of Catholic Secondary Schooling on Educational Achievement, (University of Chicago, Harris School of Public Policy and National Bureau for Economic Research, 1996). Critiques of these studies have been prepared by Arthur S. Goldberger and Glen G. Cain, The Causal Analysis of Cognitive Outcomes in the Coleman, Hoffer, and Kilgore Report, *Sociology of Education* , vol. 55 (April-July 1982), pp. 103-22; Douglas J. Wilms, Catholic School Effects on Academic Achievement: New Evidence from the High School and Beyond Follow-up Study, *Sociology of Education* , 58 (1985), 98-114.

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.[22]

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 two-thirds of the students who participated in the lottery, and baseline data provided information that allowed the analysts to adjust for non-response.

Effects of Switching from a Public to a Private School

The average test score results from all three cities provide a better indication of the effects of switching from a public to a private school than do the results from any one city, because minor fluctuations in data collection may influence results in any one site. 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 .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 random, idiosyncratic

[22] Results from these evaluations are reported in Paul E. Peterson and Bryan C. Hassel, eds., *Learning from School Choice* (Brookings, 1998).

variations in observations of student performance.[23] For these reasons, the average from all three cities of student overall test score performance, as presented in Table 1, provides the best available information concerning the effect on test scores of switching from a public to a private school by students from low-income families.

As can be seen in Table 1, overall results differ depending on whether the student is African American or from some other ethnic group. One finds 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 students in the control group--either after one or two years. Nor were significant differences observed in the test-score performance of these students on reading and math tests, considered separately. Nor were significant differences observed for these students in any one of the three cities.

The effects of switching to a private school on African American students, however, 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, 3.3 NPR points higher on the combined math and reading tests, and, after two years, 6.3 percentile points higher, than the African American students in the control group. As can be seen in Tables 2A, 2B, and 2C, these differences in combined test score performance were statistically significant in all three cities. These are the average results for the three cities combined, weighting each city in inverse proportion to the standard error of the estimate for that city. Unweighted estimates for the impact on black students after two years are slightly larger.[24]

---

[23] This procedure was also employed in Krueger. 1999.

[24] The average unweighted impact of switching schools is, for blacks, 2.7 after one year, 6.6 after two years. For math, unweighted impacts are 4.8 after one year, 6.5 after two years; for reading, 0.6 after one year, and 6.8 after two years.

Although the overall test performance is the most reliable information, Table 1 also shows that, for the three cities taken together, differences after two years are approximately the same for the reading and math tests. On average, African American students in the three cities who switched from public to private schools achieved 6.3 percentile points higher on the reading test and 6.2 points higher on the math test than did the African American students in the control group.

The findings for each city are reported in tables 2A, 2B, and 2C. No effects on students from ethnic backgrounds other than African American were observed in any city. The largest differences between African American students who switched from public to private schools and those in the control group, were observed in Washington, D. C. In this city, black students attending private schools for two years scored 9.0 percentile points higher on the two tests combined than did students in the control group. The smallest differences after two years were observed in New York City. In this city, African American students attending private schools scored 4.3 percentile points higher on the reading and math tests combined. In Dayton, the difference in combined test-score performance was 6.5 percentile points, nearly at the mid-point between the differences observed in the other two cities.

The trend over time also varies from one city to the other. As can be seen in Table 2A, 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 but then attenuate slightly in the second year. The combined score difference after two years is 4.3 percentile points, which is slightly but not significantly (in statistical terms) less than the 5.8 percentile points observed after one year. In this city, one may reasonably conclude

that the initial gains from the school voucher program for African Americans are preserved but do not increase between year one and year two.

In Dayton, there seems to be a steady upward trend in the combined test score performance of African Americans in reading and math. Table 2C shows that African American students who switched from public to private school performed 3.3 percentile points higher on the combined test in year one and 6.5 percentile points higher in year two.[25]

In some ways the most striking results are for African Americans in Washington, D. C. As can be seen in Table 2B, in general, no significant differences were observed in year one, but a large impact of attendance at a private school was observed after two years. Also in D. C., clear differences between the impact of the program on older and younger students were observed in year one.[26] Younger students may have benefited slightly from the voucher program after one year, but older students did not. In fact the older students who switched to private schools performed poorly at the end of the first year. Many of these students were quite unhappy with their new school, as they communicated frankly in their questionnaire responses.[27] We think the low performance of these students in year one reflects more their discontent than their actual achievement level. By the end of the

[25] If those whose scores jumped or dropped dramatically between baseline and year one are excluded from the analysis, then the gains in year one are larger than those reported here. See Howell and Peterson, 2000. Now that data are available for two years, we have chosen not to exclude these students from the analysis, because it is more difficult to justify such exclusions after two years than after just one. After all, students might make striking gains that are real or suffer genuinely serious losses--over a two-year time period. Changes of this magnitude over one year seem less plausible. Given our decision not to exclude cases with significant changes in year two, it was desirable, for the sake of consistency, to apply the same framework to the analysis of year-one data.

[26] The D.C. program offered the only opportunity to examine the effect on test scores of an offer of a school voucher to older students. While vouchers were offered to middle-school students in Dayton, there were not enough cases to justify a separate analysis.

[27] Parent and student surveys corroborate this intuition. Black students in grades 6-8 who attended private schools expressed less satisfaction and lower morale, and reported a higher frequency of expulsions and fewer friends than students attending private schools in grade 2-5. See Wolf, Peterson, and Howell, 2000.

second year, the transformation in the test-score performance of these same students suggests that this discontent seems to have subsided. When data from the responses to the student questionnaire for the second year become available, it will be possible to ascertain whether or not this in fact was the case. In any case, by the end of year two, younger and older African American students were benefiting similarly from the switch to private schools. Younger students who had changed schools performed on the combined test 9.3 percentile points higher than those remaining in public schools. Older African American students in private school scored 10.3 percentile points higher.

Effects of the Voucher Offer

Table 3 provides information on the effects of a voucher offer in the three cities. In general, these effects, after two years, are about half those of actually switching to a private school for the simple reason that only about half the students offered a voucher switched to and continued to remain in a private school. In the case of ethnic groups other than African Americans, no significant effects of a voucher offer were observed. For African Americans, however, statistically significant effects on combined test scores were observed after two years in all three cities. The average effect of an offer of a voucher in the three cities on combined test scores was 2.0 percentile points after one year, and 3.5 percentile points after two years. In math the average effect in the three cities was 3.2 points after one year, and 3.4 points after two years. In reading, the average effect in the three cities was 0.8 after one year and 3.6 after two years. Tables 4A, 4B, and 4C provide information by city on the effect of a simple offer of a voucher on test scores.

As suggested previously, the effects of an offer, as reported in these tables, are approximately one-half the size of the effects of switching to a private school simply

because only about one-half the students offered a voucher were using the voucher at the end of the second year. Low-income families in central cities are a highly mobile population and their resources are limited. The vouchers offered by the program sponsors in these three cities did not exceed $1,700 and were often less. Presumably, in a larger, government-funded voucher program with an unlimited time span, voucher participation rates and, therefore, voucher offer effects--would be larger and increase over time, especially if gains from vouchers were apparent to families and the number of spaces in private schools increased.

Discussion

Randomized field trials are the best available tool for detecting the effects of an educational intervention, because random assignment to test and control groups assures that all significant effects may be attributed to the intervention, not to the students initial abilities or their family backgrounds. Nonetheless, when interpreting 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. But 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 a larger context. 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 students from African American backgrounds. Black 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 Iowa Test of Basic Skills than comparable blacks who remained in public schools.

These effects are moderately large. As can be seen in Table 5, black students who switch to private schools score, after one year, 0.17standard deviations higher than the students in the control group. After two years, the size of the effect grows to 0.33 standard deviations, about a third of the difference in test score performances between blacks and whites. If this trend line should continue in subsequent years, even at an attenuated rate, it would eliminate the test-score difference between blacks and whites for those who switch to private schools.[28] Continuing evaluation of voucher programs may provide information on whether or not these gains can be consolidated and extended.

Another way of obtaining a sense of the magnitude of these effects is to compare them to the effects 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 larger than the estimated effect of a 7-student reduction in class size. According to a recent reanalysis of data from Tennessee, the class-size reduction effect for African Americans after two years was, on average, 4.9 percentile points, somewhat less than the 6.7 percentile effect of switching to a private school.[29]

Another way of thinking about the magnitude of the effects observed here is to compare them with effects reported in the RAND study entitled *Improving School*

[28] Christopher Jencks and Meridith Phillips, eds., *The Black-White Test Score Gap* (Washington, D. C.: Brookings, 1999).
[29] Krueger, p. 525.

Achievement released in August 2000.[30] Identifying the most successful states, Texas and North Carolina, which have introduced rigorous accountability systems that involve state-wide 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 0.12 to 0.14 over two years. The two-year effects of the school voucher intervention on black students observed here are over twice as large.

At this point we do not know why the gains from switching to a private school are evident for black students after two years, but not for students from other ethnic backgrounds. In earlier reports from these evaluations, parents have 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 slightly smaller classes (about 3 fewer pupils). It remains to be seen whether any or all of these factors are especially associated with black test-score performance. Given the widespread concern about racial differences in academic performance, our research is particularly salient in that it suggests that school voucher programs may have the capacity to shrink the black-white test-score gap for participating students. We plan to explore this topic further in future reports.

Further studies are also needed to ascertain whether the benefits for African Americans observed when voucher plans are small and experimental also occur when much larger voucher interventions are undertaken.

[30] Also, see Ann Flanagan, Jennifer Kawata and Stephanie Williamson. 2000. *Improving Student Achievement: What NAEP Test Scores Tell Us* (Santa Monica, CA: RAND Corporation, 2000), p. 59.

TABLE 1: The Impact in Three Cities of Switching to a Private School on Test Score Performances

Test Score Performance	Year 1 (Percentiles)	Year 2 (Percentiles)
African Americans		
Overall	3.3	6.3**
Math	5.5*	6.2*
Reading	1.3	6.3**
All Other Ethnic Groups		
Overall	0.2	-1.0
Math	-0.2	-1.2
Reading	0.4	-0.8

** significant at the .05 level, two tailed test; * .10 level. 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 standard errors of the point estimates. 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.

TABLE 2A: Impact in New York of Switching to a Private School on Test Score Performance

Test Score Performance	Year 1 (Percentiles)	(N)	Year 2 (Percentiles)	(N)
African Americans				
Overall	5.8**	623	4.3**	497
Math	7.0***	623	4.1*	497
Reading	4.6**	623	4.5**	497
All Other Ethnic Groups				
Overall	-1.7	817	-1.5	699
Math	-2.1	817	-3.2	699
Reading	-1.3	817	0.2	699

* significant at .10 level, 2-tailed test; ** .05 level; *** .01 level. Weighted two-stage least squares regressions performed; treatment status used as instrument. All models control for baseline test scores and lottery indicators. Impacts expressed in terms of national percentile rankings. 2.8 percent of the African American control group in the year 2 models attended a private school for one of two years. When using bootstrapped standard errors, the year 2 math score is not statistically significant; the significance levels of all other estimates remain the same when significance levels are estimated using the bootstrap technique. See Robert Stine, 1990. "An Introduction to Bootstrap Methods: Examples and Ideas" in J. Fox and J. S. Long, eds., Modern Methods of Data Analysis.˚ p. 325-373.˚ Newbury Park. CA: Sage Publications. Bradley Effron, 1982. "The Jackknife, the Bootstrap and Other Resampling Plans." Philadelphia, PA: Society for Industrial and Applied Mathematics.

TABLE 2B: Impact in D.C. of Switching to a Private School on Test Score Performance

Test Score Performance	Year 1 (Percentiles)	(N)	Year 2 (Percentiles)	(N)
African Americans				
Overall	-0.9	891	9.0***	700
Math	7.3**	891	9.9***	700
Reading	-9.0**	891	8.1**	700
All Other Ethnic Groups				
Overall	7.4	39	0.1	44
Math	8.5	39	5.8	44
Reading	6.3	39	-5.6	44
African Americans, Grades 2-5				
Overall	2.4	620	9.3***	490
Math	9.8***	620	10.0***	490
Reading	-5.1	620	8.6**	490
African Americans, Grades 6-8				
Overall	-8.8*	270	10.3*	210
Math	1.5	270	12.8*	210
Reading	-19.0***	270	7.8	210

* significant at .10 level, 2-tailed test; ** .05 level; *** .01 level. Weighted two-stage least squares regressions performed; treatment status used as instrument. All models control for baseline test scores; in year 1, models also control for initial testing session. Impacts expressed in terms of national percentile rankings. Grade levels refer to 1998-1999 school year. 3.7 percent of the African American control group in the year 2 models attended a private school in the second year but not the first year.

TABLE 2C: Impact in Dayton of Switching to a Private School on Test Score Performance

Test Score Performance	Year 1 (Percentiles)	(N)	Year 2 (Percentiles)	(N)
African Americans				
Overall	3.3	296	6.5*	273
Math	0.4	296	5.3	273
Reading	6.1	296	7.6*	273
All Other Ethnic Groups				
Overall	1.0	108	-0.2	96
Math	-0.8	108	0.0	96
Reading	2.8	108	-0.4	96

* significant at .10 level, 2-tailed test; ** .05 level; *** .01 level. 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. 2.0 percent of the African American control group in the year 2 models attended a private school in the second but not the first year.

TABLE 3: The Impact in Three Cities of Being Offered a Voucher on Test Score Performances

Test Score Performance	Year 1 (Percentiles)	Year 2 (Percentiles)
African Americans		
Overall	2.0	3.5**
Math	3.2**	3.4*
Reading	0.8	3.6**
All Other Ethnic Groups		
Overall	0.1	-0.6
Math	-0.2	-0.8
Reading	0.3	-0.4

Figures represent the average impact of being offered a voucher 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 standard errors of the point estimates. For African Americans, the unweighted average effects after one year are 2.0 overall, 2.9 in math, and 1.1 in reading; after two years, the unweighted average effect sizes are 3.5 overall, 3.3 in math, and 3.6 in reading.

TABLE 4A: Impact in New York of Being Offered a Voucher on Test Score Performance

Test Score Performance	Year 1 (Percentiles)	(N)	Year 2 (Percentiles)	(N)
African Americans				
Overall	4.5**	642	3.3**	497
Math	5.4***	642	3.1*	497
Reading	3.5**	642	3.4**	497
All Other Ethnic Groups				
Overall	-1.2	817	-1.0	699
Math	-1.5	817	-2.2	699
Reading	-0.9	817	0.1	699

* significant at .10 level, 2-tailed test; ** .05 level; *** .01 level. Weighted OLS regressions performed. All models control for baseline test scores and lottery indicators. Impacts expressed in terms of national percentile rankings. When using bootstrapped standard errors, the year 2 math score is not statistically significant; When using bootstrapped standard errors, the year 2 math score is not statistically significant; the significance levels of all other estimates remain the same when significance levels are estimated using the bootstrap technique. See Robert Stine, 1990. "An Introduction to Bootstrap Methods: Examples and Ideas" in J. Fox and J. S. Long, eds., Modern Methods of Data Analysis.˚ p. 325-373.˚ Newbury Park. CA: Sage Publications. Bradley Effron, 1982. "The Jackknife, the Bootstrap and other Resampling Plans." Philadelphia, PA: Society for Industrial and Applied Mathematics.

TABLE 4B: Impact in D.C. of Being Offered a Voucher on Test Score Performance

Test Score Performance	Year 1 (Percentiles)	(N)	Year 2 (Percentiles)	(N)
African Americans				
Overall	-0.3	891	3.6***	700
Math	2.9**	891	4.0***	700
Reading	-3.6**	891	3.3**	700
All Other Ethnic Groups				
Overall	4.7	39	0.0	44
Math	5.5	39	3.2	44
Reading	4.0	39	-3.1	44
African Americans, Grades 2-5				
Overall	1.0	620	3.9***	490
Math	4.2***	620	4.2***	490
Reading	-2.2	620	3.6**	490
African Americans, Grades 6-8				
Overall	-3.0*	270	3.3*	210
Math	0.5	270	4.1*	210
Reading	-6.5***	270	2.5	210

* significant at .10 level, 2-tailed test; ** .05 level; *** .01 level. Weighted OLS regressions performed. All models control for baseline test scores; in year 1, models also control for initial testing session. Impacts expressed in terms of national percentile rankings. Grade levels refer to 1998-1999 school year.

TABLE 4C: Impact in Dayton of Being Offered a Voucher on Test Score Performance

Test Score Performance	Year 1 (Percentiles)	(N)	Year 2 (Percentiles)	(N)
African Americans				
Overall	1.9	296	3.5*	273
Math	0.2	296	2.8	273
Reading	3.5	296	4.1*	273
All Other Ethnic Groups				
Overall	0.7	108	-0.1	96
Math	-0.5	108	0.0	96
Reading	1.8	108	-0.2	96

* significant at .10 level, 2-tailed test; ** .05 level; *** .01 level. Weighted OLS regressions performed. All models control for baseline test scores. Impacts expressed in terms of national percentile rankings.

Table 5: Size of the Effects of Switching to a Private School on African Americans Overall Test Score Performances

Test Score Performance	Effect Size Year One (Standard Deviations)	Effect Size Year Two (Standard Deviations)
Overall	0.17	0.33
Math	0.29	0.30
Reading	0.07	0.26

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.

Appendix
Baseline Characteristics of Respondents and Non-Respondents in Second-Year Follow-Up Testing Sessions

Washington D. C.

	OFFERED VOUCHER (TREATMENT)		NOT OFFERED VOUCHER (CONTROL)	
	Attended Session	Didn t Attend Session	Attended Session	Didn t Attend Session
% 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
Test Scores (ave.)	26.5	26.4	26.9	26.7
Family Size (ave.)	3.1	3.1	3.3	3.0
Residential Mobility (ave.)	3.4	3.5	3.5	3.4
Church Attendance (ave.)	3.7	3.5	3.7	3.7
School Satisfaction (ave.)	2.5	2.7	2.7	2.6
Mother s Education (ave.)	5.4	5.0	5.3	5.2

Items used to generated year 2 weights in D.C. Averages refer to the mean score of scaled items on survey.

Baseline Characteristics of Respondents and Non-Respondents in Second-Year Follow-Up Testing Sessions (Continued)

Dayton

	OFFERED VOUCHER (TREATMENT)		NOT OFFERED VOUCHER (CONTROL)	
	Attended Session	Didn t Attend Session	Attended Session	Didn t Attend Session
% Catholic	7.1	13.7	13.2	17.6
% Protestant	65.5	60.1	64.6	59.1
% Employed Full Time	56.1	51.3	54.8	52.6
% Welfare Recipients	22.5	23.1	24.8	21.7
% Learning Disabled	12.3	6.2	4.0	9.4
% African American	73.0	66.7	70.4	66.7
Mother s Education (ave.)	6.0	5.9	5.7	5.8
Church Attendance (ave.)	3.4	3.3	3.4	3.6
Residential Mobility (ave.)	3.4	3.4	3.6	3.7
Parental Involvement (ave.)	4.4	4.5	4.4	4.4
Test Scores (ave.)	26.3	24.2	27.4	24.3

Items used to generated year 2 weights in Dayton. Averages refer to the mean score of scaled items on survey.