

LOST OPPORTUNITIES: LAWMAKERS THREATEN D.C. SCHOLARSHIPS DESPITE EVIDENCE OF BENEFITS

By Patrick J. Wolf

August 24, 2009

Unabridged version of article published in *Education Next* 9:4, Fall 2009, pp. 48-56

School choice supporters, including hundreds of private school students in crisp uniforms, filled Washington, D.C.'s Freedom Plaza last May to protest a congressional decision to eliminate the city's federally funded school voucher program after the next school year. That afternoon, President Obama announced a compromise proposal to grandfather the more than 1,700 students currently in the District of Columbia Opportunity Scholarship Program, funding their vouchers through high school graduation, but denying entry to additional children. Both program supporters and opponents cite evidence from an ongoing congressionally mandated Institute of Education Sciences (IES) evaluation of the program, for which I am principal investigator,¹ to buttress their positions, rendering the evaluation a Rorschach test for one's ideological position on this fiercely debated issue.

Opportunity Scholarships

Opportunity Scholarships are a formal name for school vouchers. Vouchers provide funds to parents to enable them to enroll their children in private schools and, as a result, are one of the most controversial education reforms in the United States. Currently, 13 directly funded voucher programs operate in four U.S. cities and six states, serving approximately 65,000 students.² Another seven programs indirectly fund private K–12 scholarship organizations through government tax credits to individuals or corporations. About 100,000 students receive school vouchers funded through tax credits.³ All of the directly funded voucher programs are targeted to students with some educational disadvantage, such as low family income, disability, or status as a foster child.

Among the many points of contention is whether voucher programs improve student achievement. Most evaluations of such programs have found at least some positive achievement effects, but not always for all types of participants and not always in both reading and math.⁴ This pattern of results has so far failed to generate a scholarly consensus regarding the beneficial

¹ My coauthors include Babette Gutmann and Lou Rizzo of Westat, Michael Puma of Chesapeake Research Associates, Brian Kisida of the University of Arkansas, and Nada Eissa of Georgetown University. Although our government report was a collaborative product of this talented research team, my interpretation of the report's results provided here are my own.

² The programs operate in the states of Arizona, Florida, Georgia, Maine, Ohio, Utah, and Vermont as well as the cities of Cleveland, the District of Columbia, Milwaukee, and New Orleans. See <http://www.friedmanfoundation.org/schoolchoice/ShowProgram.do>

³ The tax-credit scholarship programs operate in the states of Arizona, Florida, Georgia, Iowa, Indiana, Pennsylvania, Rhode Island. See <http://www.friedmanfoundation.org/schoolchoice/ShowProgram.do>

⁴ Patrick J. Wolf, "School Voucher Programs: What the Research Says About Parental School Choice," *Brigham Young University Law Review*, 2008:2, pp. 415-446, <http://lawreview.byu.edu/archives/2008/2/90WOLF.FIN.pdf>

effects of school vouchers on student achievement. The policy and academic communities seek more definitive guidance.

Nineteen of the 20 school voucher programs in the U.S. are funded by state and local governments. The OSP is the only federal voucher initiative. Established in 2004 as part of compromise legislation that also included new spending on charter and traditional public schools in the District of Columbia, the OSP is a means-tested program. Initial eligibility is limited to K–12 students in D.C. with family incomes at or below 185 percent of the poverty line. Congress has appropriated \$14 million annually to the program, enough to support about 1,700 students at the maximum voucher amount of \$7,500. The voucher covers most or all of the costs of tuition, transportation, and educational fees at any of the 66 D.C. private schools that have participated in the program. By the spring of 2008, a total of 5,331 eligible students had applied for the limited number of Opportunity Scholarships. Recipients are selected by lottery, with priority given to students applying to the program from public schools in need of improvement (SINI) under No Child Left Behind.

The Institute of Education Sciences (IES) in the U.S. Department of Education released the third-year impact evaluation of the Opportunity Scholarship Program (OSP) in March 2009. The results showed students who participated in the program performed at significantly higher levels in reading than the students in an experimental control group. Here are the study findings and my own interpretation of what they mean.

Study Background

Our evaluation of the OSP uses the most rigorous research method available for determining the impact of this school choice program. Parents who seek schooling options for their children are likely to be highly motivated to promote their children's educational success. That high level of parental motivation that leads parents to participate in school choice programs probably also contributes to greater student achievement over time, leading to what we call "self-selection bias" in the research world.

To ensure that parent motivation does not bias studies of school choice programs, researchers over the past decade have focused on evaluating them using experimental research designs called Randomized Control Trials (RCTs). With an RCT design, a group of students that all qualify for a voucher or scholarship program and whose parents are equally motivated to exercise school choice are subject to a scholarship lottery. The students who win the lottery become the experimental "treatment" group. The students who lose the lottery become the experimental control group. Since only a school voucher and mere chance distinguish the treatment students from their control counterparts, any subsequent difference in student outcomes for the treatment students can be reliably attributed to the voucher intervention. That is, the outcomes from the control group represent what would have happened to the treatment group absent the program, and the treatment impact is therefore the treatment outcomes minus the control outcomes. Because of the rigor of experimental designs they are often dubbed the "gold standard" for policy evaluations and are widely used to evaluate the efficacy of medical drugs and procedures prior to such treatments being made available to the public.

Student and School Participation

Two cohorts of students were followed for purposes of this evaluation. All of the students were attending public schools or rising Kindergartners at the time of application. Cohort 1 consisted of 492 students entering grades 6-12 in 2004. Cohort 2 consisted of 1,816 students entering grades K-12 in 2005. The characteristics and outcomes of these two groups, combined into an impact sample of 2,308 students by lagging the Cohort 1 outcomes by one year, have been the focus of our impact evaluation. At total of 1,387 students in the impact sample won the scholarship lottery and were thereby assigned to the treatment group, while the remaining 921 students who did not win the lottery were thereby assigned to the control group. Over the five years of program operation, other students have received scholarships without having to go through the lottery. These students were not included in the rigorous impact evaluation because no appropriate comparison group is available for them.

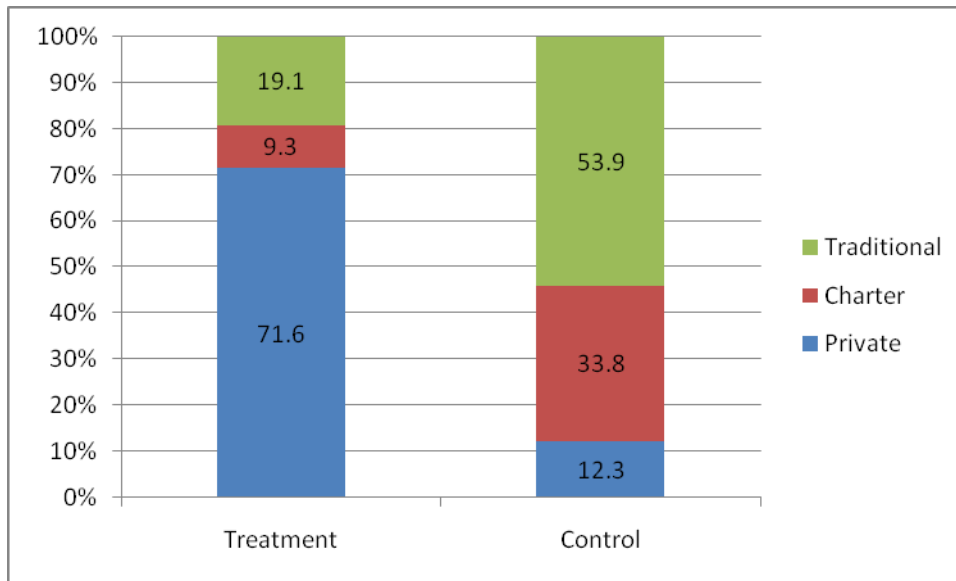
Evidence from the study confirms that the OSP serves a highly disadvantaged group of DC students. Descriptive information from the first two annual reports about program participation indicates that over 90 percent of students are African American and nine percent are Hispanic. Their family incomes averaged less than \$20,000 in the baseline year in which they applied for the program. Overall, participating students were performing well below national norms in reading and math when they applied to the program. Forty-four percent of students in both cohorts were attending a public school designated as “in need of improvement” (SINI) between 2003 and 2005.

The Opportunity Scholarship Program is designed to facilitate the enrollment of low-income District students in private schools of their parents’ choosing. It does not and cannot guarantee enrollment in a private school, but the \$7,500 voucher should make such enrollments relatively common among the students who won the scholarship lottery. The eligible students who lost the scholarship lottery and therefore were assigned to the control group still might attend a private school but they would have to do so by drawing upon resources outside of the OSP. At the same time, students in both the scholarship treatment group and the control group have access to a large number of public charter schools in the District.

The implications of these realities is that, for this evaluation of the OSP, assignment to the treatment group does not necessarily mean private schooling and assignment to the control group does not necessarily mean education in a traditional public school. Members of both the treatment and control groups attended all three types of schools – private, public charter, and traditional public – in year 3 of the voucher experiment, though the proportions that attended each type differed significantly based on whether or not they won the scholarship lottery (figure 1). Almost 72 percent of the students who won the voucher lottery and provided outcome data in year 3 were attending private schools. Only 12 percent of the students who lost the voucher lottery were enrolled in private schools in year 3. Over nine percent of the treatment students chose to attend a public charter school three years after receiving a scholarship offer, compared to almost 34 percent of the control group who opted for that public school choice option. About 19 percent of the treatment group students were enrolled in traditional public schools three years after the scholarship lottery, compared with nearly 54 percent of control group students in such schools.

I see these data as underscoring that the desire for an alternative to a neighborhood public school was strong for the families who applied to the OSP in 2004 and 2005. About 81 percent of them placed their child in a private or public school of choice three years after winning the scholarship lottery and 46 percent of them did likewise even if they lost the lottery. This was a group of families with a strong motivation to exercise parental school choice.

Figure 1. Types of Schools Attended by the Treatment and Control Groups in Year 3



Source: Wolf et al., *The Evaluation of the DC Opportunity Scholarship Program: Impacts After Three Years* (NCEE 2009-4050), Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2009, Table 2-7.

The enrollment pattern of students in the evaluation also highlights the fact that the comparison of the treatment and control groups in year 3 does not amount to a comparison between “all choice” and “no choice.” Instead, it is a comparison of outcomes between a group exercising lots of private school choice and some public school choice with a group exercising a small amount of private school choice and a substantial amount of public school choice. Any differences between the outcomes of the treatment and control groups therefore indicate the incremental impact of adding private school choice through the OSP to the existing schooling options for low-income DC families.

If one’s purpose is to evaluate the effects of a specific public policy, such as the OSP, then the comparison of the average outcomes of the treatment and control groups, regardless of what proportion attended which types of school, is most appropriate. A school voucher program cannot force scholarship recipients to use a voucher, nor can it preclude control group students from attending private schools at their own expense. A voucher program only can offer students scholarships that they subsequently may or may not use. Nevertheless, the mere offer of a scholarship, in and of itself, clearly has no impact on the educational outcomes of students. A scholarship could only change the future of a student if it were actually used.

Fortunately, two statistical techniques are available that draw upon the unbiased results of the pure experimental analysis of treatment and control group differences. In the opinion of many researchers, including myself, these methodological approaches produce reliable estimates of the average effect of using a voucher compared to not being offered one and the average effect of attending private school with or without a voucher compared to not attending private school. The technique that produces the estimate of the effect of using a voucher is called a Bloom adjustment. Since lottery winners who never used a scholarship could not have been affected by it, the average impact of the voucher program on student outcomes that was generated by the entire sample of treatment students – users and non-users alike – is simply re-scaled by dividing it by the percentage of the treatment group that actually availed themselves of the treatment. For example, if 80 percent of the treatment students used their scholarships at any time since the voucher lottery and the treatment group as a whole averaged test score outcomes that were 4 points higher than the control group, the Bloom-adjusted estimated effect of using a scholarship on test scores would be $4/.8$ or 5 points.

The method for estimating the effect of attending versus not attending private schools, called Instrumental Variable (IV) analysis, produces estimates that tend to be larger than Bloom-adjusted estimates because they adjust for both non-use of the scholarship by the treatment group and private school attendance by members of the control group. As such, an IV analysis of the effect of private schooling is not an evaluation of a school voucher program *per se* but, instead, is an evaluation of the effect of the condition (private school enrollment) that a voucher program seeks to facilitate. Because such analyses place heavy demands on the underlying data, smaller differences that are found to be statistically significant at the purely experimental stage can end up as larger differences that are not statistically significant when estimated through IV analysis. All three effect estimates – purely experimental, effect of use, and effect of private schooling – are provided in the remainder of this testimony so that individual readers can decide which outcomes are most relevant to their considerations.

The Opportunity Scholarship Program and Student Achievement

Our analysis of the data after three years of participation in the OSP revealed that the program had a statistically significant positive impact on the test scores of students in reading (table 1). The positive impact of the voucher program on student reading scores after three years amounted to an average gain of 4.5 scale score points across the entire treatment group, 5.3 scale score points for scholarship users in the treatment group, and 7.1 scale score points for attending private versus public school by the third year of the evaluation. These results are statistically significant at the 95 percent confidence level. We know from this study that participating DC students are reading at higher levels as a result of the Opportunity Scholarship Program.

No statistically significant impacts were observed in math and therefore no estimate of the effect of private schooling on math achievement was attempted by the evaluation team (table 1).

Table 1. Year 3 Overall Achievement Impact Estimates of the Scholarship Offer, Use of a Scholarship, and Private Schooling

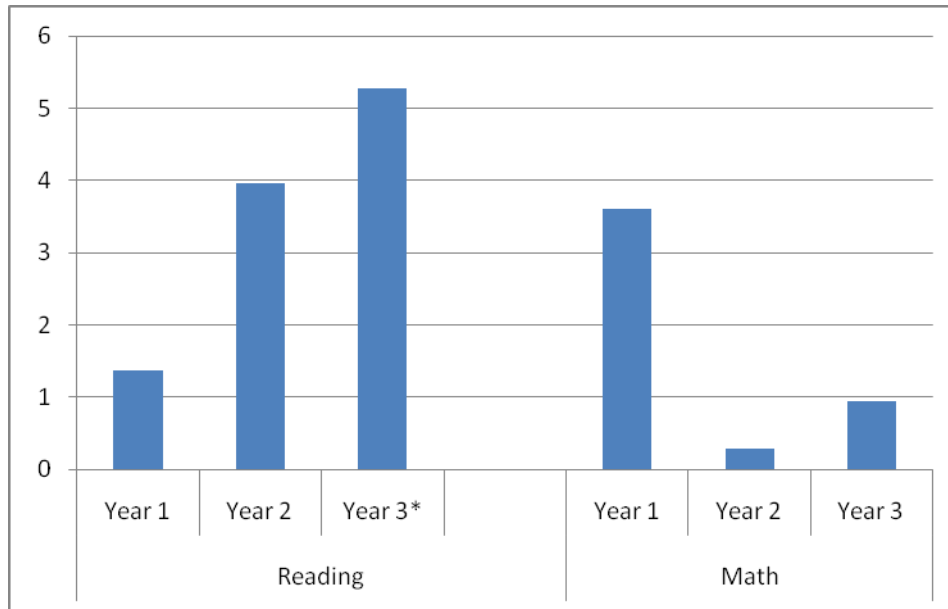
Student Achievement	Impact of the Scholarship Offer			Impact of Scholarship Use	Effect of Private Schooling
	Treatment Group Mean	Control Group Mean	Difference (Estimated Impact)	Adjusted Impact Estimate	IV Estimate
Reading	635.4	631.0	4.5*	5.3*	7.1*
Math	630.2	629.4	.8	1.0	NA

*Statistically significant at the 95 percent confidence level.

SOURCE: Wolf et al., *Evaluation of the DC Opportunity Scholarship Program: Impacts After Three Years...*, Tables 3-2 and E-1.

Examined over time, the experimental impacts of using an Opportunity Scholarship for any length of time appear to show a trend towards larger reading gains cumulating for students in the program (figure 2). Especially when one considers that students who used their scholarship in year 1 needed to adjust to a new and different school environment, the reading impacts of using a scholarship of 1.4 scale score point (not significant) in the first year, 4.0 scale score points (not significant) in the second year, and 5.3 scale score points (significant) in the third year suggest that students are consistently gaining in reading performance relative to their control group peers the longer they experience the OSP. No such trend is apparent regarding math achievement.

Figure 2. Experimental Achievement Impacts of Scholarship Use in All Three Years of Evaluation



*Statistically significant at the 95 percent confidence level.

SOURCE: Calculated by applying the reported scholarship usage rates to the experimental impacts of the scholarship offer reported in Wolf et al., *Evaluation of the DC Opportunity Scholarship Program: Impacts After Three Years...*, Figure 3-3.

In sum, the OSP generated clear achievement gains in reading after three years but no discernible impacts in math.

Were There Any Subgroup Differences?

Several commentators on the DC OSP evaluation have made strong claims that certain subgroups of participants experienced benefits from the program treatment that were significantly different than those of other subgroups of participants. Martin Carnoy states that “the treated students in Cohort 1 were concentrated in middle schools and the effect on their reading score was significantly higher than for treated students in Cohort 2.”⁵ Henry Levin asserts that “the evaluators found that receiving a voucher resulted in no advantage in math or reading test scores for either [low achievers or students from SINI schools].”⁶

The actual results of the evaluation provide no scientific basis for such claims that students with certain characteristics experienced benefits from the voucher treatment that differed from students without those characteristics. The impact of the program on the reading achievement of Cohort 1 students was not significantly different than the impact of the program on the reading achievement of Cohort 2 students, Carnoy’s claim notwithstanding. Low achievers and applicants from SINI schools did not experience significantly different reading gains from the program than high achievers and non-SINI applicants, contrary to Levin’s insinuation that they had. In statistical parlance, the interaction between the treatment impact and being a member of any of the subgroups examined was never, itself, statistically significant. These commentators took a set of subgroup results in which some (but not all) treatment subgroups demonstrated statistically significant reading achievement *differences from their comparable control subgroups* and mischaracterized it as proving that some subgroups of treatment students benefited academically from the program but others did not.

A hypothetical example from the field of public opinion will help make this subtle point clearer. Suppose a national poll found that the Democrat candidate for President was preferred by 54 percent of likely voters but the Republican candidate was preferred by only 46 percent of them. With a large enough sample of, say 1,000 respondents, that difference of 8 percentage points would be outside of the margin of error of plus or minus 3 percentage points and the pollsters would say that the nation as a whole prefers the Democrat candidate to the Republican candidate. Suppose that results were also presented by gender, and the Democrat held a 57-43 edge among women and a 52-48 advantage among men. Such a subgroup analysis of poll results would produce a higher margin of error, say plus or minus 7 percentile points, meaning that the advantage of the Democrat among women is statistically significant at the subgroup level whereas his advantage among men represents a statistical dead heat. The Democrat's 14 point advantage among women, itself, would not likely be significantly different, in the statistical sense, from his 4 point advantage among men.

⁵ Martin Carnoy, "Review of *Evaluation of the DC Opportunity Scholarship Program: Impacts After Three Years*," Boulder and Tempe: Think Tank Review Project, 2009, p. 4.

⁶ Henry Levin, "Can Educational Vouchers Solve the Problem of Low Achievement in Washington DC? Results of a Three Year Experiment," Press release of the National Center for the Study of Privatization in Education, Teachers College, New York, NY, received April 17, 2009.

The correct interpretation of these polling results would be that the nation as a whole prefers the Democrat candidate, that the preference for the Democrat is not necessarily higher among women than it is among men, but that the preference is sufficiently large among women alone that it is statistically significant even at the subgroup level. Carnoy and Levin, in contrast, would say incorrectly that the Democrat's advantage over his rival is significantly larger among women than among men and that women are the sole source of the Democrat's overall edge in popularity -- claims that would undoubtedly cheer the beleaguered Republican candidate but would have no scientific merit.

From a scientific standpoint, three conclusions are valid about the achievement results in reading from the year 3 impact evaluation of the OSP:

1. That it improved the reading achievement of the treatment group students overall;
2. That overall reading gains from the program were not significantly different across the various subgroups examined;
3. That certain subgroups of students did demonstrate statistically significant reading impacts from the program even at the smaller subgroup level of analysis.

Any other interpretations of those results, such as that some types of students benefited more in reading from the voucher program than other types of students, or that some student subgroups didn't benefit at all from the program, are mere speculation wholly divorced from the scientific evidence in the rigorous evaluation.

Why examine distinct achievement impacts at the subgroup level, if the evidence indicates only an overall reading gain for the entire sample? The reasons are that Congress mandated an analysis of subgroup impacts, at least for SINI and non-SINI students, and because analyses at the subgroup level can yield more specific information about the size of achievement gains for certain types of students, so long as those specific results are interpreted properly.

Such subgroup analyses require that evaluators carve up the overall study sample into smaller constituent parts. As a result, less data inform each impact estimate, making them less precise and therefore less likely to identify statistically significant impacts. In addition, evaluating the impact of a program on different sub-groupings of participants requires multiple significance tests, any one of which, at the 95 percent confidence level, has about a 5 percent chance of being a false discovery. With each subgroup impact examined, the risk of at least one false discovery increases somewhat. The impact evaluation conducted statistical tests to determine which subgroup findings could be false discoveries. As a result, three types of findings emerge from the subgroup analysis in the evaluation: findings that are clearly statistically significant, findings that are statistically significant with reservations (because "false discovery" could not be ruled out), and findings that are not statistically significant.

When examined as separate subgroups, three types of students clearly experienced significant reading impacts as a result of three years in the OSP (table 2). Public school students who were not attending schools in need of improvement prior to entering the program gained an average of 6.6 scale score points in reading if in the treatment group, 7.7 scale score points from

using a scholarship, and 10.3 scale score points from private schooling. Students in the higher two-thirds of the performance distribution, whose average reading test score was at the 37th National Percentile Rank at baseline, gained 5.5, 6.2, and 9.5 scale score points in reading achievement from the scholarship offer, scholarship use, and private schooling respectively. Students entering K-8 at baseline, where slots were plentiful in a wide variety of participating private schools, gained 5.2, 6.0, and 8.3 scale score points in reading from the scholarship offer, scholarship use, and private schooling after three years. These impact estimates were statistically significant and remained so after adjustments for multiple comparisons.

Two other individual subgroups of students demonstrated reading impacts from the program that were not as robust. Female students gained an average of 5.1 scale score points in reading from the scholarship offer, 5.9 scale score points from using a scholarship, and 6.1 scale score points from private schooling. Students in Cohort 1 – the eager “first movers” into the program – gained 8.7, 11.7, and 15.8 scale score points in reading from the scholarship offer, scholarship use, and private schooling respectively. However, the estimation of the private schooling effect was not statistically significant and statistical tests indicated that the impacts of the scholarship offer and scholarship use could have been false discoveries for both of these subgroups.

Reading impacts for the other five subgroups examined individually – applicants from schools in need of improvement (i.e. SINI), students in the lower one-third of the performance distribution at baseline, males, students entering high school grades at baseline, and students in Cohort 2 – were not statistically significant after three years. This does not mean that those subgroups of students did not benefit from the program, as research results never prove a negative, but it does mean that reading gains were not clearly evident at the subgroup level for those types of students. The fact that significant reading impacts were not observed for the subgroup of SINI students is noteworthy, since Congress designated SINI students as the highest service priority for the program. Math impacts were not statistically significant for any of the 10 subgroups examined.

Table 2. Year 3 Subgroup Achievement Impact Estimates of the Scholarship Offer, Use of a Scholarship, and Private Schooling

Student Achievement	Impact of the Scholarship Offer			Impact of Scholarship Use	Effect of Private Schooling
	Treatment Group Mean	Control Group Mean	Difference (Estimated Impact)	Adjusted Impact Estimate	IV Estimate
SINI never	625.3	618.7	6.6**	7.7**	10.3*
Higher performance	644.7	639.3	5.5*	6.2*	9.5*
<i>Female</i>	<i>639.3</i>	<i>634.2</i>	<i>5.1*</i>	<i>5.8*</i>	<i>6.1</i>
K-8	627.3	622.1	5.2**	6.0**	8.3*
<i>Cohort 1</i>	<i>672.9</i>	<i>664.2</i>	<i>8.7*</i>	<i>11.7*</i>	<i>15.8</i>

*Statistically significant at the 95 percent confidence level. Subgroup results in italics were not statistically significant after adjustments for multiple comparisons.

SOURCE: Wolf et al., *Evaluation of the DC Opportunity Scholarship Program: Impacts After Three Years...*, Tables 3-3 and E-1.

Overall Impacts on Parent and Student Satisfaction

Whenever school choice researchers have asked about satisfaction with schools, parents who were given the chance to select their child's school have reported much higher levels of satisfaction. Students themselves, for any number of possible reasons, have rarely described themselves as more satisfied with the new schools chosen by their parents. The year 3 satisfaction results from the OSP evaluation fit this pattern of previous studies. The proportion of parents who assigned a high grade of A or B to their child's school was 11 percentile points higher if they were in the treatment group, 12 percentile points higher if their child actually used a scholarship, and 21 percentiles higher if they were attending private school in year 3 of the study. Parents also were significantly more confident of the safety of their children in school if they had been awarded an Opportunity Scholarship. Students in grades 4-12, when asked similar questions, were no more likely to be satisfied with their school or describe it as safe if they were in the treatment compared to the control group.

Additional evidence of parental satisfaction with the OSP comes from the series of focus groups conducted independently of the congressionally mandated evaluation. One parent was satisfied with the expanded freedom inherent in school choice:

"[The OSP] gives me the choice to, freedom to attend other schools than D.C. public schools....I just didn't feel that I wanted to put him in DC public school and I had the opportunity to take one of the scholarships, so, therefore, I can afford it and I'm glad that I did do that." (Cohort 1 Elementary School Parent, Spring 2008)⁷

Another parent with two children in the OSP may have hinted at a reason achievement impacts were observed specifically in reading:

"They really excel at this program, 'cause I know for a fact they would never have received this kind of education at a public school....I listen to them when they talk, and what they are saying, and they articulate better than I do, and I know it's because of the school, and I like that about them, and I'm proud of them." (Cohort 1 Elementary School Parent, Spring 2008)⁸

These parents of OSP students clearly see their families as having benefited from this program.

Interpreting the Findings

What does this pattern of results suggest about the effectiveness of the OSP? Any answer to that question is bound to be somewhat subjective, so I think the best way to begin is to compare the achievement impacts from the OSP with those from randomized control trials of other education programs.

⁷ Thomas Stewart, Patrick Wolf, Stephen Q. Cornman, Kenann McKenzie-Thompson, and Jonathan Butcher, *Family Reflections on the District of Columbia Opportunity Scholarship Program*, Report of the School Choice Demonstration Project, University of Arkansas, Fayetteville, AR, January 2009, p. 32, available at http://www.uaedreform.org/SCDP/DC_Research/2009_Final.pdf

⁸ Ibid., p. 33.

The National Center for Educational Evaluation (NCEE) at the Institute of Education Sciences has released the results of 11 studies that, like this one, employ the methodological rigor of random assignment to treatment and control groups. The DC OSP evaluation is one of only three of these 11 NCEE studies to report overall statistically significant positive achievement impacts in either reading or math (table 3). The other two discreet federal education programs which have been confirmed to deliver overall achievement impacts are Enhanced Reading Opportunities and After-School Programs and Enhanced Academic Instruction. The size of the reading gain from Enhanced Reading Opportunities is 40 percent smaller than, and the math gain from After School Programs is less than half of, the reading gain from the DC OSP, which has shown the largest statistically significant impact of any NCEE experimental study.

Six other education programs -- covering approaches such as student mentoring, Reading First, classroom interventions in Even Start, alternative teacher certification, initial teacher training, and professional development -- have not demonstrated statistically significant achievement impacts overall. Two programs demonstrated a mix of non-significant and negative impacts on achievement. Several of these education programs have only been evaluated for one or two years, and could show significant achievement impacts in subsequent reports. The larger point is that many federal education programs targeted at disadvantaged students are now the subjects of rigorous evaluations. Most of these programs have yet to demonstrate the ability to move disadvantaged students to significantly higher levels of academic achievement. In my opinion, by demonstrating statistically significant impacts overall in reading based on an experimental evaluation, the DC OSP has met a tough standard for efficacy in serving low-income inner-city students.

At the same time, the fact that students who had been attending public schools in need of improvement (SINI), as a distinct subgroup, have yet to show significant gains from the program should not be ignored. The SINI students were designated by Congress as a service priority for the program. The data suggest that, as a subgroup, SINI students thus far are no better or worse off academically if they were offered a scholarship. It also is important to repeat that the statistically significant OSP gains, overall and for half of the subgroups, thus far have been limited to reading. Although significant gains were observed for two subgroups in math after one year, statistical tests suggested they might be false discoveries and no significant math impacts have been detected since.

Table 3. NCEE Intervention Studies in Order of Significance of Achievement Impacts, May 2009

NCEE Single Intervention Study	Overall Significant Impact	Partial or Subgroup Sig. Impact
1 DC Opportunity Scholarship Program	positive (reading), no effect (math) (3 years) Effect size = 0.13 (reading)	some positive subgroups, some no effect
2 Enhanced Reading Opportunities	positive (1 year) Effect size = 0.08 (reading)	some positive subgroups, some no effect
3 After-School Programs and Enhanced Academic Instruction	positive (math), no effect (reading) (1 year) Effect size = 0.06 (math)	n/a
4 Student Mentoring Program	no effect	some positive subgroups, some no effect
5 Reading First	no effect (3 years)	improvements in student decoding skills
6 Classroom Literacy Interventions and Outcomes in Even Start	no effect (literacy measures)	improvements in parenting skills and children's social skills
7 Teacher Certification Routes	no effect	n/a
8 Comprehensive Elementary Teacher Induction	no effect (student achievement, teachers' practices, or teacher retention rates)	n/a
9 Professional Development Interventions for Early Reading	no effect (test scores)	no subgroup effects
10 Impact of Selective Supplemental Curricula on Reading Comprehension	3 no effect, 1 negative (1 year)	Some no subgroup effects, some negative
11 Closing the Gap Impacts on Reading (Title I)	2 no effect (3rd reading & math), 2 negative (5th reading & math) (1 year)	some positive subgroups, some no effect
Totals:	1 positive, 2 some positive/ some no effect, 6 no effect, 2 some no effect/some neg.	

SOURCE: Calculated from review of the most recent evaluation reports where interventions were compared to a control group (see <http://ies.ed.gov/ncee/pubs/>). Evaluations that merely compared interventions to each other are excluded.

How large are the statistically significant reading gains observed in the OSP overall and for half of the subgroups after three years? The magnitude of the gains may lie in the eyes of the beholder. One constructive way to view achievement gains, however, is in terms of additional months of instruction. The overall gains from the OSP observed after three years mean that members of the control group, who represent what scholarship students would have experienced absent the program, would need to remain in school an extra 3.7 months on average to catch up to the level of reading achievement obtained by scholarship users (table 4). When the IV

procedure is used to adjust for control group students attending private schools, we see that private schooling added nearly 5 months of achievement to the reading skills of students over the three years of the study.

Table 4. Estimated Impacts in Months of Schooling of the Scholarship Offer, Use of a Scholarship, and Private Schooling for Statistically Significant Reading Impacts After 3 Years

Student Achievement: Reading	Months of Schooling		
	Impact of the Scholarship Offer	Impact of Scholarship Use	Effect of Private Schooling
Full sample	3.1	3.7	5.0
SINI never	4.1	4.9	6.5
Higher performance	4.0	4.6	7.0
<i>Female</i>	<i>3.1</i>	<i>3.6</i>	<i>3.7</i>
K-8	2.9	3.3	4.6
<i>Cohort 1</i>	<i>14.1</i>	<i>18.9</i>	<i>25.5</i>

Subgroup results in italics were not statistically significant after adjustments for multiple comparisons.

SOURCE: Wolf et al., *Evaluation of the DC Opportunity Scholarship Program: Impacts After Three Years...*, Table 3-4.

At the subgroup level, private schooling increased the reading achievement of Cohort 1 students by over 25 months, though that result for this relatively small subgroup of students is not very robust and should be interpreted with caution. Female students gained 3.7 months (also not very robust), K-8 students 4.6 months, applicants from non-SINI schools 6.5 months, and higher baseline performers 7 months of reading, respectively, due to the private schooling opportunities made possible by the OSP.

If these trends were to continue over the entire educational experience of a typical treatment student who entered the program in Kindergarten, my calculation is that the student would be reading two-and-a-half years ahead of her peers in the control group who did not receive an Opportunity Scholarship by the time she graduated from high school. The reading gains from private schooling demonstrated by OSP students after three years are equivalent to about one-quarter of the notorious Black-White achievement gap. Although it is mere speculation at this point whether the impacts we have observed will continue at these levels, over 13 years of K-12 education, similar results would eliminate the racial gap in reading performance entirely.

Future Research

The current rigorous evaluation of the DC Opportunity Scholarship Program makes important contributions to our understanding of the effects of voucher programs and of private school attendance. However, there is much more that could be learned from the OSP – either through new data collection or even new analyses of what we have already obtained. The most important questions that remain to be explored include:

1. *What are the impacts of the program after four or five years of participation?*

The research team is well along in the process of collecting follow-up data from Cohort 1 after five years and Cohort 2 after four years since students were randomly assigned to the treatment and control groups. Analysis of those data will indicate if the achievement impacts observed for the OSP students after three years grow or fade, and will be an important topic of our final evaluation report, planned for release next spring.

2. *Does the OSP improve high school graduation and college enrollment rates?*

Unlike many other scholarship programs, the OSP enrolled older students (beyond grade 6). Although in our final report next year we will estimate the impacts of the program on educational attainment, only a relatively small proportion of students are old enough to be included in that analysis. In the coming years, there will be a substantial group of OSP students of high school age or older. Recent studies of charter schools suggest that their biggest impact may be on educational attainment.⁹ It is important to know whether or not voucher programs have the same effects.

3. *How do participating private schools differ from the public school students would have attended?*

The current evaluation is not the first to find academic benefits for students who use scholarships, but none have been able to adequately explore – much less determine empirically – why or how these impacts happen. There are many hypotheses: better instruction, better peer group of students, higher expectations set, more discipline, a smaller more nurturing school community, greater parent involvement. We surveyed parents, students, and public and private school principals about these school characteristics, but because of resource and time constraints, have only begun to tap the surface of examining the environments and organization of the two types of schools. Such an investigation would likely require site visits, classroom observations, and surveys of teachers of students in the impact sample, which all have been beyond the scope of the current federal evaluation.

4. *Who participates in the OSP and who drops out?*

We have learned that, by year 3, almost half of the students who received an OSP scholarship have graded out (graduated from high school), earned out (change in family income makes them ineligible), moved out of the District, or left the program for other reasons. It would be possible to use the current evaluation data to explore what types of students initially applied to the program, how and why students move in and out of scholarship use and private school enrollment, and how these patterns relate to program impacts. This information might help organizations that run voucher or private school scholarship programs identify students who might need additional programmatic supports and what types of supports might be helpful.

⁹ Ron Zimmer, Brian Gill, Kevin Booker, Stephane Lavertu, Tim R. Sass, John Witte, *Charter Schools in Eight States: Effects on Achievement, Attainment, Integration, and Competition*, Santa Monica, CA: The RAND Corporation, 2009, available at http://www.rand.org/pubs/monographs/2009/RAND_MG869.pdf

5. *Does the OSP have any effect on racial integration?*

Many people are concerned that school choice programs may affect the racial diversity of schools. It might be possible, with the current evaluation data, to estimate the impacts of the program on the racial composition of District of Columbia public and private schools. We could address two important policy questions surrounding scholarship programs. First, do the students who participate in such programs end up enrolled in schools with greater racial diversity than they otherwise would have experienced? Second, are the schools that scholarship participants leave as a result of the program better integrated racially as a consequence. The combination of school-reported and individual level data that we have collected provides a unique opportunity to examine these important questions.

Conclusion

For the past five years, the District of Columbia Opportunity Scholarship Program has provided income-disadvantaged students with government-financed scholarships or vouchers to facilitate their enrollment in participating private schools selected by their parents. Having collected and analyzed data from the first three years of student experiences with the OSP, we have learned much about the program. Students overall are reading at higher levels as a result of the OSP. No impacts have been observed in Math achievement. When the data are parsed into smaller subgroups, half of those individual subgroups of students are demonstrating reading gains as a result of the program. The SINI students, who are a service priority of the program, and four other student subgroups have not shown significant achievement impacts as an individual subgroup to date. Parents, but not students, say that they are more satisfied with their schools if offered an Opportunity Scholarship and they view those schools as safer.

Through a rigorous evaluation, much knowledge has been gained about the nation's first federally-funded school voucher program. Nevertheless, questions about the program's future remain. Will the Opportunity Scholarship Program be reauthorized and new students, including the siblings of existing students, permitted to join this means-tested school voucher initiative? Will future funding be restricted to existing students, causing the program to wither on the vine as those students graduate? Will the program be ended entirely, forcing nearly 1,700 low-income DC students from their private schools of choice? What will policymakers see as they continue to gaze at the results of this government-sponsored independent evaluation? The educational futures of a group of low-income D.C. schoolchildren hinge on the answer.

Patrick J. Wolf is professor of education reform at the University of Arkansas and principal investigator of the D.C. Opportunity Scholarship Program Impact Evaluation. The opinions expressed in this article are his own.