EVALUATION OF THE DC OPPORTUNITY SCHOLARSHIP PROGRAM: SUMMARY OF EXPERIMENTAL IMPACTS AFTER THREE YEARS

Patrick Wolf
University of Arkansas

Babette Gutmann,
Westat

Michael Puma
Chesapeake Research Associates

Brian Kisida
University of Arkansas

Lou Rizzo
Westat

Nada Eissa
Georgetown University

EVALUATION OF THE DC OPPORTUNITY SCHOLARSHIP PROGRAM: SUMMARY OF EXPERIMENTAL IMPACTS AFTER THREE YEARS [1]

The *District of Columbia School Choice Incentive Act of 2003*,[2] passed by the Congress in January 2004, established the first federally funded, private school voucher program in the United States. The purpose of the new scholarship program is to provide low-income parents, particularly those whose children attend schools identified for improvement or corrective action under the *Elementary and Secondary Education Act*, with "expanded opportunities to attend higher performing schools in the District of Columbia" (Sec. 303). According to the statute, the key components of the Program include:

- To be eligible, students entering grades K-12 must reside in the District and have a family income at or below 185 percent of the federal poverty line.

- Participating students receive scholarships of up to $7,500 to cover the costs of tuition, school fees, and transportation to a participating private school.

- Scholarships are renewable for up to 5 years (as funds are appropriated), so long as students remain eligible for the Program and remain in good academic standing at the private school they are attending.

- In a given year, if there are more eligible applicants than available scholarships or open slots in private schools, applicants are to be awarded scholarships by random selection (e.g., by lottery).

- In making scholarship awards, priority is given to students attending public schools designated as in need of improvement (SINI) under the *No Child Left Behind (NCLB) Act* and to families that lack the resources to take advantage of school choice options, operationally defined as students attending public school at the time of application.

- Private schools participating in the Program must be located in the District of Columbia and must agree to requirements regarding nondiscrimination in admissions, fiscal accountability, and cooperation with the evaluation.

---

[2] Title III of Division C of the *Consolidated Appropriations Act*, 2004, P.L. 108-199.

As part of this legislation, the Congress mandated a rigorous evaluation of the impacts of the Program, now called the DC Opportunity Scholarship Program (OSP). This paper presents findings from the evaluation on the impacts 3 years after families who applied were given the option to move from a public school to a participating private school of their choice.

The evaluation is based on a randomized controlled trial design that compares the outcomes of eligible applicants randomly assigned to receive (treatment group) or not receive (control group) a scholarship through a series of lotteries. The main findings of the evaluation so far include:

- **Four years after the initial Program launch and 3 years after full implementation, a total of 8,578 students had applied to the OSP.** A total of 5,769 applicants were confirmed to be eligible for the Program; 4,176 were offered scholarships. A total of 2,881 students used their scholarship to attend a participating private school in the fall immediately after receiving it, and 1,714 of them were using scholarships in the fall of 2008 (table 1).

- **After 3 years, the evaluation found that the OSP improved reading achievement by 4.5 scale score points—equivalent to 3.1 months of additional learning** (tables 5 and 7).

- **There was no statistically significant difference in math test scores between students who were offered an OSP scholarship and students who were not offered a scholarship.** Overall, those in the treatment and control groups were performing at comparable levels in mathematics (table 5).

- **The Program had a positive impact on overall parent satisfaction and parent perceptions of school safety, but not on students' reports of satisfaction and safety** (tables 8-11). Parents were more satisfied with their child's school and viewed the school as less dangerous if the child was offered a scholarship. Students had a different view of their schools from that of their parents. Student reports of dangerous incidents in school were comparable for students in the treatment and control groups. Overall, student satisfaction was unaffected by the Program.

- **This same pattern of findings holds when the analysis is conducted to determine the impact of *using* a scholarship rather than being *offered* a scholarship.** Fourteen percent of students in our impact sample who were randomly assigned by lottery to receive scholarship and who responded to year 3 data collection chose not to use their scholarship at any point over the 3-year period after applying to the Program. We use a common statistical technique to take those "never users" into account; it assumes that the students had zero impact from the OSP, but it does not change the statistical significance of the original impact estimates. Therefore, the positive impacts on reading achievement, parent views of school safety and climate, and parent views of satisfaction all increase in size, and there remains no impact on math achievement and no impact on students' perceptions of school safety or satisfaction from using an OSP scholarship.

3

- **The OSP improved reading achievement for 5 of the 10 subgroups examined.[3]**
  Being offered or using a scholarship led to higher reading test scores for participants who applied from schools that were not classified as "schools in need of improvement" (non-SINI). There were also positive impacts for students who applied to the Program with relatively higher levels of academic performance, female students, students entering grades K-8 at the time of application, and students from the first cohort of applicants. These impacts translate into 1/3 to 2 years of additional learning growth. However, the positive subgroup reading impacts for female students and the first cohort of applicants should be interpreted with caution, as reliability tests suggest that they could be false discoveries.

- **No achievement impacts were observed for five other subgroups of students.**
  Subgroups of students who applied from SINI schools (designated by Congress as the highest priority group for the Program) or were in the lower third of the test score distribution among applicants did not demonstrate significant impacts on reading test scores if they were offered a scholarship. In addition, male students, those entering high school grades upon application, and those in application cohort 2 showed no significant impacts in either reading or math after 3 years. Interaction terms in the equations that estimated the subgroup impacts indicated that the experience of the treatment by subgroups that did not demonstrate impacts at the subgroup level were not significantly different from the experience of the treatment by subgroups that did demonstrate impacts at the subgroup level. In other words, although some treatment subgroups showed reading outcomes that were significantly different from their peers in the control group, their reading impacts were not significantly different from the reading impacts of other subgroups. The implication of this statistical result is that the overall achievement effects of the Program—positive impacts in reading and no impacts in math—are more clear and reliable than any of the subgroup results.

DC Opportunity Scholarship Program

The OSP is operated by the Washington Scholarship Fund (WSF), a 501(c)3 organization based in the District of Columbia, under contract to the U.S. Department of Education's Office for Innovation and Improvement. To date, there have been five rounds of applications to the OSP (table 1). Applicants in spring 2004 (cohort 1) and spring 2005 (cohort 2) represent the majority of Program applicants; the evaluation sample was drawn from these two groups.[4] There were a smaller number of applicants in spring 2006 (cohort 3), spring 2007 (cohort 4), and spring 2008 (cohort 5) who were

---

[3] The subgroups that are analyzed in this study were designated prior to the collection and analysis of data and are of particular policy interest based on the Program statute and education policy literature. The subgroups are: (1) whether or not students attended a SINI school under the *No Child Left Behind Act* prior to application to the Programs—students were either attending a SINI-ever or SINI-never school; (2) whether students were relatively lower performing or relatively higher performing at baseline—students were either in the bottom one-third or the top two-thirds of the test score distribution; (3) student gender; (4) whether students were entering grades K-8 or 9-12 at the time of application; and (5) whether students were in application cohort 1 (applied in 2004) or application cohort 2 (applied in 2005).

[4] Descriptive reports on each of the first 2 years of implementation and cohorts of students have been previously prepared and released (Wolf, Gutmann, Eissa, Puma, and Silverberg 2005; Wolf, Gutmann, Puma, and Silverberg 2006) and are available on the Institute of Education Sciences' website at http://ies.ed.gov/ncee.

recruited and enrolled by WSF in order to keep the Program operating at capacity each year. These relatively late enrollees are not formally part of this impact analysis.

Table 1.   OSP Applicants by Program Status, Cohorts 1 Through 5, Years 2004-2008

	Cohort 1 (Spring 2004)	Cohort 2 (Spring 2005)	Total Cohort 1 and Cohort 2	Cohort 3 (Spring 2006), Cohort 4 (Spring 2007), and Cohort 5 (Spring 2008)	Total, All Cohorts
Applicants	2,692	3,126	5,818	2,760	8,578
Eligible applicants	1,848	2,199	4,047	1,722	5,769
Scholarship awardees	1,366	1,088	2,454	1,722	4,176
Scholarship users in initial year of receipt	1,027	797	1,824	1,057	2,881
Scholarship users fall 2005	919	797	1,716	NA	1,716
Scholarship users fall 2006	788	684	1,472	333	1,805
Scholarship users fall 2007	678	581	1,259	671	1,930
Scholarship users fall 2008	498	411	909	807	1,714

NOTES: Because most participating private schools closed their enrollments by mid-spring, applicants generally had their eligibility determined based on income and residency, and the lotteries were held prior to the administration of baseline tests. Therefore, baseline testing was not a condition of eligibility for most applicants. The exception was applicants entering the highly oversubscribed grades 6-12 in cohort 2. Those who did not participate in baseline testing were deemed ineligible for the lottery and were not included in the eligible applicant figure presented above, though they were counted in the applicant total. In other words, the cohort 2 applicants in grades 6-12 had to satisfy income, residency, and baseline testing requirements before they were designated eligible applicants and entered in the lottery.

The initial year of scholarship receipt was fall 2004 for cohort 1, fall 2005 for cohort 2, fall 2006 for cohort 3, fall 2007 for cohort 4, and fall 2008 for cohort 5.

SOURCES: OSP applications and WSF's enrollment and payment files.

Mandated Evaluation of the OSP

In addition to establishing the OSP, Congress mandated an independent evaluation of it be conducted, with annual reports on the progress of the study. The legislation indicated the evaluation should analyze the effects of the Program on various academic and nonacademic outcomes of concern to policymakers and use ". . . the strongest possible research design for determining the effectiveness" of the Program. The current evaluation was developed to be responsive to these requirements. In particular, the foundation of the evaluation is a randomized controlled trial (RCT) that compares outcomes of eligible applicants (students and their parents) randomly assigned to receive or not receive a scholarship. This decision was based on the mandate to use rigorous evaluation methods, the expectation that there would be more applicants than funds and private school spaces available, and the statute's requirement that

random selection be the vehicle for determining who receives a scholarship. An RCT design is widely viewed as the best method for identifying the independent effect of programs on subsequent outcomes (e.g., Boruch, de Moya, and Snyder 2002, p. 74). Random assignment has been used by researchers conducting impact evaluations of other scholarship programs in Charlotte, NC; New York City; Dayton, OH; and Washington, DC (Greene 2001; Howell, Wolf, Campbell, and Peterson 2002; Mayer, Peterson, Myers, Tuttle, and Howell 2002).

The recruitment, application, and lottery process conducted by WSF with guidance from the evaluation team created the foundation for the evaluation's randomized trial and determined the group of students for whom impacts of the Program are analyzed in this report. Because the goal of the evaluation was to assess both the short-term and longer term impacts of the Program, it was necessary to focus the study on early applicants to the Program (cohorts 1 and 2) whose outcomes could be tracked over at least 3 years during the evaluation period. During the first 2 years of recruitment, WSF received applications from 5,818 students. Of these, approximately 70 percent (4,047 of 5,818) were eligible to enter the Program (table 1). Of the total pool of eligible applicants, 2,308 students who were rising kindergarteners or currently attending public schools entered lotteries (492 in cohort 1; 1,816 in cohort 2), resulting in 1,387 students assigned to the treatment condition and 921 assigned to the control condition. These students constitute the evaluation's impact analysis sample and represent three-quarters of all students in cohorts 1 and 2 who were not already attending a private school when they applied to the OSP.[5]

Data Collection

The evaluation gathers information annually from students and families in the study, as well as from their schools, in order to address the key research questions. These data include:

- **Student assessments.** Measures of student achievement in reading and math for public school applicants come from the Stanford Achievement Test-version 9 (SAT-9)[6] administered by either the District of Columbia Public Schools (DCPS) (cohort 1 baseline) or the evaluation team (cohort 2 baseline and all follow-up data collection). The evaluation testing takes place primarily on Saturdays, during the spring, in

---

[5] Students in the District living in families with incomes below 185 percent of the poverty level who were already attending private schools were eligible for the Program. Since they were the lowest service priority, WSF only offered scholarships to existing private school students during the first year of partial program implementation. In that first year, 505 existing private school students were eligible applicants to the Program, and 216 were awarded scholarships in a lottery separate from the lottery designed for public school students. The private school students subject to a lottery in that first year were not followed for purposes of the evaluation because the nature of the treatment intervention was distinctive for them. Existing private school students sought a scholarship in order to remain in a private school and not to switch from a public to a private school.

[6] *Stanford Abbreviated Achievement Test (Form S)*, Ninth Edition. San Antonio, TX: Harcourt Educational Measurement, Harcourt Assessment, Inc., 1997.

locations throughout DC arranged by the evaluators. The testing conditions are similar for members of the treatment and control groups.

- **Parent surveys.** The OSP application included baseline surveys for parents applying to the Program. These surveys were appended to the OSP application form and therefore were completed at the time of application to the Program. Each spring after the baseline year, surveys of parents of all applicants are being conducted at the Saturday testing events, while parents are waiting for their children to complete their outcome testing. The parent surveys provide the self-reported outcome measures for parental satisfaction and safety.

- **Student surveys.** Each spring after the baseline year, surveys of students in grades 4 and above are being conducted at the outcome testing events. The student surveys provide the self-reported outcome measures for student satisfaction and safety.

- **Principal surveys.** Each spring, surveys of principals of all public and private schools operating in the District of Columbia are conducted. Topics include self-reports of school organization, safety, and climate; principals' awareness of and response to the OSP; and, for private school principals, why they are or are not OSP participants.

Several methods were used to encourage high levels of response to year 3 data collection in spring 2007 (year 3 cohort 1 outcomes) and spring 2008 (year 3 cohort 2 outcomes). Study participants were invited to at least three different data collection events if a member of the treatment group and at least five different data collection events if a member of the control group. Impact sample members received payment for their time and transportation costs if they attended a data collection event. The events were held on Saturdays except for one session that was staged on a weeknight. Multiple sites throughout DC were used for these events, and participants were invited to the location closest to their residence. When the address or telephone number of a participant was inaccurate, such cases were submitted to the tracing office at Westat and subject to intensive efforts to update and correct the contact information. Treatment and control group students were tested under the same conditions.

After these initial data collection activities were completed, the test score response rate for year 3 was 63.9 percent—the treatment group response rate was 67.8 percent, and the control group response rate was 57.8 percent, a response rate differential of 10 percentage points lower for the control group compared to the treatment group. To reduce this response rate differential, a random subsample of half of the control nonrespondents was drawn and subjected to intensive efforts at nonrespondent conversion (Wolf et al. 2009, pp. A-16–A-25). Since these initial nonrespondents were selected at random, each one that was successfully converted to a respondent counts double, as he or she "stands in" for an approximately similar control nonrespondent that was not subsampled (Kling, Ludwig, and Katz 2005; Sanbonmatsu, Kling, Duncan, and Brooks-Gunn 2006). The "effective" response rate after subsample conversion is the number of actual respondents prior to the subsample plus two times the number of subsampled respondents, all divided by the total number of students in the impact sample.

As a result of the subsample conversion process, the final effective test score response rate for year 3 was 68.5 percent, and the differential rate of response between the treatment and control groups was reduced to 1.7 percentage points higher for the control group.[7] The effective parent survey response rate was 67.9 percent.[8] The effective student survey response rate was 67.0 percent.[9] The principal survey was the data collection instrument with the lowest response rate in year 3, ranging from 51.8 percent to 57.3 percent depending on school sector and calendar year of administration.[10]

Missing outcome data create the potential for nonresponse bias in a longitudinal evaluation such as this one, if the nonrespondent portions of the sample are different between the treatment and control groups. Response rates for the various data collection instruments differed by less than 2 percent between the treatment and control groups, meaning that similar proportions of the treatment and control groups provided outcome data. In addition, nonresponse weights were used to equate the two groups on important baseline characteristics, thereby reducing the threat of nonresponse bias in this case.[11]

The test score response rate of 69 percent for this year 3 analysis of the OSP is higher than the response rates obtained in any of the three previous experimental evaluations of privately funded K-12 scholarship programs 3 years after random assignment. The previous evaluations of such programs in

---

[7] Specifically, the overall effective response rates were 67.8 percent for the treatment group and 69.5 percent for the control group. Prior to drawing the subsample, response rates for the control group were 41.4 percent (cohort 1) and 61.9 percent (cohort 2). Response rates (after drawing the subsample) for the control group were 51.1 percent (cohort 1) and 66.8 percent (cohort 2). After subsample weights were applied, the effective response rates for the control group were 60.9 percent (cohort 1) and 71.7 percent (cohort 2). Actual and effective response rates for the treatment group were 63.8 percent (cohort 1) and 68.9 percent (cohort 2). Eighty-five impact sample students awarded scholarships were entering grades 10 or higher at baseline and, therefore, were no longer grade-eligible for the OSP by year 3; these students are excluded from response rate calculations. See Wolf et al. 2009, appendix A, figure A-1 and tables A-4 and A-6 for a detailed breakdown of the response rates and a further discussion of the subsampling procedure.

[8] Specifically, the overall effective response rates were 67.4 percent for the treatment group and 68.6 percent for the control group. Response rates (after drawing the subsample) for the control group were 51.1 percent (cohort 1) and 66.2 percent (cohort 2). After subsample weights were applied, the effective response rates for the control group were 60.9 percent (cohort 1) and 70.5 percent (cohort 2). Actual and effective response rates for the treatment group were 65.2 percent (cohort 1) and 68.0 percent (cohort 2).

[9] Specifically, the overall effective response rates were 67.0 percent for the treatment group and 67.1 percent for the control group. Response rates (after drawing the subsample) for the control group were 50.9 percent (cohort 1) and 65.8 percent (cohort 2). After subsample weights were applied, the effective response rates for the control group were 60.7 percent (cohort 1) and 69.3 percent (cohort 2). Actual and effective response rates for the treatment group were 65.9 percent (cohort 1) and 67.4 percent (cohort 2).

[10] Since the principal survey is designed to gather information about all public and private schools in DC, as opposed to a defined set of students in the impact sample, the response rates for this instrument are broken down by school sector (public or private) and by academic year (since cohort 1 students in the study experienced year 3 in these schools in 2006-07 but cohort 2 students experienced year 3 in 2007-08). For the principal survey, response rates for the 2007-08 school year were 56.4 percent (public schools) and 57.3 percent (private schools). For the 2006-07 school year, response rates were 53.2 percent (public schools) and 51.8 percent (private schools).

[11] For additional details about data sources, collection methods, scale construction, response rates, subsampling for nonresponse conversion, and final nonresponse sample weights see Wolf et al. 2009, pp. A-5–A-25.

New York City and Washington, DC, reported year 3 test score response rates of 67 percent and 60 percent, respectively (Howell et al. 2006, p. 47). A previous experimental evaluation of the publicly funded Milwaukee Parental Choice Program reported test score response rates in year 3 of 47 percent for the treatment group and 23 percent for the control group (Rouse 1998, p. 555).

Participation in the OSP

In interpreting the impacts of the OSP, it is useful to examine the characteristics of the private schools that participate in the Program and the extent to which students offered scholarships (the treatment group) moved into and out of them during the first 3 years.

School Participation

The private schools participating in the OSP represent the choice set available to parents whose children received scholarships. That group of schools had mostly stabilized by the 2005-06 school year. The schools that offered the most slots to OSP students, and in which OSP students and the impact sample's treatment group were clustered, have characteristics that differed somewhat from the average participating OSP school. Although 56 percent of all participating schools were faith-based (39 percent were part of the Catholic Archdiocese of Washington), 82 percent of the treatment group attended a faith-based school, with 59 percent of them attending the 22 participating Catholic parochial schools. Twenty-two percent of treatment group students were attending a school that charged tuition above the statutory cap of $7,500 during their third year in the Program (table 2) even though 38 percent and 46 percent of participating schools charged tuitions above that cap in 2006-07 and 2007-08, respectively. The average OSP student in the treatment group attended a school with 261 students, while the school averages were 242 (2006-07) and 265 (2007-08) students across the set of all participating OSP schools.

While the characteristics of the participating private schools are important considerations for parents, in many respects it is how the schools differ from the public school options available to them that matters most. In the third year after applying to the OSP, students in the treatment and control groups did not differ significantly regarding the proportion attending schools that offered a separate library (88 vs. 91 percent), gyms (71 and 72 percent), and art programs (89 and 87 percent). Differences in school characteristics between the treatment and control groups 2 years after they applied to the OSP that were statistically significant at the .01 level included:

- Students in the treatment group were more likely than those in the control group to attend schools with a computer lab (96 vs. 87 percent), with special programs for advanced learners (48 vs. 32 percent), and that offered a music program (89 vs. 82 percent).

- Students in the treatment group were less likely than those in the control group to attend a school with a cafeteria facility (79 vs. 88 percent) or a nurse's office (30 vs. 81 percent).

- Students in the treatment group were also less likely than those in the control group to attend a school that offered special programs for non-English speakers (26 vs. 57 percent), special programs for students with learning problems (71 vs. 88 percent), counselors (69 vs. 82 percent), tutors (50 vs. 67 percent), and after-school programs (86 vs. 92 percent).

Table 2. **Features of Participating OSP Private Schools Attended by the Treatment Group in Year 3**

Characteristic	Weighted Mean	Highest	Lowest	Valid *N*
Archdiocesan Catholic schools (percent of treatment students attending)	59.2	NA	NA	66
Other faith-based schools (percent of treatment students attending)	22.5	NA	NA	66
Charging over $7,500 tuition (percent of treatment students attending)	22.3	NA	NA	48
Tuition	$6,620	$29,902	$3,600	48
Enrollment	260.5	1,072	10	43
Student *N*	701			

NOTES: "Valid *N*" refers to the number of schools for which information on a particular characteristic was available. When a tuition range was provided, the mid-point of the range was used. The weighted mean was generated by associating each student with the characteristics of the school he/she was attending and then computing the average of these student-level characteristics.

SOURCE: OSP School Directory information, 2004-05, 2005-06, 2006-07, and 2007-08, Washington Scholarship Fund.

Student Participation

As has been true in similar programs, not all students offered an OSP scholarship actually used it to enroll in a private school. For students assigned to the treatment group, during the first 3 years of the Program (figure 1):

- 346 out of 1,387 (25 percent) treatment group students never used the OSP scholarships offered to them;
- 473 treatment students (34 percent) used their scholarships during some but not all of the first 3 years after the scholarship award. Among these students are 142 students estimated to be "forced decliners," meaning that they could not continue to use their scholarship because they "graded out" (graduated high school), "earned out" (their

family income grew to exceed the Program's eligibility requirements), or there was no space for them in a participating high school;[12] and

- The remaining 568 treatment group students (41 percent) used their scholarship during the entire 3 years after the scholarship lottery.

Figure 1.   Proportions of Treatment Group Students Who Experienced Various Categories of Usage in First 3 Years
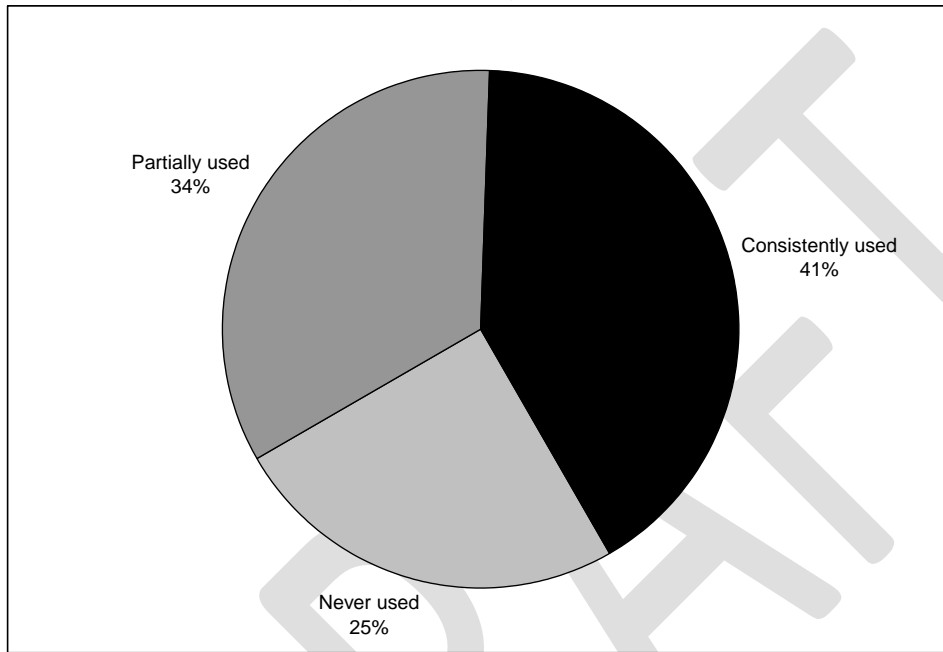
---

[12] The calculations regarding likely forced decliners were made using information from the baseline application/survey and administrative data provided by the WSF. A total of 85 students awarded scholarships in cohort 1 or cohort 2 were entering grades 10 or higher at baseline and therefore were no longer grade-eligible for the scholarship by the third year. A total of seven treatment students initially qualified for the Program but later reported family income of over 300 percent of the poverty level, thereby "earning out" of subsequent Program eligibility. The estimate of the number of students forced to decline their scholarships due to the lack of high school slots was calculated by comparing the higher rate of scholarship continuation for 7th graders moving to 8th grade with the lower rate of scholarship continuation for 8th graders moving to 9th grade. The difference between those two continuation rates, applied to the number of OSP students moving from 8th to 9th grade generates the estimate of forced decliners due to high school slot constraints of 50 (20 in year 2 plus 30 new ones in year 3). It is impossible to know for certain if all 50 of these students declined to use the scholarship solely or primarily because of high school slot constraints, and not for other reasons, or if some treatment students were forced to decline their scholarship at the very start due to high school slot constraints. Therefore, the total estimate of 142 forced decliners for outcome year 3 is simply an estimate based on the limited data available.

The reasons for not using the scholarship—either initially or consistently—varied. The most common reasons cited in year 3 by parents whose child did not use his/her scholarship that year and who completed surveys were (figure 2):

- Lack of available space in the private school they wanted their child to attend (22 percent of these parents);

- Child moved out of DC (21 percent of these parents);

- Child was accepted into a public charter school (19 percent of these parents); and

- Participating schools did not offer services for their child's learning or physical disability or other special needs (16 percent of these parents).

Figure 2.   Most Common Reasons Given by Parents for Declining to Use the OSP Scholarship in Year 3
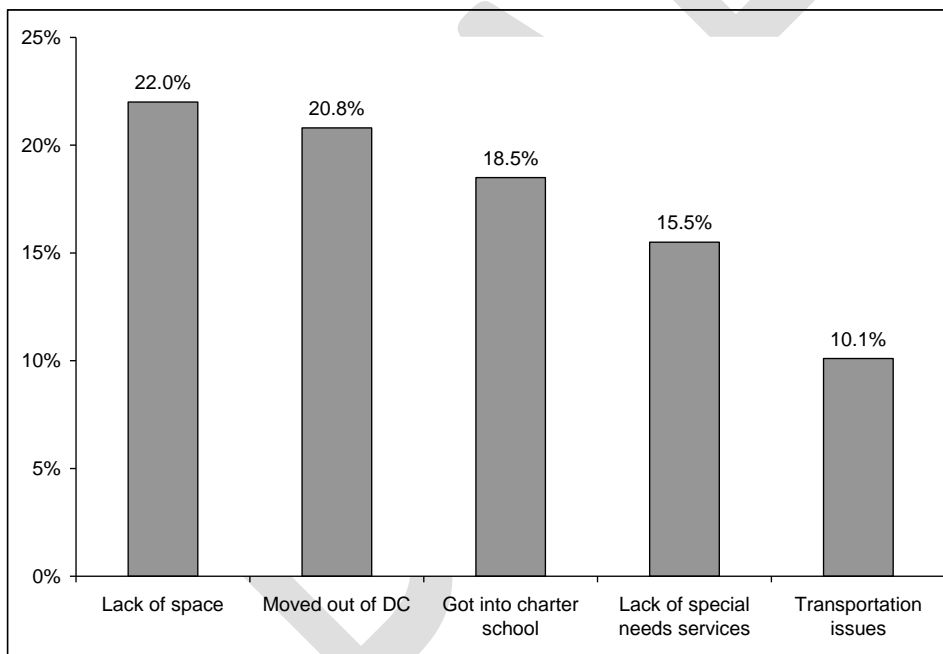
NOTES: Responses are unweighted. Respondents were able to select multiple responses, which generated a total of 180 responses provided by 153 parents. This equates to an average of 1.2 responses per parent.

SOURCE: Impact Evaluation Parent Surveys.

Overall Movement Into and Out of Private and Public Schools

Where did students who declined to participate in the OSP attend school instead? Children in the treatment group who never used the OSP scholarship offered to them, or who did not use the scholarship consistently, could have remained in or transferred to a public charter school or a traditional DC public school or enrolled in a non-OSP-participating private school. The same alternatives were available to students who applied to the OSP, were entered into the lottery, but were never offered a scholarship (the impact sample's control group); they could remain in their current DC public school (traditional or charter), enroll in a different public school, or try to find a way to attend a participating or nonparticipating private school. As indicated earlier, these choices could affect program impacts because traditional public, public charter, and private schools are presumed to offer different educational experiences and because previous studies suggest that switching schools has an initial short-term negative effect on student achievement (Hanushek, Kain, and Rivkin 2004).

To examine what types of schools students in both the treatment and control groups attended throughout the evaluation, it is necessary to access information from sources other than the WSF payment file, which informed the estimations of scholarship usage rates (figure 1). The WSF payment file contains comprehensive information about which students used an Opportunity Scholarship and what participating private school they attended. The payment file does not, however, include any information about the schools attended by scholarship decliners or members of the control group. To obtain comparable information about the type of school attended by all groups of students -- scholarship users, scholarship decliners, and members of the control group -- an alternative data source is required. The surveys administered annually to parents of students in the impact sample asked what school or schools the student attended that year. Although only 62 percent of study participants responded to that question in year 3 of the evaluation, response rates were approximately equal between members of the treatment and control groups, so that descriptive comparisons across the two groups should be valid. Still, readers should interpret the specific proportions presented for each sector as applying to treatment and control survey respondents and not necessarily to all members of those populations. Readers are further cautioned not to draw conclusions about the impact of the OSP in causing these descriptive patterns of school-sector enrollments.

Of the students in the evaluation who were entering grades 1-12 at the time of application[13], approximately three-fourths were attending traditional DC public schools, while the remaining one-fourth

---

[13] Rising kindergarteners are omitted from the data for this comparison, as many of them were not attending a school at baseline so their initial "school sector status" was ambiguous.

were attending public charter schools. Three years after random assignment, there was substantial variation across educational sectors (table 3).

Table 3. Percentage of the Impact Sample by Type of School Attended: At Baseline and in Year 3

| | Baseline | | 3 Years After Random Assignment | | |
| | Public | | Public | | Private |
	Traditional	Charter	Traditional	Charter	
Treatment	75.8	24.2	19.1	9.3	71.6
Control	73.7	26.3	53.9	33.8	12.3
Difference	2.1	-2.1	-34.7	-24.6	59.3

NOTES: The longitudinal statistics presented in this table exclude data from students who were rising kindergarteners at baseline to reduce the risk of compositional bias across the years examined. As a result, the type of school attended reported here may vary slightly from other cross-sectional descriptions of school attended found in this report. Student N = 1,985. Percent missing baseline: Treatment = 5.4, Control = 9.9; percent missing year 3: Treatment = 31.3, Control = 47.6. Some of the missing data rates for year 3 are a product of students naturally grading-out of the Program's eligibility requirements: 85 students are considered to have graded out during the third year, including 54 control group members and 31 treatment group members. Data are unweighted and represent actual responses. Given the rates of missing data, readers are cautioned against drawing firm conclusions.

SOURCES: Program applications and Evaluation Parent Surveys.

Based on these data from survey respondents, in the third year:

- Nineteen percent of the treatment group and 54 percent of the control group attended a traditional public school;

- Nine percent of the treatment group and 34 percent of the control group were enrolled in public charter schools; and

- Seventy-two percent of the treatment group and 12 percent of the control group attended a private school.

These data show how assignment to treatment is not perfectly correlated with private school attendance and that assignment to the control group does not necessarily entail attendance at a traditional public school. A number of school choices are available in DC to parents who seek alternatives to their neighborhood public school, and many members of the control group availed themselves of school choice options even if they were not awarded an Opportunity Scholarship. In a sense, these cases of treatment students remaining in public schools and control students finding their own way into private schools do not necessarily represent "non-compliance" with the treatment assignment. The lottery only randomly assigned students to the *offer* of a scholarship or not. It could not require that they use a scholarship if offered one. The control group remains the ideal counter-factual. Presumably a small proportion of the students in the evaluation would have attended private schools absent the intervention of the scholarship program, since some members of the control group did just that. For those readers interested in the effect of actually using a scholarship compared to the control condition, or the effect of attending private school

14

compared to attending public school, estimates of those programmatic effects are provided as a complement to the main experimental impacts by adjusting the experimental impacts to account for scholarship non-use and private school attendance by the control group.

The enrollment patterns of students who attended SINI schools is a special focus of this evaluation, given that Congress assigned that specific group of students to be the highest service priority of the OSP (Section 306). Among the applicant parents in the impact sample who provided the identity of their child's school (table 4):

- Fifty-six percent of the treatment and 52 percent of the control parents reported that, at the time they applied to the Program, their child was attending a school designated in need of improvement between 2003 and 2005 (SINI ever).

- Three years after random assignment, the number of treatment group students reported to be attending SINI-ever schools was 15 percent, while the number of control group students in such schools was 42 percent.

Table 4.   Percentage of the Impact Sample Attending SINI Schools: Baseline and Year 3

	Baseline		3 Years After Random Assignment		
	SINI-Ever Schools	SINI-Never Schools	SINI-Ever Schools	SINI-Never Schools	Private
Treatment	55.7	44.3	14.6	13.8	71.6
Control	52.3	47.8	42.0	45.7	12.3
Difference	3.4	-3.4	-27.4	-31.9	59.3

NOTES: Schools were identified as SINI ever if they were officially designated as in need of improvement under the *Elementary and Secondary Education Act* between 2003 and 2005. The longitudinal statistics presented in this table exclude data from students who were rising kindergarteners at baseline to reduce the risk of compositional bias across the years examined. As a result, the type of school attended reported here may vary slightly from other cross-sectional descriptions of school attended found in this report. Student N = 1,985. Percent missing baseline: Treatment = 5.4, Control = 9.9; percent missing Year 3: Treatment = 31.3, Control = 47.6. Some of the missing data rates for year 3 are a product of students naturally grading out of the Program's eligibility requirements: 85 students are considered to have graded out during the third year, including 54 control group members and 31 treatment group members. Data are unweighted and represent actual responses. Given the rates of missing data, readers are cautioned against drawing firm conclusions.

SOURCES: Program applications and Evaluation Parent Surveys.

The movement of impact sample students between public (both traditional and charter) and private schools or between SINI and non-SINI schools masks some additional transitions because students can change schools within the same sector. That is, some students moved from one charter school to another charter school, or one private school to another one. Over the course of all 3 years since random assignment:

- Among the treatment group, 3 percent remained in the same public school they were in when they applied to the Program, 46 percent switched schools once, 40 percent switched twice, and 11 percent switched three times; and

- Among the control group, 15 percent remained in the same public school they were in when they applied to the Program, 40 percent switched schools once, 37 percent switched schools twice, and 8 percent switched three times.

Both groups experienced higher rates of school mobility than the typical annual rate for urban students. The treatment group students switched schools at an annual rate of 53 percent and the control group at an annual rate of 46 percent from the baseline year to year 3.[14] In contrast, other studies of urban school populations report annual school-switching rates of 22 to 28 percent (Witte 2000, p. 144; Wong et al. 1997, p. 17). The higher school-switching rate of the treatment group compared to the control group of 7 percentiles annually (21 percentiles cumulatively over 3 years) is itself statistically significant.[15] The treatment group students have switched schools more frequently than the control group students, but both groups have switched schools more often than is the norm for inner-city K-12 students.

School-switching is both an inherent aspect of the K-12 experience and a component part of the scholarship program treatment. Students in inner-city school districts with limited grade-range schools (e.g. K-5, 6-8, 9-12) and many public charter schools will tend to switch schools frequently during their educational careers. Such school mobility is a natural aspect of the counterfactual demonstrated by our control group. Moreover, public school students offered private school scholarships must switch schools, at least initially, in order to avail themselves of such an intervention. In the context of an experimental evaluation that begins when all of the students are in public schools[16], separating the effect of the school choice intervention from the effect of an action *necessitated by use of the school choice intervention* (i.e. switching schools) would produce only a partial and misleading estimate of the true and total impact of the intervention. As a result, we merely describe the school-switching that the students in our evaluation experienced. We do not control for the effects of school switching after random assignment because doing so would undermine the validity of our impact estimates.

Impact of the Program After 3 Years: Key Outcomes

---

[14] Annual school-switching rates were calculated by totaling the number of switches since baseline (2,205 for the treatment and 1,271 for the control students), dividing by 3 years to generate the average annual number of switches, and further dividing by the number of students in each group to generate the average annual rate.

[15] In an Ordered Logit estimation of the number of school switches experienced by students in the impact sample, the treatment variable was a statistically significant predictor of school switching ($Z=2.15$, $p=.03$).

[16] For longitudinal evaluations of voucher programs using observational data (e.g. Witte at al., 2009), the use of control variables for school-switching can be justified as an attempt to capture unmeasured characteristics (e.g. family instability) that otherwise might confound the analysis. Since random assignment approximately equalizes all student and family characteristics at baseline, such controls are unnecessary and, as discussed above, actually would bias the evaluation of program impact.

Research Methodology

The statute that authorized the OSP mandated that the Program be evaluated with regard to its impact on student test scores and school safety, as well as the "success" of the Program, which, in the design of this study, includes satisfaction with school choices. The impacts of the Program on these outcomes are presented in three ways: (1) the impact of the *offer* of an OSP scholarship, derived straight from comparing outcomes of the treatment and control groups, and (2) the impact of *using* an OSP scholarship, calculated from the unbiased treatment-control group comparison, but statistically netting out students who declined to use their scholarships,[17] and (3) the effect of private schooling, calculated through Instrumental Variable (IV) analysis.

As with previous experimental analyses of the impacts of voucher or voucher-like programs (e.g. Howell et al. 2002), the estimates of the impact of the scholarship offer, called "intent-to-treat" or ITT, **include in the treatment average the outcomes for all scholarship recipients who provided outcome data, including recipients who never used their scholarships.** In other words, scholarship "decliners" remain in the study, as full members of the treatment group, for the purposes of generating these experimental estimates of the impact of the scholarship offer on student outcomes.

Because the RCT approach has the important feature of generating comparable treatment and control groups, we used a common set of analytic techniques, designed for use in social experiments, to estimate the Program's impact on test scores and the other outcomes listed above. These analyses began with the estimate of simple mean differences using the following equation, illustrated using the test score of student i in year t (Y_{it}):

(1) $Y_{it} = \alpha + \tau\, T_{it} + \varepsilon_{it}$ if $t > k$ (period after Program takes effect),

where T_{it} is equal to 1 if the student *receives an offer to participate* in the OSP (i.e., the award rather than the actual use of the scholarship) and is equal to 0 otherwise. Equation (1) therefore estimates the effect of the **offer** of a scholarship on student outcomes. Under this ITT model, all students who were randomly

---

[17] This analysis uses straightforward statistical adjustments to account for the impact sample respondents who received the offer of a scholarship but declined to use it (the "decliners") and also a small number of control group members who never received a scholarship offer but who, by virtue of having a sibling with an OSP scholarship, ended up in a participating private school (we call this "program-enabled crossover"). These adjustments essentially re-scale the impact of the scholarship offer across the subset of treatment students that actually used the scholarship, increasing the size of the effect estimates. Since the adjustment is merely arithmetic, and does not involve any new statistical estimation of impact, it cannot make a statistically insignificant result significant.

assigned by virtue of the lottery are included in the analysis, regardless of whether a member of the treatment group used the scholarship to attend a private school or for how long.

Proper randomization renders experimental groups approximately comparable, but not necessarily identical. In the current study, some modest differences, almost all of which are not significant, exist between the treatment group and the control group counterfactual at baseline (Wolf et al. 2007, p. 13). The basic regression model can, therefore, be improved by adding controls for observable baseline characteristics to increase the reliability of the estimated impact by accounting for minor differences between the treatment and control groups at baseline and improving the precision of the overall model. This yields the following equation to be estimated:

$$(2) \ Y_{it} = \alpha + \tau \ T_{it} + X_i \gamma + \delta_R \ R_{it} + \delta_M \ M_{it} + \varepsilon_{it}.$$

where X_i is a vector of student and/or family characteristics measured at baseline and known to influence future academic achievement, and R_{it} and M_{it} refer to **baseline** reading and mathematics scores, respectively.[18] In this model, τ—the parameter of sole interest—represents the effect of scholarships on test scores for students in the Program, conditional on X_i and the baseline test scores. The δ's reflect the degree to which test scores are, on average, correlated over time. With a properly designed RCT, baseline test scores and controls for observable characteristics that predict future achievement should improve the precision of the estimated ITT impact.

To estimate the magnitude of the impact of actually using a scholarship, if offered one, we employed conventional Bloom adjustments (Bloom 1984), re-scaling the ITT impacts over the subgroup of treatment members who actually used their scholarships.

The method for estimating the effect of attending versus not attending private schools, 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

---

[18] The consistent set of covariates used to generate impact estimates were: student's baseline reading scale score, student's baseline math scale score, student attended a school designated SINI 2003-05 indicator, student's age (in months) at the time of scholarship application, student's forecasted entering grade for the next school year, student's gender, student's race (African-American indicator), special needs indicator, mother employed part-time or full-time indicator, household income, total number of children in student's household, the number of months the family has lived at its current address, and the number of days from September 1 to the date of outcome testing for each student. Some missing baseline data were imputed by fitting stepwise models to each covariate using all of the available baseline covariates as potential predictors.

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.[19]

The main focus of this study was on the impacts of the OSP on the overall group of students who were randomly assigned. The study provides additional consideration of the programmatic impacts on policy-relevant subgroups of students. The subgroups were designated prior to data collection and include students who were attending SINI versus non-SINI schools at application, those performing relatively higher or lower at baseline, girls or boys, elementary versus high school students, and those from application cohort 1 or cohort 2.

Previous reports released in spring 2007 and spring 2008 indicated that 1 and 2 years after application, there were no statistically significant impacts on overall academic achievement or on student perceptions of school safety or satisfaction (Wolf et al. 2007; Wolf et al. 2008). Parents were more satisfied if their child was in the Program and viewed their child's school as safer and more orderly. Among the secondary analyses of subgroups, there were impacts on math test scores in year 1 for students who applied from non-SINI schools and those with relatively higher pre-Program test scores and impacts in reading test scores (but not math) in year 2 for those same two subgroups plus those who applied in the first year of Program implementation. Statistical adjustments for multiple comparisons suggested there is a possibility that the subgroup achievement impacts in years 1 and 2 were chance discoveries. The analyses in this report were conducted using data collected on students 3 years after they applied to the OSP.

Impacts on Student Achievement

- Across the full sample, there was a statistically significant impact on reading achievement of 4.5 scale score points from the offer of a scholarship and 5.3 scale score points from the use of a scholarship (table 5). These impacts are equivalent to 3.1 and 3.7 months of additional learning, respectively (table 7).[20]

- Attending a private school in year 3 had a statistically significant effect of 7.1 scale score points, equivalent to 5 months of additional learning (tables 5 and 8).

---

[19] IV analyses were limited to achievement impacts that were found to be statistically significant during the estimation of ITT impacts.

[20] Scale score impacts were converted to approximate months of learning first by dividing the impact effect size by the effect size of the weighted (by grade) average annual increase in reading scale scores for the control group. The result was the proportion of a typical year of achievement gain represented by the Programmatic impact. That number was further divided by nine to convert the magnitude of the gain to months, since the official school year in the District of Columbia comprises nine months of instruction.

- There was no statistically significant impact on math achievement overall from the offer of a scholarship or from the use of a scholarship (tables 5 and 8).[21]

---

[21] The magnitudes of these estimated achievement effects are below the threshold of .12 standard deviations, estimated by the power analysis to be the study's Minimum Detectable Effect size.

Table 5. **Year 3 Impact Estimates of the Offer and Use of a Scholarship on the Full Sample: Academic Achievement**

| Student Achievement | Impact of the Scholarship Offer (ITT) | | | Impact of Scholarship Use | *p-value* | Effect of Private Schooling | *p-value* |
	Treatment Group Mean	Control Group Mean	Difference (Estimated Impact)	Adjusted Impact Estimate		IV Estimate	
Reading	635.44	630.98	4.46*	5.27*	.01	7.1*	.04
Math	630.15	629.35	.81	.95	.62	NA	NA

*Statistically significant at the 95 percent confidence level.

NOTES: Means are regression adjusted using a consistent set of baseline covariates. Impacts are displayed in terms of scale scores. Valid *N* for reading = 1,460; math = 1,468. Separate reading and math sample weights used. Robust regression calculations generated by clustering at the family level.

Subgroup Impacts on Student Achievement

In addition to determining the general impacts of the OSP on all study participants, this evaluation also reports Programmatic impacts on policy-relevant subgroups of students. The subgroups were designated prior to data collection and include students who were attending SINI versus non-SINI schools at application, those performing relatively higher or lower at baseline, girls or boys, elementary versus high school students, and those from application cohort 1 or cohort 2.

The statistical significance of impacts for particular subgroups of students in year 3 are consistent with those for students overall in math but not in reading. There were no impacts on math achievement for any of the 10 subgroups examined, as was true for the full impact sample. The offer of a scholarship, and therefore the use of a scholarship, had a statistically significant positive impact on reading achievement in the third year for half of the student subgroups, at the subgroup including at least two subgroups that applied with a relative advantage in academic preparation (table 6). The subgroups with positive reading impacts include:[22]

- Students in the treatment group who had attended non-SINI public schools prior to the Program, who scored an average of 6.6 scale score points higher in reading than students in the control group from non-SINI schools (the impact of the offer of a scholarship); the calculated impact of using a scholarship was 7.7 scale score points.

- Students in the treatment group who entered the Program in the higher two-thirds of the applicant test-score performance distribution—averaging a 43 National Percentile Rank in reading at baseline—who scored an average of 5.5 scale score points higher in reading than students in the control group who applied to the OSP in the higher two-

---

[22] Each of these findings refers to one subgroup and not to the paired subgroup categories, so that a significant finding refers to a treatment vs. control difference, not a difference between, for example, males and females.

thirds of the test-score distribution; the impact of using a scholarship for this group was 6.2 scale score points.

- Female students in the treatment group, who scored an average of 5.1 scale score points higher in reading than females in the control group; the impact of using a scholarship was 5.8 scale score points.

- Students in the treatment group who entered the Program in grades K-8, who scored an average of 5.2 scale score points higher in reading than students in the control group who applied to the OSP entering grades K-8; the impact of using a scholarship was 6.0 points.

- Students in the treatment group from the first cohort of applicants, who scored an average of 8.7 scale score points higher in reading than students in the control group from cohort 1; the impact of using a scholarship was 11.7 scale score points.

The analysis did not show statistically significant subgroup impacts for students who applied from a school designated SINI between 2003 and 2005, students who entered the program in the lower one-third of the applicant test-score performance distribution, male students, students who entered the Program from high school, and cohort 2 students.[23]

The five statistically significant subgroup impacts of the OSP on reading test scores in year 3 were the product of an analysis involving multiple comparisons of treatment and control group members. Under such conditions, statistically significant findings can emerge by chance, even as a result of comparatively imprecise analyses at the subgroup level (Schochet 2008). Statistical adjustments to account for the multiple comparisons suggest that two of the five significant subgroup achievement impacts in reading—the impact on female students and the impact on cohort 1 students—may be false discoveries and therefore should be interpreted with caution. Another way to interpret the subgroup results in the light of the multiple comparison adjustments is that the reading impacts at the subgroup level are robust to adjustments for multiple comparisons in the case of the non-SINI, higher baseline performing, and K-8 student subgroups but not robust to such adjustments in the case of the female and cohort 1 subgroups.

Interaction terms used to test the significance of the differences between the treatment impacts on each subgroup pair (e.g., SINI versus non-SINI) all proved to be insignificant. The practical meaning of that result is that, although some district subgroups demonstrated statistically significant reading impacts at the subgroup level, the impact of the treatment was not significantly different across

---

[23] Since the analysis of impacts on subgroups draws on smaller samples of students, such analyses inevitably have less power to detect statistically significant impacts than do analyses on the entire impact sample. The five subgroup impacts on reading discussed here were detectable, even with lower power, because they were larger in magnitude than the overall reading impact.

subgroups. Moreover, these estimations of program impacts at the subgroup level inevitably are less precise, and therefore less likely to identify statistically significant impacts, than the overall analysis of impacts because each subgroup is merely a portion of the much larger impact sample. For these reasons, the overall achievement results of a statistically significant program impact in reading and no significant impact in math are the clearest and most reliable achievement results from the year 3 analysis and should be given the greatest weight when evaluating the Program on the metric of student achievement.

Table 6.   Year 3 Impact Estimates of the Offer and Use of a Scholarship on Subgroups: Reading Achievement

	Reading						
	Impact of the Scholarship Offer (ITT)			Impact of Scholarship Use (IOT)		Effect of Private Schooling	
Student Achievement: Subgroups	Treatment Group Mean	Control Group Mean	Difference (Estimated Impact)	Adjusted Impact Estimate	*p*-value	IV Estimate	*p*-value
SINI ever	649.77	648.25	1.52	1.81	.59	NA	NA
SINI never	625.29	618.72	6.57**	7.72**	.01	10.25*	.04
Lower performance	614.48	612.38	2.10	2.68	.47	NA	NA
Higher performance	644.74	639.29	5.45*	6.21*	.02	9.52*	.02
Male	631.32	627.48	3.83	4.67	.15	NA	NA
Female	639.30	634.24	5.07*	5.81*	.04	6.08	.15
K-8	627.30	622.07	5.23**	6.04**	.01	8.28*	.02
9-12	682.41	682.50	-.10	-.14	.98	NA	NA
Cohort 2	625.64	622.27	3.37	3.87	.09	NA	NA
Cohort 1	672.87	664.17	8.70*	11.67*	.04	15.75	.05

*Statistically significant at the 95 percent confidence level.
**Statistically significant at the 99 percent confidence level.

NOTES: Means are regression-adjusted using a consistent set of baseline covariates. Impacts are displayed in terms of scale scores. Effect sizes are in terms of standard deviations. Total valid N for reading = 1,460. Reading sample weights used. Robust regression calculations generated by clustering at the family level.

Table 7. **Year 3 Impact Estimates of the Offer and Use of a Scholarship on Subgroups: Math Achievement**

	Math						
	Impact of the Scholarship Offer (ITT)				Impact of Scholarship Use (IOT)		
Student Achievement: Subgroups	Treatment Group Mean	Control Group Mean	Difference (Estimated Impact)	Effect Size	Adjusted Impact Estimate	Effect Size	*p*-value
SINI ever	646.73	646.56	.17	.01	.20	.01	.95
SINI never	618.39	617.12	1.27	.04	1.49	.04	.58
Lower performance	615.43	615.08	.35	.01	.45	.02	.90
Higher performance	636.64	635.65	.98	.03	1.12	.04	.64
Male	629.35	629.31	.04	.00	4.67	.00	.99
Female	630.92	629.38	1.54	.05	1.76	.06	.50
K-8	621.74	620.73	1.01	.03	1.16	.03	.57
9-12	678.77	679.18	-.41	-.02	-.60	-.03	.92
Cohort 2	619.32	619.53	-.21	-.01	-.24	-.01	.91
Cohort 1	671.48	666.74	4.74	.23	6.37	.31	.19

NOTES: Means are regression-adjusted using a consistent set of baseline covariates. Impacts are displayed in terms of scale scores. Effect sizes are in terms of standard deviations. Total Valid N for math = 1,468. Math sample weights used. Robust regression calculations generated by clustering at the family level.

It is useful to place the estimated effect sizes for these overall and subgroup impacts in context (table 8). The overall reading impact of the scholarship offer (ITT) of 4.5 scale score points is equivalent to 3.1 months of additional learning for members of the treatment group. The overall impact of the actual use of a scholarship (IOT) of 5.3 scale score points is equivalent to 3.7 additional months of learning. The SINI-never impacts on reading are equivalent to 4.1 additional months of learning for the offer and 4.9 additional months of learning for the use of a scholarship. With the exception of cohort 1, the positive reading achievement impacts on the other subgroups ranged from 3 to 5 additional months of learning, or one-third to one-half of a typical 9-month school year. The reading impacts for cohort 1 are equivalent to 1.5 to 2 years of extra learning (14 to 19 months).

Table 8. **Estimated Impacts in Months of Schooling From the Offer and Use of a Scholarship for Statistically Significant Reading Impacts After 3 Years**

Student Achievement: Reading	Impact of the Scholarship Offer (ITT)		Impact of Scholarship Use (IOT)		Effect of Private Schooling	
	Effect Size	Months of Schooling	Effect Size	Months of Schooling	Effect Size	Months of Schooling
Full sample	.13	3.1	.15	3.7	.22	5.0
SINI never	.19	4.1	.22	4.9	.30	6.5
Higher performance	.17	4.0	.19	4.6	.30	7.0
Female	.15	3.1	.17	3.6	.19	3.7
K-8	.15	2.9	.17	3.3	.25	4.6
Cohort 1	.31	14.1	.42	18.9	.57	25.5

NOTES: Effect sizes are expressed as a proportion of a standard deviation of the distribution of values observed for the study control group. One full standard deviation above and below the average value for a variable such as outcome test scores contains 64 percent of the observations in the distribution. Two full standard deviations above and below the average contain 95 percent of the observations. Scale score impacts were converted to approximate months of learning first by dividing the impact effect size by the effect size of the weighted (by grade) average annual increase in reading scale scores for the control group. The result was the proportion of a typical year of achievement gain represented by the Programmatic impact. That number was further divided by nine to convert the magnitude of the gain to months, since the official school year in the District of Columbia comprises 9 months of instruction.

Impacts on Reported Safety and an Orderly School Climate

School safety is a valued feature of schools for the families who applied to the OSP. A total of 17 percent of cohort 1 parents at baseline listed school safety as their most important reason for seeking to exercise school choice — second only to academic quality (48 percent) among the available reasons (Wolf et al. 2005, p. C-7). A separate study of why and how OSP parents chose schools, which relied on focus group discussions with participating parents, found that school safety was among their most important educational concerns (Stewart, Wolf, and Cornman 2005, p. v).

Unlike student achievement, there are no specific tests to evaluate the safety of a school. There are various indicators of the relative orderliness of the school environment, such as the presence or absence of property destruction, cheating, bullying, and drug distribution to name a few. Students and parents can be surveyed regarding the extent to which such indicators of disorder are or are not a problem at their or their child's school. The responses then can be consolidated into an index of safety and an orderly school climate and analyzed, as we do here.

Parent Self-Reports

Overall, the parents of students offered an Opportunity Scholarship in the lottery subsequently reported their child's school to be safer and more orderly than did the parents of students in the control group. The impact of the offer of a scholarship on parental perceptions of safety and an

orderly school climate was 1.01 on a 10-point index of indicators of school safety and orderliness, an effect size of 0.29 standard deviations (table 9). The impact of using a scholarship was 1.20 on the index, with an effect size of .34 standard deviations.

Table 9.   Year 3 Impact Estimates of the Offer and Use of a Scholarship on the Full Sample and Subgroups: Parent Perceptions of Safety and an Orderly School Climate

Safety and an Orderly School Climate: Parents	Impact of the Scholarship Offer (ITT)				Impact of Scholarship Use (IOT)		
	Treatment Group Mean	Control Group Mean	Difference (Estimated Impact)	Effect Size	Adjusted Impact Estimate	Effect Size	*p*-value of estimates
Full sample	**8.08**	**7.07**	**1.01****	**.29**	**1.20****	**.34**	**.00**
SINI ever	7.91	6.74	1.16**	.32	1.38**	.38	.00
SINI never	8.20	7.30	.90**	.27	1.06**	.32	.00
Lower performance	7.81	6.80	1.02**	.28	1.30**	.36	.01
Higher performance	8.20	7.19	1.01**	.29	1.15**	.33	.00
Male	8.12	7.06	1.06**	.30	1.29**	.37	.00
Female	8.04	7.08	.96**	.28	1.10**	.32	.00
K-8	8.22	7.29	.93**	.27	1.08**	.32	.00
9-12	7.31	5.80	1.51*	.40	2.15*	.56	.02
Cohort 2	8.29	7.33	.97**	.29	1.11**	.33	.00
Cohort 1	7.28	6.08	1.20*	.31	1.61*	.42	.03

*Statistically significant at the 95 percent confidence level.

**Statistically significant at the 99 percent confidence level.

NOTES: Means are regression-adjusted using a consistent set of baseline covariates. Effect sizes are in terms of standard deviations. Valid *N* = 1,423. Parent survey weights used. Robust regression calculations generated by clustering at the family level.

This impact of the offer of a scholarship on parental perceptions of safety and an orderly school climate was consistent across all subgroups of students examined, including parents of students from SINI and non-SINI schools, parents of students who entered the program with relatively higher and lower levels of academic achievement, parents of male and female students, parents of students in grades K-8 and 9-12, and parents of both cohort 1 and cohort 2. All of these subgroup impacts on parental views of school safety remained statistically significant after adjustments to account for multiple comparisons.

Because the impacts of the scholarship offer on perceptions of safety and an orderly school climate were statistically significant for these subgroups of parents, the Programmatic impacts on actual scholarship users were also statistically significant. For example, the impact of using a scholarship on parental perceptions of school safety for these affected subgroups ranged from 1.06 for parents whose students were in SINI-never schools to 2.15 for parents of students entering grades 9-12 at baseline, which equates to subgroup effect sizes ranging from .32 to .56 standard deviations (table 9).

Student Self-Reports

The students in grades 4-12 who completed surveys paint a different picture about school safety at their school than do their parents. The student index of school climate and safety asked students if they personally had been a victim of theft, drug-dealing, assaults, threats, bullying, or taunting or had observed weapons at school. On average, reports of school climate and safety by students offered scholarships through the lottery were not statistically different from those of the control group, either overall or for any specific subgroup (table 10).

Table 10.   Year 3 Impact Estimates of the Offer and Use of a Scholarship on the Full Sample and Subgroups: Student Reports of Safety and an Orderly School Climate

Safety and an Orderly School Climate: Students	Impact of the Scholarship Offer (ITT)				Impact of Scholarship Use (IOT)		*p*-value of estimates
	Treatment Group Mean	Control Group Mean	Difference (Estimated Impact)	Effect Size	Adjusted Impact Estimate	Effect Size	
Full sample	**6.17**	**6.05**	**.12**	**.06**	**.14**	**.07**	**.36**
SINI ever	6.06	5.98	.08	.04	.10	.05	.66
SINI never	6.24	6.10	.14	.08	.17	.10	.41
Lower performance	6.04	5.89	.14	.07	.19	.09	.55
Higher performance	6.22	6.12	.10	.06	.12	.07	.50
Male	5.96	5.86	.10	.05	.12	.06	.59
Female	6.37	6.23	.13	.07	.16	.09	.44
4-8	6.12	6.00	.12	.06	.14	.07	.39
9-12	6.42	6.31	.11	.06	.15	.08	.74
Cohort 2	6.16	6.01	.16	.08	.18	.09	.29
Cohort 1	6.16	6.21	-.04	-.02	-.06	-.03	.88

NOTES: Means are regression-adjusted using a consistent set of baseline covariates. Effect sizes are in terms of standard deviations. Valid N = 1,098. Survey given to students in grades 4-12. Student survey weights used. Robust regression calculations generated by clustering at the family level.

Impacts on School Satisfaction

Economists have long used customer satisfaction as a proxy measure for product or service quality (see Johnson and Fornell 1991). While not specifically identified as an outcome to be studied, it is an indicator of the "success of the Program in expanding options for parents," which Congress asked the evaluation to consider.[24] Satisfaction is also an outcome studied in the previous evaluations of K-12 scholarship programs, all of which concluded that parents tend to be significantly more satisfied with their

---

[24] Section 309 of the *District of Columbia School Choice Incentive Act of 2003*.

child's school if they have had the opportunity to select it (see Greene 2001, pp. 84-85). Satisfaction of both parents and students was measured by the percentage that assigned a grade of A or B to their child's or their school.

Parent Self-Reports

About 30 months after the start of their experience with the OSP, parents overall are more satisfied with their child's school if they were offered a scholarship and if their child used a scholarship to attend a participating private school. A total of 74 percent of treatment parents assigned their child's school a grade of A or B compared with 63 percent of control parents ─ a difference of 11 percentage points (impact of the offer of a scholarship); the impact of using a scholarship was a difference of 12 percentage points in parent's likelihood of giving their child's school a grade of A or B. The effect sizes of these impacts were .22 and .26, respectively (table 11).

There were also statistically significant positive impacts of the Program on school satisfaction for 7 of 10 subgroups, including parents of students from non-SINI schools, those who had higher test-score performance at baseline, were male or female, entering grades K-8, or were in cohort 1 or cohort 2. Parents of these students were significantly more likely to give their child's school a grade of A or B if they were in the treatment group. The effect sizes ranged from .16 to .41 standard deviations for the offer of, and from .19 to .55 standard deviations for the use of, a scholarship. All seven of the parent satisfaction subgroup impacts that initially were statistically significant remained significant after adjustments for the fact that they were the product of multiple comparisons.

Three groups of parents were no more satisfied with their child's school if they had been offered a scholarship. Parents of students who entered the program from SINI schools, with lower levels of academic performance at baseline, or in grades 9-12 were just as likely to grade their child's school A or B if they were in the treatment as the control group. As described previously, those subgroups of students did not exhibit reading achievement impacts from the Program at the subgroup level of analysis.

Table 11. **Year 3 Impact Estimates of the Offer and Use of a Scholarship on Subgroups: Parent Reports of Satisfaction with Their Child's School**

Parents Who Gave Their School a Grade of A or B	Impact of the Scholarship Offer (ITT)				Impact of Scholarship Use (ITT)		p-value of estimates
	Treatment Group Mean	Control Group Mean	Difference (Estimated Impact)	Effect Size	Adjusted Impact Estimate	Effect Size	
Full sample	**.74**	**.63**	**.11****	**.22**	**.12****	**.26**	**.00**
SINI ever	.69	.63	.06	.13	.07	.15	.16
SINI never	.78	.64	.14**	.29	.17**	.35	.00
Lower performance	.63	.57	.06	.12	.08	.16	.21
Higher performance	.79	.67	.13**	.27	.15**	.31	.00
Male	.71	.60	.11**	.22	.13**	.27	.01
Female	.77	.67	.10**	.22	.12**	.25	.01
K-8	.77	.64	.13**	.27	.15**	.31	.00
9-12	.59	.60	-.01	-.03	-.02	-.04	.88
Cohort 2	.75	.68	.08*	.16	.09*	.19	.02
Cohort 1	.68	.48	.20**	.41	.27**	.55	.00

*Statistically significant at the 95 percent confidence level.

**Statistically significant at the 99 percent confidence level.

NOTES: Means are regression-adjusted using a consistent set of baseline covariates. Impact estimates are reported as marginal effects. Effect sizes are in terms of standard deviations. Valid $N = 1,410$. Parent survey weights used. Robust regression calculations generated by clustering at the family level.

Student Self-Reports

As was true with the school safety and climate measures, students had a different view of their schools than did their parents. Three years after random assignment, there were no significant differences between the treatment group and the control group in their likelihood of assigning their schools a grade of A or B (table 12).[25] Student reports of school satisfaction were statistically similar between the treatment and control groups for all 10 subgroups examined.

---

[25] Only students in grades 4-12 were administered surveys, so the satisfaction of students in early elementary grades is unknown.

Table 12.   Year 3 Impact Estimates of the Offer and Use of a Scholarship on the Full Sample and Subgroups: Student Reports of Satisfaction with Their School

Students Who Gave Their School a Grade of A or B	Impact of the Scholarship Offer (ITT)				Impact of Scholarship Use (IOT)		*p*-value of estimates
	Treatment Group Mean	Control Group Mean	Difference (Estimated Impact)	Effect Size	Adjusted Impact Estimate	Effect Size	
Full sample	**.71**	**.73**	**-.03**	**-.06**	**-.03**	**-.07**	**.41**
SINI ever	.64	.72	-.08	-.18	-.09	-.21	.07
SINI never	.76	.74	.02	.05	.03	.07	.60
Lower performance	.63	.69	-.06	-.13	-.08	-.17	.29
Higher performance	.74	.75	-.01	-.02	-.01	-.02	.81
Male	.67	.72	-.05	-.11	-.06	-.14	.27
Female	.74	.74	-.00	-.01	-.00	-.01	.95
4-8	.73	.76	-.03	-.06	-.03	-.07	.44
9-12	.55	.57	-.02	-.04	-.03	-.06	.76
Cohort 2	.73	.78	-.05	-.11	-.05	-.13	.24
Cohort 1	.59	.56	.03	.05	.04	.07	.63

NOTES: Means are regression adjusted using a consistent set of baseline covariates. Impact estimates are reported as marginal effects. Effect sizes are in terms of standard deviations. Valid N = 1,014. Survey given to students in grades 4-12. Student survey weights used. Robust regression calculations generated by clustering at the family level.

The Impact of the Program on Intermediate Outcomes

Understanding the mechanisms through which the OSP does or does not affect student outcomes requires examining the expectations, experiences, and educational environments made possible by Program participation. The analysis here estimates the impact of the Program on a set of "intermediate outcomes" that may be influenced by parents' choice of whether to use an OSP scholarship and where to use it, but are not end outcomes themselves. The method used to estimate the impacts on intermediate outcomes is identical to that used to estimate impacts on the key Program outcomes, such as academic achievement.

Prior to data analysis, possible intermediate outcomes of the OSP were selected based on existing research and theory regarding scholarship programs and educational achievement. Because 24 intermediate outcome candidates were identified through this process, the variables were organized into four conceptual groups or clusters, as described below, to aid in the analysis.

There is no way rigorously to evaluate the linkages between the intermediate outcomes and achievement—students are not randomly assigned to the experience of various educational conditions and

programs. That is why any findings from this element of the study do not suggest that we have learned what specific factors "caused" any observed test score impacts, only that certain factors emerge from the analysis as possible candidates for mediating influence because the Program affected students' experience of these factors. The analyses are exploratory, and, given the number of factors analyzed, some of the statistically significant findings may be "false discoveries" (due to chance).

Overall, 3 years after applying to the Program, the offer of an Opportunity Scholarship appears to have had an impact on 8 of the 24 intermediate outcomes examined, 7 of which remained statistically significant after adjustments for multiple comparisons:

- **Home Educational Supports.** Of the four intermediate outcomes in this category, the offer of a scholarship had an impact on one of them. There was a significant negative impact on tutor usage outside of school, and this impact remained statistically significant after adjustments for multiple comparisons. There were no statistically significant differences between the treatment and control groups on parents' reports of their involvement in school in year 3, parents' aspirations for how far in school their children would go, or time required for the student to get to school.

- **Student Motivation and Engagement.** Of the six intermediate outcomes in this category, the offer of a scholarship may have had an impact on one of them. Based on student surveys, the offer of a scholarship seems to have had a significant negative impact on whether students read for fun. Adjustments for multiple comparisons, however, indicate that this result could be a false discovery, so it should be interpreted with caution. There were no statistically significant differences between the treatment and control groups in their reported aspirations for future schooling, engagement in extracurricular activities, and frequency of doing homework or in their parents' reports of student attendance or tardiness rates.

- **Instructional Characteristics.** The offer of a scholarship had a statistically significant impact on 5 of the 10 intermediate outcomes in this group of indicators. Students offered a scholarship experienced a lower likelihood that their school offered tutoring, special programs for children who were English language learners, or special programs for students with learning problems compared to control group students; these impacts remained statistically significant after adjustments for multiple comparisons. Students offered a scholarship experienced a higher likelihood that their school offered programs for advanced learners and such enrichment programs as art, music, and foreign language; these two impact estimates also remained statistically significant after adjustments for multiple comparisons. There were no significant differences between the treatment and control groups in student/teacher ratio, how students rated their teacher's attitude, the school's use of ability grouping, in-school tutor usage, or the availability of before and after school programs.

- **School Environment.** The offer of a scholarship affected one of four measures of school environment. Students offered a scholarship experienced schools that were smaller by an average of 182 students than the schools attended by students in the control group; this impact remained statistically significant after adjustments for multiple comparisons. There were no statistically significant differences between the

treatment and control groups, on average, in school reports of parent/school communication practices, the percentage of minority students at the school, or the classroom behavior of peers based on student reports.

Conclusions

This paper presents results at an intermediate point in a longitudinal experimental evaluation of a school voucher program. Three years after random assignment to the treatment (offer of a scholarship) or control (not offered) group, some statistically significant programmatic impacts have emerged. On average, students are performing higher in reading if they were offered an Opportunity Scholarship. However, the impact of the OSP on student test scores in math is not significantly different from zero. Parents are much more satisfied with their children's school and view it as safer as a result of the scholarship offer. Additionally, students attended smaller schools and experienced a higher likelihood that their school offered programs for advanced learners and such enrichment programs as art, music, and foreign language if they were offered an Opportunity Scholarship.

The results of this analysis of a school choice program emerged from an evaluation structured as a Randomized Control Trial (RCT). Because mere chance determined whether eligible applicants received the treatment of a scholarship offer or a place in the control group, any subsequent differences between the outcomes of the two groups that are statistically significant can be attributed to the program intervention.

Although RCTs are widely viewed as the most rigorous research designs for determining outcome differences caused by programs, they can face certain threats to validity. The two greatest threats to the validity of the impact estimates presented here are non-response bias and inadequate study power. As discussed in the section of Data Collection, the year 3 effective test-score response rates were 69 percent for the entire impact sample and the separate rates for the treatment and control group were almost identical. Non-response bias is possible in any longitudinal study that fails to obtain outcome data from 100 percent of the participants. In order for non-response bias to explain the test-score impacts of the OSP that we report here, the 30 percent of control group students who did not respond to data collection would have had to differ markedly in their reading achievement from the 32 percent of treatment group students who likewise did not respond to data collection. Although such a large non-response bias in such a small portion of the impact sample seems unlikely, to further guard against such an eventuality we weighted the observations from the students who did respond to data collection so that the outcome sample, in the aggregate, shares the same baseline characteristics as the original sample of randomized students. As a result of these methodological steps, it is our opinion that non-response bias is unlikely to be an explanation for the test-score impacts that we report here.

At the beginning of the evaluation, a power analysis was conducted to determine if the impact sample was sufficiently large to detect test-score impacts of a reasonable magnitude if such impacts were actually produced by the OSP. That power analysis forecasted that the analysis would be able to detect year 3 test score impacts that were .12 standard deviations or larger using the overall sample, and impacts that were between .14 and .38 standard deviations or larger at the subgroup level, depending on the size of each subgroup (Wolf et al, 2009, p. A-5). Although the practical importance of educational impacts of certain magnitudes is inherently subjective and debatable, in our opinion, the ability to detect an impact as small as .12 standard deviations is uncommon in rigorous education evaluations. The wide range in the size of the minimum detectable effects for the test-score analysis at the subgroup level is yet another reason for caution in interpreting those disaggregated results and for a greater emphasis on the impacts for the overall sample when judging the results of the OSP.

The power analysis that was conducted prior to year 3 data analysis proved to be highly reliable. As a result of the impact analysis, a treatment-group advantage in reading achievement of .15 standard deviations was found to be statistically significant whereas a treatment-group advantage of just .03 standard deviations in math was found not to be statistically significant. All of the subgroup impacts in both reading and math that were found to be statistically significant as a result of the actual analysis were larger than the minimum level necessary for detection as estimated by the power analysis, and all of the subgroup test-score differences not found to be statistically significant were smaller than the minimum level. The analysis might have failed to identify some program impacts at the subgroup level, due to a lack of analysis power, but this impact evaluation had the power to be quite precise, especially when estimating overall program impacts, and the actual results of the statistical analysis mapped very closely the forecasts of the preliminary power analysis.

It is important to note that the findings regarding the impacts of the OSP reflect the particular Program elements that evolved from the law passed by Congress and the characteristics of students, families, and schools—public and private—that exist in the Nation's capital. The same program implemented in another city could yield different results, and a different scholarship program in Washington, DC, might also produce different outcomes. Evaluation of the OSP continues, with the year 4 impact report expected for release in 2010.

Sources Cited

Bloom, Howard S. "Accounting for No-Shows in Experimental Evaluation Designs." *Evaluation Review* 1984, 8(2): 225-246.

Boruch, Robert, Dorothy de Moya, and Brooke Snyder. "The Importance of Randomized Field Trials in Education and Related Areas." *Evidence Matters: Randomized Trials in Education Research*, Frederick Mosteller and Robert Boruch, editors. Washington, DC: The Brookings Institution Press, 2002.

Greene, Jay P. "Vouchers in Charlotte." *Education Matters* 2001, 1(2): 55-60.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. "Disruption Versus Tiebout Improvement: The Costs and Benefits of Switching Schools." *Journal of Public Economics* 2004, 88: 1721-1746.

Howell, William G., and Paul E. Peterson, with Patrick J. Wolf and David E. Campbell. *The Educational Gap: Vouchers and Urban Schools.* Revised Edition, Washington, DC: The Brookings Institution Press, 2006.

Howell, William G., Patrick J. Wolf, David E. Campbell, and Paul E. Peterson. "School Vouchers and Academic Performance: Results from Three Randomized Field Trials." *Journal of Policy Analysis and Management* 2002, 21(2): 191-217.

Johnson, Michael D., and Claes Fornell. "A Framework for Comparing Customer Satisfaction Across Individuals and Product Categories." *Journal of Economic Psychology* 1991, 12(2): 267-286.

Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz, "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment." *Quarterly Journal of Economics* 2005, 120(1): 87-130.

Mayer, Daniel P., Paul E. Peterson, David E. Myers, Christina Clark Tuttle, and William G. Howell. *School Choice in New York City After Three Years: An Evaluation of the School Choice Scholarships Program.* MPR Reference No. 8404-045. Cambridge, MA: Mathematica Policy Research, 2002.

Rouse, Cecilia Elena. "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program." *Quarterly Journal of Economics* 1998, 113(2): 553-602.

Sanbonmatsu, Lisa, Jeffrey R. Kling, Greg J. Duncan, and Jeanne Brooks-Gunn. "Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment." *Journal of Human Resources* 2006, 41(4): 649-691.

Schochet, Peter Z. *Guidelines for Multiple Testing in Experimental Evaluations of Educational Interventions.* MPR Reference No. 6300-080. Cambridge, MA: Mathematica Policy Research, 2008.

Stewart, Thomas, Patrick J. Wolf, and Stephen Q. Cornman. *Parent and Student Voices on the First Year of the DC Opportunity Scholarship Program.* SCDP Report 05-01. Washington, DC: School Choice Demonstration Project, Georgetown University, 2005. Available online at [http://www.georgetown.edu/research/scdp/PSV-FirstYear.html].

Witte, John F. *The Market Approach to Education: An Analysis of America's First Voucher Program.* Princeton, NJ: Princeton University Press, 2000.

Witte, John F., Patrick J. Wolf, Joshua M. Cowen, David J. Fleming, and Juanita Lucas-McLean, *MPCP Longitudinal Educational Growth Study Second Year Report*, Report of the School Choice Demonstration Project, University of Arkansas, Fayetteville, AR, March 2009, Milwaukee Evaluation Report #10, pp 40, available at http://www.uark.edu/ua/der/SCDP/Milwaukee_Eval/Report_10.pdf.

Wolf, Patrick, Babette Gutmann, Nada Eissa, Michael Puma, and Marsha Silverberg. *Evaluation of the DC Opportunity Scholarship Program: First Year Report on Participation.* U.S. Department of Education, National Center for Education Evaluation and Regional Assistance. Washington, DC: U.S. Government Printing Office, 2005. Available online at [http://ies.ed.gov/ncee/].

Wolf, Patrick, Babette Gutmann, Michael Puma, and Marsha Silverberg. *Evaluation of the DC Opportunity Scholarship Program: Second Year Report on Participation.* U.S. Department of Education, National Center for Education Evaluation and Regional Assistance, NCEE 2006-4003. Washington, DC: U.S. Government Printing Office, 2006. Available online at [http://ies.ed.gov/ncee/].

Wolf, Patrick, Babette Gutmann, Michael Puma, Lou Rizzo, Nada Eissa, and Marsha Silverberg. *Evaluation of the DC Opportunity Scholarship Program: Impacts After One Year.* U.S. Department of Education, National Center for Education Evaluation and Regional Assistance, NCEE 2007-4009. Washington, DC: U.S. Government Printing Office, 2007. Available online at [http://ies.ed.gov/ncee/].

Wolf, Patrick, Babette Gutmann, Michael Puma, Brian Kisida, Lou Rizzo, and Nada Eissa. *Evaluation of the DC Opportunity Scholarship Program: Impacts After Two Years*. U.S. Department of Education, National Center for Education Evaluation and Regional Assistance, NCEE 2008-4023. Washington, DC: U.S. Government Printing Office, 2008. Available online at [http://ies.ed.gov/ncee/].

Wolf, Patrick, Babette Gutmann, Michael Puma, Brian Kisida, Lou Rizzo, and Nada Eissa. *Evaluation of the DC Opportunity Scholarship Program: Impacts After Three Years*. U.S. Department of Education, National Center for Education Evaluation and Regional Assistance, NCEE 2009-4050. Washington, DC: U.S. Government Printing Office, 2009. Available online at [http://ies.ed.gov/ncee/].