

An Analysis of the Effects of North Carolina’s Opportunity Scholarship Program on Student Achievement

Anna J. Egalite
D. T. Stallings
Stephen R. Porter

North Carolina State University, Raleigh, NC

The North Carolina Opportunity Scholarship Program is a private school voucher program that provides state-funded vouchers worth up to \$4,200 to eligible students entering kindergarten through 12th grade. Because the public and private school sectors administer different assessments, we recruited approximately 700 students to take a common, nationally normed, standardized test. Matching on baseline achievement and rich demographic data, we use a quasi-experimental inverse propensity weighting approach to maximize comparability between the public and private school student samples. Our preferred specification examines first-year effects for new Opportunity Scholarship students, revealing positive estimates of .36 SD in math and .44 SD in language; there is no effect on reading scores. Results for renewal students are statistically significant in language scores only. In further analyses, we estimate separate effects for private schools that regularly administer another version of the assessment used in this study, the Iowa Test of Basic Skills. We conclude by discussing policy implications.

Keywords: *school choice, education policy, vouchers, academic achievement, private schooling*

LAUNCHED in the 2014–2015 school year, the North Carolina Opportunity Scholarship Program provides state-funded vouchers worth up to \$4,200 per year for eligible students in kindergarten through 12th grade to attend participating private schools. Since its inception, nearly \$53 million dollars have been disbursed in Opportunity Scholarship funds. This analysis focuses on identifying any changes in the math, reading, and language achievement of students attending a subset of participating private schools with the assistance of the North Carolina Opportunity Scholarship Program. Although the legislation that established the program calls for an evaluation of students’ learning gains or losses, no such study has been conducted until now. Two primary challenges have inhibited an analysis of program effects to date. First, the accountability requirements associated with the program require participating private schools to administer and submit scores from any nationally normed standardized test of their choosing, whereas comparable public school students take the criterion-referenced state test, the North Carolina End of Grade (EOG) exam. Thus, participating students do not have scores on a common metric that could be pooled for comparison to students in the public school sector. Second, although the establishing legislation calls for an evaluation of the program, it does not incentivize applicants to participate in

one, which introduces recruitment challenges with regard to the collection of original test score data.

In the spring of 2017, we traveled to the four regions of North Carolina with the highest number of Opportunity Scholarship users. Relying on a volunteer sample of students in the public and private school sectors in those regions, we administered Form E of the Iowa Test of Basic Skills (ITBS)—a nationally normed, standardized assessment of math and reading achievement. Using inverse propensity weighting (IPW) in a regression model to compare similar public school students to students attending private schools with an Opportunity Scholarship, we generate estimates of changes in math, reading, and language scores associated with participating in the Opportunity Scholarship Program. The overall effect sizes observed are positive, large, and statistically significant in the aggregate, ranging from .25 to .49 standard deviations (henceforth *SDs*); however, there is important heterogeneity beneath the surface of these analyses. Specifically, an analysis comparing schools that regularly use ITBS tests (which are majority Catholic schools) to those that regularly use an alternative assessment regime (e.g., the Terra Nova or Stanford assessments) reveals no differences in the academic achievement of voucher students in non-ITBS-using private schools. This suggests that the positive coefficients observed for ITBS-using private schools



could be driving the overall results, which may have important policy implications.

Background

Description of North Carolina's Voucher Program

North Carolina's Opportunity Scholarship Program joins a wave of recently enacted private school choice programs that feature relatively broad eligibility criteria and expand access to an entire state, unlike earlier programs, which have been limited in scope to a single city (e.g., Cleveland, OH; Milwaukee, WI; and Washington, D.C.). Students participating in North Carolina's voucher program are eligible to receive a private school voucher worth up to \$4,200 per year for tuition and fees for books, transportation, or school equipment. Eligibility is determined by a two-part test. First, the student must meet at least one of the following initial criteria: Students must have been enrolled in a North Carolina public school during the previous semester; students must be entering kindergarten or first grade; students must have received an Opportunity Scholarship for the previous semester; students must be in foster care or recently adopted; or students must have a parent or legal guardian who is on full-time active duty in the military. Second, students must reside in a household with an income level that does not exceed 133% of the amount required for the federal free or reduced-price lunch program.

The most recent data indicate that program participation continues to expand annually. The Opportunity Scholarship Program received 12,553 new applications for the 2019–2020 school year, of which 8,959 were deemed eligible for assistance. Even though there is sufficient funding to award vouchers to all eligible applicants, for various reasons, not all eligible students end up enrolling in the program (Egalite et al., 2017). Ultimately, 4,511 new students enrolled in the most recent school year and 7,498 students renewed last year's voucher, bringing the total number of recipients in 2019–2020 to 12,009 students. Of this group, 13% identify as Hispanic, 28% are Black, 52% are White, and the rest are Asian, American Indian or Alaskan Native, Native Hawaiian or other Pacific Islander, two or more races, or chose not to identify a race.

Recipients come from every county in North Carolina, although the greatest number of recipients come from Cumberland County ($n = 1,349$ in 2019–2020), followed by Mecklenburg County ($n = 943$) and Wake County ($n = 932$). The private school educating the greatest number of Opportunity Scholarship recipients is Trinity Christian School of Fayetteville (Cumberland County), which received \$640,500 in disbursements in 2019–2020. Full-price tuition at this school was \$5,000 per student in 2018–2019, the most recent year of available data.

To better understand program accessibility, it is helpful to know more about the size of the voucher relative to the

average private school tuition. Unfortunately, the state agency that oversees private schools in North Carolina—the Division of Non-Public Education—does not collect or publicly report these data. We worked instead with the State Education Assistance Authority (SEAA; the state agency that disburses voucher funds) to access, digitize, and analyze the tuition information that was submitted voluntarily by those private schools that participate in the Opportunity Scholarship Program. We believe this is the most comprehensive database of private school tuition that has been assembled to date in North Carolina, even though it suffers from nontrivial missing data. We successfully manually extracted the 2016–2017 tuition data for 374 private schools. The median tuition charged was \$5,483. The minimum value was \$2,025 and the maximum value was \$27,500.

Theoretical Framework

Proposals to subsidize private education can be traced back to John Stuart Mill (1869), who argued that compulsory education should be compelled by the state but not provided by that entity. To avoid the inevitable conflicts that would arise over what should be taught and how, Mill proposed breaking the financing of education from its provision. Milton Friedman (1955, 1962) added an economic dimension to the argument for school choice by suggesting that the competition resulting from this arrangement would raise the quality of education across the board. In the 1980s, James Coleman added a sociological dimension to the argument for school choice by pointing to the success of Catholic schools for urban students in particular (Coleman & Hoffer, 1987). He theorized that the “social capital” present in these communities was driving differences in student achievement. John Chubb and Terry Moe (1990) added a political dimension to the theoretical framework with the publication of *Politics, Markets, & America's Schools*. They reasoned that the American public school system had become excessively bureaucratic and political, which is particularly problematic when it comes to teaching and learning, as a successful education system requires autonomy and flexibility at the point of service delivery. A system of choice would be more efficient, they argued, and the most promising strategy to improve school quality. In more recent years, attention to the opportunity disparities between students from different economic backgrounds have inspired a social justice argument for school choice (Fuller & Page, 2014), the exercise of which allows for self-determination among underprivileged families who often do not otherwise enjoy the wide range of choices exercised by more affluent families.

Across many dimensions, North Carolina's program presents promise for testing these theories. Private school participation is high, with 405 of the state's 769 private schools (53%) enrolling voucher students in 2018–2019.

Program regulations are light, with private schools applying their own admissions criteria and administering a standardized test of their choosing. Participating private schools are diverse, at least in terms of what we can observe in administrative data, which corresponds with Chubb and Moe's (1990) assertion that a successful education program must be personalized to suit a particular context. In 2017–2018, for example, 30% of participating private schools were Christian, 22% nonsectarian, 22% Baptist, and 10% Roman Catholic, with the rest representing a diverse mixture of faith backgrounds, including Islamic and Jewish. This brings to mind Coleman and Hoffer's (1987) argument that social capital forms in close communities, such as religious communities of shared values. We also see great diversity in the standardized assessments schools rely upon, which may signal differences in curricula and the overall course of study. These tests include the ITBS, the Terra Nova, the Stanford Achievement Test, the Woodcock Johnson, and the North West Evaluation Association's Measures of Academic Progress. Finally, North Carolina's Opportunity Scholarship Program targets disadvantaged families by design, thus expanding opportunity to lower income families across the state, an outcome promoted by Fuller and Page (2014).

Literature Review

Since the establishment of the nation's first private school choice program in Milwaukee in 1990, researchers have conducted numerous lottery-based studies to quantify the achievement impact of these programs, which have been both publicly funded (e.g., the voucher programs in Milwaukee, WI; Cleveland, OH; and Washington D.C.) and privately funded (e.g., the New York City School Choice Scholarships Foundation Program and the Charlotte Scholarship Fund). In general, the majority of the 17 lottery-based studies conducted to date have revealed small positive impacts overall, with larger positive impacts observed for subgroups of interest, such as African American students, students of low socioeconomic status, or students who initially were low-achieving (Table 1). A review of the attainment literature is provided in a supplemental online appendix.

Specifically, among the 17 random-assignment studies detailed in Table 1, 10 have revealed positive impacts in math or reading overall or for at least one subgroup of interest, five studies revealed null impacts, and two studies revealed negative impacts for the overall sample or for at least one subgroup. Although the weight of the experimental evidence points toward a positive impact associated with voucher usage, two recent lottery-based studies of the Louisiana Scholarship Program reveal sizable negative impacts.

Abdulkadiroglu et al. (2018) leveraged Louisiana's over-subscribed application lottery to examine experimentally the impact of the Louisiana Scholarship Program in the first year

after its statewide expansion in 2012, finding large and statistically significant negative impacts across all subjects examined. Using a voucher to attend a private school reduced student achievement in math by 41% of an *SD*, with reductions of 8%, 26%, and 33% in reading, science, and social studies achievement, respectively. Fourth-year impact estimates by Mills and Wolf (2019) reveal similarly large and statistically significant negative impacts. Specifically, voucher users scored 28% of an *SD* lower in math and 31% lower in science after 4 years in the program. Many have questioned what differs about the Louisiana context that might explain these negative impacts. Abdulkadiroglu et al. (2018) raise questions about the quality of private schools accepting voucher students in Louisiana, as participating schools appear to have been experiencing declining enrollment prior to the passage of the voucher law. Further evaluation work will be necessary to determine how the quality of participating private schools is evolving as the program ages and as regulations to address private school quality are implemented (White, 2016).

Two other evaluations of statewide voucher programs have been conducted in recent years that do not rely on a lottery-based design but employ alternative approaches to causal inference that attempt to approximate a gold-standard design. First, Figlio and Karbownik (2016) employ a propensity score matching approach to evaluate Ohio's EdChoice Scholarship program. Using student records from 2003–2004 to 2012–2013, Figlio and Karbownik report substantial negative program impacts in both math and English language arts, with the largest negative effects observed in math. The authors note that these negative effects cannot be explained by the transition effects of switching to a new private school because they persist over time.

Second, Waddington and Berends (2018) employ a difference-in-differences model with propensity score matched students to evaluate achievement changes for students in Grades 2 through 8 who participate in the Indiana Choice Scholarship Program. After 1 year in the program, there are no differences in English language arts outcomes but math scores decline by 15% of an *SD*, on average—a negative effect that persists across all four years of the study.

In summary, the existing body of evidence on the impact of vouchers is growing both in sophistication and in scope, but it remains incomplete. As a result, much remains unknown about the contemporary school choice context. Crucially, most of the prior evaluations of private school voucher programs have been limited to cities, but recent evaluations of the relatively young, statewide voucher programs in Louisiana, Ohio, and Indiana point to potentially substantial negative test score impacts associated with initial voucher use, contrary to a consistent body of evidence prior to 2017 that showed null to positive impacts associated with voucher usage in the United States. A high-quality analysis of North Carolina's Opportunity Scholarship Program, a

TABLE 1
Studies of the Achievement Impacts of School Vouchers

Citation	Program Name	Program Funding	Program Scope	Impact
Lottery-based studies				
Cowen (2008)	Charlotte Scholarship Fund	Private	Citywide	Positive Overall
Howell et al. (2002)	The D.C. Opportunity Scholarship Program	Public	Citywide	Positive Overall
Greene (2001)	Charlotte Scholarship Fund	Private	Citywide	Positive Overall
Greene et al. (1999)	Milwaukee Parental Choice Program	Public	Citywide	Positive Overall
Rouse (1998)	Milwaukee Parental Choice Program	Public	Citywide	Positive Overall
Kisida & Wolf (2015)	The D.C. Opportunity Scholarship Program	Public	Citywide	Positive Overall
Jin et al. (2010)	New York City School Choice Scholarships Foundation Program	Private	Citywide	Positive for Subgroups
Barnard et al. (2003)	New York City School Choice Scholarships Foundation Program	Private	Citywide	Positive for Subgroups
Howell et al. (2002)	The Dayton, Ohio Parents Advancing Choice in Education Program	Private	Citywide	Positive for Subgroups
Howell et al. (2002)	New York City School Choice Scholarships Foundation Program	Private	Citywide	Positive for Subgroups
Webber et al. (2019)	The D.C. Opportunity Scholarship Program	Public	Citywide	Null
Bitler et al. (2014)	New York City School Choice Scholarships Foundation Program	Private	Citywide	Null
Wolf et al. (2013)	The D.C. Opportunity Scholarship Program	Public	Citywide	Null
Bettinger & Slonim (2006)	The Children's Scholarship Fund of Toledo, Ohio	Private	Citywide	Null
Krueger & Zhu (2004)	New York City School Choice Scholarships Foundation Program	Private	Citywide	Null
Mills & Wolf (2017)	Louisiana Scholarship Program	Public	Statewide	Negative Overall and for Subgroups
Abdulkadiroglu et al. (2018)	Louisiana Scholarship Program	Public	Statewide	Negative Overall and for Subgroups
Recent quasi-experimental studies				
Waddington & Berends (2018)	Indiana Choice Scholarship Program	Public	Statewide	Negative Overall
Figlio & Karbownik (2016)	Ohio EdChoice Scholarship Program	Public	Statewide	Negative Overall

Note. The designations “Positive Overall” and “Negative Overall” indicate that a statistically significant positive or negative achievement impact was observed for the overall sample in at least one subject.

close cousin of these three state-level programs, provides valuable new data on the effect of contemporary private school choice programs.

Prior Evaluations of North Carolina’s Opportunity Scholarship Program

The 2013 legislation that established North Carolina’s Opportunity Scholarship Program includes an evaluation mandate, which is described in a supplemental online appendix. Nevertheless, the only evaluation of the academic achievement of participants in this program is a report by the Children’s Law Clinic at Duke Law School (Wettach, 2017). Instead, the author presents the percentage of voucher students scoring above average on any of the standardized tests given in a subset of private schools (schools with greater than 25 voucher enrollees—just 12% of the total population of voucher-accepting schools in North Carolina in 2015–2016) and compares this percentage to the average National Assessment of Educational Progress performance of all low-income students in the public schools. Vastly different sample sizes in the two groups under comparison raise questions about uneven measurement error in the two sectors being compared. Furthermore, students in these two groups differ in terms of gender, race, age, family background, prior academic achievement, and numerous other background characteristics that are not accounted for with a statistical model. This is an important omission because descriptive data on program participants demonstrate that voucher students represent a distinct subgroup of students in the state (Egalite et al., 2017). For example, voucher students reside in some of the lowest income households in North Carolina, with an adjusted median household income in 2016–2017 of \$16,213 for new voucher recipients.¹ In what follows, we describe the methodology we rely upon to address many of these shortcomings.

Methodology

Our analytic approach requires academic performance data for comparable groups of voucher and public school students. We first collected outcome data from voucher students in their first or second year of voucher use and low-income public school students who agreed to take the ITBS, our measure of academic performance. We then excluded public school students from the analysis who did not qualify for free and reduced-price meals. Because the Opportunity Scholarship Program uses eligibility guidelines similar to those of the federal free and reduced-price meals program, our exclusion helps us to create relatively comparable groups of students in terms of their economic backgrounds (explained in more detail below). We then use IPW to create comparable groups within this sample of test takers, using state standardized tests from the prior year

and demographic characteristics in the model used to generate the propensities.

Student Recruitment

In spring 2017, we collaborated with the leadership of public schools, private schools, and partner organizations in four geographic regions of North Carolina to recruit a volunteer sample of low-income students in Grades 4 through 8 who either were awarded a voucher or likely would have qualified for the Opportunity Scholarship Program.

Recruitment on the public school side adhered to the following protocol: We first applied for separate institutional review board approval in each participating school district, as well as at the university level, then commenced a four-stage approach to recruiting unpaid student volunteers to participate in data collection. Starting at the district level, we approached the public school superintendents of those geographic regions that featured high numbers of Opportunity Scholarship students to request permission to recruit volunteers in their schools to serve as our comparison group. Once a cooperation agreement was established with four public school superintendents, we coordinated with the head of testing in each district to identify the highest poverty public schools in those districts. We next reached out to individual school principals to request their cooperation with the study. Once principals consented, school personnel distributed explanatory letters to families on our behalf, describing the goals of the research project and requesting parental consent for student participation in the research project. Finally, on the day of testing, students were given a children’s version of the parental information letter, written in simple language that explained the purpose of the research and any expected risks or benefits, and notifying them that their participation was voluntary and they were free to opt out at any time.

Recruitment procedures on the private school side adhered to the following protocol: We started by making courtesy calls to the private school associations in the state, notifying them about the project so they would have information for answering questions from private school leaders in their networks. We then scheduled one-on-one recruitment phone calls with individual school principals and meetings with both policy-supportive and policy-opposed advocacy groups and state legislators. The bulk of recruitment support ultimately was provided by one policy-supportive organization, Parents for Educational Freedom in North Carolina, which encouraged private schools to respond to our request to participate in the study and helped schools schedule data collection sessions with us. Parent and student consent procedures followed the protocols already described and were identical for private and public school students.

All test administration was conducted by our research team and by third-party, independent research contractors—unaffiliated with any research or advocacy organizations—on school grounds during the school day. The

TABLE 2
Constructing the Analysis Sample

Description	Private	Public	Total
Took ITBS in 2017	297	401	698
Not included in analysis because:			
No unique match with DPI data	48	76	124
Took wrong version of ITBS in 2017	0	8	8
Did not take/finish ITBS in 2017	0	7	7
Missing 2016 EOG	1	0	1
Not free/reduced-price lunch in 2016	—	58	58
Could not verify OS receipt	3	—	3
Final student count	245	252	497
Status of OS recipients			
New	90	—	—
Renewal	155	—	—
Analytic data set <i>n</i>	245	252	497

Note. ITBS = Iowa Test of Basic Skills; DPI = Department of Public Instruction; EOG = North Carolina End of Grade assessments (our measure of baseline achievement); OS = Opportunity Scholarship.

test administrators all were trained in common protocols and ethical considerations to ensure that the testing conditions were consistent across sites. The team arrived at each school site at the start of the school day and coordinated with school personnel to set up the room with testing booklets, scratch paper, and pencils. Parental consent letters were collected prior to our arrival by school personnel. In addition, students were provided with an assent form on the day that further allowed them to opt out of testing, if necessary. Testing sessions took approximately 2 hours in total. Unfortunately, students with special educational needs were excluded from all data collection efforts as the researchers did not have the manpower to provide special testing accommodations as needed.

Analytic Sample

Students. In total, there were 698 low-income students tested in spring 2017: 297 in private schools and 401 in public schools (Table 2). Student answer sheets were machine-scored by Houghton Mifflin Harcourt and the scores were returned to us in a digital format. To construct the final analysis sample, we merged these records with student data from the North Carolina Department of Public Instruction (DPI), which allows us to add rich demographic data and prior test scores to each student record. Because the voucher students tested were in their first or second year of voucher usage, and because the vast majority were enrolled in a public school before their enrollment in a private school, the DPI data included rich baseline information about both the public and private school students in our sample (with some exceptions, detailed below). We merged records by student name, gender, race, and grade level, and 124 observations were

dropped at this stage, either because a reliable match could not be made with DPI data (e.g., the student reported a nickname instead of their legal name on the ITBS answer sheet) or because a name matched with multiple records for that grade level in the DPI data (e.g., in the case of students with very common first and last names), creating uncertainty as to the true match. In both of these scenarios, we opted for the conservative approach of dropping records when we could not be sure of an exact match. Eight more students were dropped because they were mistakenly given the wrong test for their grade level in 2017 (i.e., some fourth graders took the fifth grade test and vice versa). An additional seven students opted to not complete the tests on testing day and were dropped from the sample as a result.

The Opportunity Scholarship Program uses the same income eligibility guidelines that the federal government uses to determine eligibility for free and reduced-price meals, with Opportunity Scholarship Program thresholds set to match the reduced-price thresholds for applicants seeking a full voucher; for applicants seeking a 90% voucher, income must not exceed 133% of the reduced-price threshold. In 2016–2017, for example, the household gross income for a family of four could not exceed \$44,955 for full tuition and \$59,7908 for 90% tuition.

One concern is that the 133% threshold could result in voucher students having higher family incomes, on average, than public school students eligible for free and reduced-price meals. We do not have reported income data for the voucher students in our analytic sample to investigate this possibility. However, we do have access to an anonymized data set from SEAA that contains household income reported on the Opportunity Scholarship Program application form for all 2016–2017 voucher recipients. Only 19% of voucher

recipients received the 90% voucher, indicating their household income marginally exceeds the free and reduced-price meal thresholds. Although the Opportunity Scholarship Program technically allows applicants whose income is 133% of the free and reduced-price meals threshold, very few families in the application pool actually meet this threshold. Specifically, the average percentage over the threshold for the 2016–2017 applicants was just 114%. This analysis suggests that the household incomes for the two groups of students in our testing sample are indeed highly similar.

To further ensure comparability between Opportunity Scholarship students and students in public schools, 58 public school students who were not listed as free and reduced-price lunch–eligible in the 2016 DPI data were removed from the analysis, along with three students who could not be verified in the Opportunity Scholarship Program records. These final two screens ensure that every student in both public and private schools in our sample was identified as eligible for free or reduced-price lunch. The final student count in the analytic data set is 497 students, with 245 in private schools and 252 in public schools.

Schools. In total, we collected test scores from volunteer students in 24 private schools and 14 public schools (Table 3). Sample sizes were larger in the public schools, with a mean of 35 tested students per site and a median of 23 students. On the private school side, there was a mean of 12 tested students per site and a median of seven students. We visited every private school that agreed to be in the study, including a school with just one student volunteer and another with 43 students. Data for 19 private schools were located in federal data files, and data for 16 schools were located in state records from the Division of Non-Public Education. The reader should note we were unable to locate descriptive data for three private schools in the testing sample, so those schools are excluded from the descriptive data presented in Table 3 and the discussion presented in the next section.

Approximately half (53%) of the private schools we visited were Catholic schools, 26% were Christian (no specific denomination), 16% were Baptist, and 5% were some other religion such as Methodist or Episcopal. For reference, Columns 3 and 4 in Table 3 present descriptive statistics for all private schools in the state and for all voucher-accepting private schools. Of note, Catholic schools represent just 8% of all private schools, statewide, and 12% of all voucher-accepting private schools, indicating that they are overrepresented in our sample, which has implications for external but not internal validity.

The private schools in our sample have a median enrollment of 225 students, with a minimum enrollment of 107 and a maximum of 1,402. These are larger than average values, as the median enrollment for private schools, statewide, is 77 students and for voucher-accepting private schools, it is 120 students. The majority of the private schools in our

sample (90%) are located in either a city (74%) or a suburb (16%). For comparison, the proportion of all private schools across the state located in a city or suburb is 75% and the proportion of all voucher-accepting private schools located in a city or suburb is 79%. The private schools in our sample are more diverse than private schools across the state, with higher percentages of school percentage Black (15%) and Hispanic (13%) in our sample, relative to the state as a whole, where Black students represent 14% of private school enrollment and Hispanic students just five%.

The median pupil:teacher ratio for private schools in our sample is 12 and the median number of full-time equivalent teachers for private schools in our sample is 26. All private schools are required to report which standardized tests they use for accountability purposes to the North Carolina Division of Non-Public Education, so we also report this information for the schools in our sample. The most commonly used test is the ITBS (used by 44% of the private schools in our sample), followed by the Terra Nova (31%).

Data for all 14 of the public schools in our sample were located in federal files, and descriptive statistics are presented in Column 2 of Table 3. For comparison, we also present descriptive statistics for all public schools in the state that lost a student to the Opportunity Scholarship Program in 2015–2016, which we term the “Sending public schools” (Column 5). The reader should note that five of the 141 sending public schools are included in our analysis sample, as a result of our efforts to recruit volunteers from the most relevant public schools.

All of the public schools in our sample are traditional public (93%) or magnet schools (7%) and all qualified for schoolwide Title 1, a common proxy for high-poverty schools. This latter selection criterion was by design, to maximize the chance of recruiting comparable low-income public school students. The public schools were larger than the private schools we visited: The median enrollment at private schools in our sample was 225, compared to 649 in the public schools in our sample. Although half of the public schools in our sample are coded as being located in a rural area, they were all located in the same geographic regions of the state as the private schools visited. This difference in urban locale may also reflect differences in how the Private School Universe Survey and the Public Elementary/Secondary School Universe Survey code urbanity (i.e., the private school survey uses the categories city, suburb, town, rural; whereas the public school survey uses the categories city, large; city, mid-size; city, small; suburb, large; suburb, midsize; suburb, small; town, fringe; town, distant; town, remote; rural, fringe; rural, distant; rural, remote). Finally, the median public school pupil:teacher ratio is 14, and the median number of full-time equivalent teachers is 45. All of the public schools administer the North Carolina EOG tests—the standardized math, English language arts, and science tests given to all North Carolina public school students in Grades 3 through 8.

TABLE 3
Description of Public and Private Schools in Sample

Description	(1) Private Schools in Sample	(2) Public Schools in Sample	(3) All North Carolina Private Schools	(4) All Voucher-Accepting Private Schools	(5) Sending Public Schools
No. of students tested per school for this study					
<i>Mdn</i>	7	23	n/a	n/a	n/a
<i>M</i>	12	35	n/a	n/a	n/a
Minimum	1	4	n/a	n/a	n/a
Maximum	43	166	n/a	n/a	n/a
Religious affiliation					
Catholic	0.53	0.00	0.08	0.12	0.00
Baptist	0.16	0.00	0.20	0.23	0.00
Christian (no specific denomination)	0.26	0.00	0.22	0.25	0.00
Other religion (e.g., Methodist, Episcopal)	0.05	0.00	0.19	0.19	0.00
Nonsectarian	0.00	1.00	0.30	0.21	1.00
School type					
Private	1.00	0.00	1.00	1.00	0.00
Traditional public school	0.00	0.93	0.00	0.00	0.81
Charter	0.00	0.00	0.00	0.00	0.08
Magnet	0.00	0.07	0.00	0.00	0.11
Schoolwide Title 1	n/a	1.00	n/a	n/a	0.90
Enrollment					
<i>Mdn</i>	225	649	77	120	628
<i>M</i>	399	640	173	203	643
Minimum	107	342	5	5	272
Maximum	1,402	933	1,603	1,603	1,309
School community type					
City	0.74	0.36	0.43	0.44	0.49
Suburb or town	0.16	0.14	0.32	0.35	0.30
Rural	0.11	0.50	0.24	0.21	0.21
Racial composition					
Average of school % White	0.64	0.14	0.73	0.71	0.35
Average of school % Black	0.15	0.49	0.14	0.15	0.37
Average of School% Hispanic	0.13	0.31	0.05	0.06	0.18
Pupil:teacher ratio					
<i>Mdn</i>	12	14	8	9	16
<i>M</i>	12	14	8	9	16
Minimum	8	12	1	1	10
Maximum	19	17	27	25	49
Full-time equivalent teachers					
<i>Mdn</i>	26	45	11	13	40
<i>M</i>	34	45	18	20	41
Minimum	11	26	1	2	10
Maximum	94	73	231	231	78
Standardized tests used					
Iowa Test of Basic Skills	0.44	n/a	n/a	n/a	n/a
Terra Nova	0.31	n/a	n/a	n/a	n/a
Stanford Achievement Test	0.25	n/a	n/a	n/a	n/a
California Achievement Test	0.06	n/a	n/a	n/a	n/a

(continued)

TABLE 3 (CONTINUED)

Description	(1) Private Schools in Sample	(2) Public Schools in Sample	(3) All North Carolina Private Schools	(4) All Voucher-Accepting Private Schools	(5) Sending Public Schools
Preliminary Scholastic Assessment Test	0.06	n/a	n/a	n/a	n/a
North Carolina End of Grade Tests	n/a	1.00	n/a	n/a	1.00

Note. n/a implies this data point was not available for schools in this sector; $n = 14$ public schools and 24 private schools in the testing sample. All testing was conducted in spring 2017. Tested students per school refers to the number of students tested by the North Carolina State research team for this project. Some schools use more than one standardized test so these individual percentages are not expected to sum to 100%. “Standardized tests used” refers to the accountability tests used to assess all students in a school, not just Opportunity Scholarship students. Sending public schools are defined as public schools that lost students to the Opportunity Scholarship Program in 2015–2016 ($n = 141$). Source. Private school data are from the 2015–2016 Private School Universe Survey and the North Carolina Division of Non-Public Education records. Public school data are from the 2015–2016 U.S. Department of Education Public Elementary/Secondary School Universe Survey.

Variables

Outcome Measures. The ITBS is a standardized, nationally norm-referenced test. This study relies upon Form E, which is the ITBS Survey Battery, a shorter version of the ITBS Complete Battery, which would take several days to administer. The Survey Battery was selected because it consists of three 30-minute tests in the areas of reading, language, and math, which we administered in a single testing session. We relied upon the tests at Levels 10 through 14, which correspond to Grade Levels 4 through 8. Students were administered the test that matched their current grade level. Students’ answer booklets were machine-scored by the test publishers, and the results were made available to us for analysis.² Table 4 displays the correlations between the ITBS and North Carolina EOG test scores.

Covariates. All of the public school students in our sample have EOG exam scores from 2016, but only 89 out of 245 private school students have EOG records from 2016. Of these, 29 (12%) attended a public school in 2016, but the number of membership days suggest they did not attend for the entire school year, which may explain why they are missing EOG exam scores from spring 2016. Another 31 students (13%) have EOG records from 2015, suggesting that their missing 2016 EOG records are due to participation in the Opportunity Scholarship Program for 2 years rather than just 1 year. In the latter scenario, we run the analyses both with and without these students in order to cleanly estimate 1-year effects, as compared to running models that would estimate a mixture of 1- and 2-year effects. In addition to EOG scores, we have access to data on gender, race/ethnicity, and whether the student had a disciplinary incident during the previous year.

Estimation Strategy

We use IPW to estimate the achievement changes associated with participation in North Carolina’s Opportunity Scholarship Program, a commonly used approach to causal inference in the voucher literature when a lottery study is not

possible. We use students’ observable characteristics to predict their likelihood of using a voucher, and these propensities are used to construct IPWs that result in comparable treatment and comparison groups. Within-study comparisons are one approach researchers have employed to assess the validity of propensity approaches such as the one employed here, with generally encouraging results.

We use IPW in a regression model to estimate the effect of private school attendance through the Opportunity Scholarship Program on students’ academic performance. IPW adjusts for differences between the treatment and comparison groups by using some form of the inverse of the predicted probability of treatment as a weight in a statistical model. IPW is one of a general class of techniques commonly referred to as matching or propensity score analysis. We provide tests of the IPW assumptions in a supplemental online appendix.

We first estimate the predicted probability of treatment using the following logistic regression model:

$$P(D = 1) = B_0 + B_1X_1 + \dots + B_kX_k + \mu,$$

where D is an indicator variable for whether the student attended private school, and X_1 to X_k are the set of covariates described below. These predicted probabilities are used to generate IPWs, with treated (private school) students assigned a weight equal to 1, and comparison (public school) students assigned a weight equal to $P(D = 1)/1 - P(D = 1)$. Because families who choose to use a voucher are typically unrepresentative of the general population (Egalite et al., 2017), we generate weights for the average treatment effect of the treated, rather than the average treatment effect. These weights are then used in the estimation of our outcome models, using multiple regression:

$$Y = B_0 + B_1D + B_2X_1 + \dots + B_{k+1}X_k + \mu,$$

where Y is one of the three ITBS measures described above. The coefficient B_1 from this model yields the effect of private school attendance on academic performance.

TABLE 4
Correlations Between State EOG Scores and ITBS Scores

Test score	Math EOG	Reading EOG	Math ITBS	Reading ITBS	Language ITBS
Math state EOG score	1.00				
Reading state EOG score	.48	1.00			
Math ITBS score	.62	.30	1.00		
Reading ITBS score	.45	.57	.53	1.00	
Language arts ITBS score	.45	.55	.56	.72	1.00

Note. EOG = End of Grade assessment; ITBS = Iowa Test of Basic Skills.

TABLE 5
Achievement Changes Associated With Participation in the North Carolina Opportunity Scholarship Program

Treatment group	OLS Results			Effect Sizes			<i>n</i>		
	Math	Reading	Language	Math	Reading	Language	Treatment	Control	Total
New OS recipients	6.49** (2.22)	3.90 [†] (2.14)	7.89** (2.48)	0.36	0.22	0.44	89	156	245
Renewal OS recipients	6.01 [†] (3.24)	7.64 [†] (3.85)	9.41** (3.35)	0.33	0.42	0.52	68	140	208
New OS recipients; school does not use ITBS	0.74 (1.98)	1.84 (2.15)	1.35 (1.87)	0.04	0.10	0.07	59	157	216
New OS recipients; school does use ITBS	11.29** (2.88)	7.11* (2.98)	15.50** (3.20)	0.63	0.39	0.86	29	191	220

Note. OLS = ordinary least square; OS = Opportunity Scholarship; ITBS = Iowa Test of Basic Skills. Coefficients are for private school status from regressions with ITBS composite scores in math, reading, and language as the dependent variables, propensity covariates as controls, and using average treatment effects of the treated inverse propensity weights. Standard errors in parentheses, clustered at the 2017 school level.

[†] $p < .10$. * $p < .05$. ** $p < .01$.

Results

We estimate the effect of the Opportunity Scholarship Program using a multiple regression model with a dummy variable indicating treatment, the propensity score covariates as control variables, and the IPWs used during estimation (Table 5). The average treatment effects of the treated IPWs are used to ensure comparability between the two groups of private and public school students. Including the propensity covariates controls for remaining differences between the two groups after weighting and increases our statistical power.

We first estimate the effect of voucher receipt on the sample of new voucher recipients. These are students in their first year of private school using the voucher during the 2016–2017 school year. The top panel of Table 5 shows the differences in performance on the ITBS for this treatment group compared to the public school sample. In general, the results show that in our testing sample, voucher recipients scored higher than their public school counterparts in all three subject areas examined—math, reading, and language. The *SD* in our sample for the three tests is around 18 points, so the effect sizes for voucher recipients range from .22 to .44 *SD*.

One downside associated with focusing on new Opportunity Scholarship students is that the estimated effect of the voucher cannot be separated from the effect of switching schools. If the treatment group students are slow to adjust to the standards, expectations, and culture of their new schooling environment, this may attenuate the estimated effect of the voucher. The bias could work in the opposite direction, too—for example, if the initial excitement of having “won” a voucher seat motivates students to put forth greater effort in the short run, temporarily boosting achievement but not resulting in sustained gains into the second year of voucher use.

Therefore, to generate estimates that are clean of these potential biases, we next estimate the achievement changes associated with voucher use for renewal students instead. Here, the treatment group consists of voucher students in their *second* year of voucher use and the propensity weights for both the public and private school groups are constructed using the 2015 EOG instead of the 2016 EOG scores. The main caveat to bear in mind when interpreting results from this sample is that selective attrition may be shaping who remains in the program beyond their first year, which could

result in upwardly biased estimates for analyses that rely on this sample.

In the models examining impacts for renewal students, the coefficients for all three subjects—math, reading, and language—remain positive, although only the language coefficient is statistically significant at $p < .05$, which may be related to the drop in sample size required to run these models. Of note, the coefficient estimates for reading and language scores actually increase by 3.74 points (i.e., $7.64 - 3.90$) and 1.52 (i.e., $9.41 - 7.89$), respectively, whereas math scores decrease slightly, by approximately one half of one point (i.e., $6.49 - 6.01$). Thus, we conclude that students in our sample who are still participating in the Opportunity Scholarship Program after 2 years perform 9 points higher on the language assessment than comparable students who remain in public schools. This is equivalent to .52 *SD*.

While it is important to note that none of the schools in our sample actually use the assessment administered for this study—the Short Battery (Form E) of the ITBS—some of the private schools in our sample do administer the Complete Battery of the ITBS to students in third, sixth, and 12th grade for internal assessment purposes. While these schools may not explicitly focus their instruction to ensure their students perform well on the ITBS, they may have selected the ITBS because the scope and sequence of tested content most closely align to their curriculum. Thus, our positive results could be explained, in part, by an alignment with the assessment used to measure student achievement or with a perception of the assessment we administered as being “a high-stakes test” by students who recognize the name of the test publisher. This potential imbalance could be partially mitigated, however, by the fact that at least one of the North Carolina public school districts in our sample also administers the ITBS as a screening tool to identify academically or intellectually gifted students, thus raising the stakes on the comparison side too.

Unfortunately, we cannot identify which, if any, of the public school students in our sample have been administered the ITBS by their school district, but we can test the theory of ITBS familiarity on the private school side by repeating our analysis of new Opportunity Scholarship recipients and splitting the treatment group into two groups, depending on whether or not the private school attended by a student administers the ITBS. The bottom panel of Table 5 shows the results of these analyses. The large positive coefficients from our previous analyses largely disappear when analyzing students from private schools that do not use ITBS, implying that Opportunity Scholarship students attending these schools perform about the same as their public school counterparts on this assessment. The treatment variable is no longer statistically significant. The results for the ITBS sample, meanwhile, remain statistically significant, with coefficients that range from .39 to .86 *SD*.

The large, positive coefficients for ITBS schools are important because they are likely contributing to the positive

coefficients observed in the main results. They may, in part, reflect curricular alignment with the ITBS, but we also cannot rule out other explanations. For instance, it could be the case that use of the ITBS is correlated with other influential school characteristics, such as school age, school size, teacher quality, school leader quality, peer group quality, and so on. It could also be driven by student selection into certain types of schools. It is also worth noting that all of the Catholic schools in our sample use the ITBS, such that half of the students in the ITBS-only private school group attend Catholic schools, versus none in the non-ITBS sample, making it challenging to tease apart any achievement benefit associated with attending a Catholic school from the benefit of attending an ITBS-using school.³ Thus, we cannot rule out the possibility that the large positive coefficients observed in this analysis may reflect factors other than curricular alignment with the assessment.

This latter test complements work by Mills and Wolf (2017) that attempts to shed light on the issue of potential “test familiarity bias” in the context of a rigorously designed evaluation of the Louisiana Scholarship Program. By taking advantage of the fact that different grade levels relied on different accountability assessments in some years, some of which were more aligned with the state curriculum than others, Mills and Wolf decompose the negative impact of Louisiana’s voucher program to identify what proportion of that impact can be attributed to the specific assessment used. They conclude that about half of the negative impact of the Louisiana voucher program on test score outcomes could be attributed to the curricular alignment of the test. Our parallel finding of differential impacts for North Carolina private schools that use the ITBS assessment bolsters this evidence from Louisiana, which invites consideration of the influence of test alignment when drawing inferences from school choice evaluations.

Discussion

The overall results presented here reveal large positive test score increases associated with voucher usage in North Carolina, although further analysis points to particularly large coefficients for private schools that use the ITBS assessment, as opposed to alternative third-party nationally normed tests such as the Terra Nova, the California Achievement Test, or the North West Evaluation Association’s Measures of Academic Progress. Because we administered the ITBS to participants in this evaluation, the concern is that perceived accountability pressure associated with test name recognition or curricular alignment, or some other factor associated with a private school’s usage of the ITBS, may be driving the overall findings.

Although we collected data from approximately 700 students, our preferred specification uses observations from only 245 students, controlling for a host of covariates.

These analyses reveal first-year coefficient estimates of .36 *SD* in math and .44 *SD* in language. Two-year coefficient estimates suggest that student achievement in math was unchanged but language scores improved by .52 *SD*. While these findings may be surprising in that they run contrary to recent evidence from Louisiana (Abdulkadiroglu et al., 2018; Mills & Wolf, 2017), Ohio (Figlio & Karbownik, 2016), and Indiana (Waddington & Berends, 2018), the positive associations observed here are in line with the bulk of the literature on voucher usage in the United States since the 1990s. It may be the case that participation in the North Carolina Opportunity Scholarship Program is associated with higher student achievement, perhaps because it reaches highly economically disadvantaged students (Egalite et al., 2017) who have few school choice options in the absence of the program and perhaps the highest potential for academic growth, as a result. Three other potential explanations, which cannot be fully ruled out because of study limitations, also merit consideration.

The first potential alternative explanation for the positive coefficients reported here relates to a limitation of the quasi-experimental research design employed. IPW allows us to mitigate potential selection bias by accounting for observable differences between students in the treatment and comparison groups. This is greatly strengthened by our inclusion of pretreatment measures of students' academic performance (Bifulco, 2012). Nonetheless, it is possible that the two groups of students are not perfectly aligned along unobservable dimensions, which could introduce potential selection bias. This is always a possibility in research designs that cannot leverage a lottery for causal inference. For instance, if parents of voucher-receiving students have more resources compared to their public school peers, these unobserved characteristics may lead to higher test scores after the student switches into private school. We do not believe this is a major concern, however, given how the sample was constructed to maximize comparability along this dimension. Furthermore, a descriptive analysis of household income suggests that voucher-receiving families are actually among the poorest households in the state (Egalite et al., 2017).

Second, it is possible that the choice of test unfairly advantaged the private school students. This would be a serious concern if the test chosen for this analysis was a criterion-referenced test, aligned to a specific set of standards and content sequence, such as the North Carolina EOG tests, but the ITBS is a norm-referenced test that is not aligned with any one curriculum. A related concern is if the private schools in our sample use the ITBS for their annual assessment and students in the study sample were seeing the same set of questions for the second time. We tried to avoid this scenario in two ways. First, we tested early enough in the spring so that the students took the test

for our study before taking the test for their own school's accountability protocol. A complementary benefit of this timing decision was that we avoided unfair test fatigue among public and private school students, because neither group had started their own spring testing yet. We also avoided this potential pitfall by selecting a short battery test that had not been reported previously to the SEAA as being in use among private schools for accountability purposes. Thus, while we believe we have made every effort to mitigate this potential issue, it raises an important design issue for the policy evaluation community: In studies of this nature, one should select an assessment instrument that is not similar to any of the tests already used by the treatment or comparison groups. In the voucher context, specifically, policymakers should heed this important issue in the program design phase by requiring common testing across public and private schools for evaluation purposes.

Third, it is also important to consider issues related to the construction of our analysis sample when interpreting these results. Even though we succeeded in administering a common assessment to almost 700 students, the final analysis sample is modestly sized. Indeed, it is even too small to allow for investigations of heterogeneity in the estimated effects by students' demographic characteristics. Furthermore, because students and schools were not randomly selected to be in the study, our sample is not representative of all voucher users or all voucher-accepting private schools in 2016–2017. For example, every voucher student in our study is attending a religious private school, even though nonsectarian private schools also participate in the Opportunity Scholarship Program (e.g., in 2017–2018, 22% of participating private schools were nonsectarian). Furthermore, on both the public and private school sides, we were forced to rely on volunteer students, which may have resulted in unusually high performers in our sample. Recruitment efforts were supported by public and private school partners including individual public and private school principals, public school district central office staff, and a school choice interest group. Thus, the results reported here are not reflective of the average expected gains for a typical voucher student attending a North Carolina private school by way of the Opportunity Scholarship Program.

The sample construction protocol also raises interesting questions about how we should conceptualize the "treatment." Students in the comparison group attend large, high-poverty public schools, whereas students in the treatment group attend smaller, less racially diverse private schools. As with any private school voucher study, test score differences between these two groups may be related to differences in peer effects, school size, and resources as much as they reflect switching from the public to private sector.

A final limitation is the short time frame in which we have been able to examine impacts. We present 1- and 2-year

estimates of changes in student achievement associated with voucher use, but North Carolina's Opportunity Scholarship Program is still in its infancy and we do not know if these effects will hold up, grow stronger, or disappear over time as the program grows and evolves.

Conclusion

This analysis of North Carolina's Opportunity Scholarship Program uses a quasi-experimental design to estimate the achievement changes associated with using a voucher to attend a religious private school in North Carolina. We use IPWs in a regression model to estimate changes in student performance in math, reading, and language. The test score increases we observe are positive, large, and statistically significant, ranging from .25 to .49 *SD*. As a robustness check, we also run models that account for preexisting trends in student achievement by controlling for both 2015 and 2016 test scores. Coefficients remain positive and statistically significant but shrink in magnitude to .21 to .39 *SD*. Although not a comprehensive analysis of all 5,624 voucher recipients in the 2016–2017 school year, this analysis presents the first estimate of the achievement changes associated with the use of a means-tested voucher in the Tar Heel state, filling a gap in the literature on the achievement changes associated with contemporary statewide voucher programs. A further contribution of this study is an analysis of the subgroup of private schools that regularly administer the complete battery of the ITBS, a nationally norm-referenced standardized test that is similar to the short battery test administered for this evaluation, pointing to the issue of “test familiarity bias” in the context of otherwise rigorously designed evaluations, an important consideration for the design and implementation of future impact studies.

Notes

1. Adjusted per-household income is calculated as household income/square root of household size, a commonly used approach to account for household savings resulting from economies of scale and resource-sharing. The unadjusted median household income for voucher recipients is \$31,485.

2. Specific information on the validity and reliability of the Iowa tests is available here: <https://itp.education.uiowa.edu/ia/documents/ITBS-Research-Guide.pdf>

3. To test this theory, we descriptively compare the outcomes of students in Catholic schools to the outcomes of students in other private schools that also use the ITBS. The coefficient on the Catholic school dummy is statistically insignificant, which sheds doubt on the theory of a “Catholic school effect” explaining higher scores. We urge the reader not to overinterpret these findings, however, given the extremely small sample size. Specifically, in our data, all of the students in other private schools that also use the ITBS attend a single private school. Thus, we cannot present a full table of results examining this question without risking identifying a participating private school or group of students.

References

- Abdulkadiroglu, A., Angrist, J. D., Cohodes, S., Dynarski, S. M., Fullerton, J., Kane, T. J., & Pathak, P. A. (2009). *Informing the debate: Comparing Boston's charter, pilot and traditional schools*. Boston Foundation Report.
- Abdulkadiroglu, A., Pathak, P. A., & Walters, C. R. (2018). Free to choose: Can school choice reduce student achievement? *American Economic Journal*, *10*(1), 175–206. <https://doi.org/10.1257/app.20160634>
- Angrist, J. D., Cohodes, S. R., Dynarski, S. M., Fullerton, J. B., Kane, T. J., Pathak, P. A., & Walters, C. R. (2011). *Student achievement in Massachusetts' charter schools*. Center for Education Policy Research, Harvard University.
- Barnard, J., Frangakis, C. E., Hill, J. L., & Rubin, D. B. (2003). Principal stratification approach to broken randomized experiments: A case study of School Choice vouchers in New York City. *Journal of the American Statistical Association*, *98*(462), 288–311. <https://doi.org/10.1198/016214503000071>
- Behrman, J. R., Kletzer, L. G., McPherson, M. S., & Schapiro, M. O. (1998). The microeconomics of college choice, careers, and wages: Measuring the impact of higher education. *Annals of the American Academy of Political and Social Science*, *559*(1), 12–23. <https://doi.org/10.1177/0002716298559001002>
- Bettinger, E., & Slonim, R. (2006). The effect of educational vouchers on academic and non-academic outcomes: Using experimental economic methods to study a randomized natural experiment. *Journal of Public Economics*, *90*(8–9), 1625–1648. <https://doi.org/10.1016/j.jpube.2005.10.006>
- Bifulco, R. (2012). Can nonexperimental estimates replicate estimates based on random assignment in evaluations of school choice? A within-study comparison. *Journal of Policy Analysis and Management*, *31*(3), 729–751. <https://doi.org/10.1002/pam.20637>
- Bitler, M. P., Domina, T., Penner, E. K., & Hoynes, H. W. (2014). Distributional analysis in educational evaluation: A case study from the New York City voucher program. *Journal of Research on Educational Effectiveness*, *8*(3), 419–450.
- Chingos, M. M. (2018). *The effect of the D.C. School Voucher Program on college enrollment*. <https://www.urban.org/research/publication/effect-dc-school-voucher-program-college-enrollment>
- Chingos, M. M., Kuehn, D., Monarrez, T., Wolf, P. J., Witte, J. F., & Kisida, B. (2019). *The effects of means-tested private school choice programs on college enrollment and graduation*. https://www.urban.org/sites/default/files/publication/100665/the_effects_of_means-tested_private_school_choice_programs_on_college_enrollment_and_graduation_2.pdf
- Chingos, M. M., Monarrez, T., & Kuehn, D. (2019). *The effects of the Florida Tax Credit Scholarship Program on college enrollment and graduation: An update*. <https://www.urban.org/research/publication/effects-florida-tax-credit-scholarship-program-college-enrollment-and-graduation>
- Chingos, M. M., & Peterson, P. E. (2015). Experimentally estimated impacts of school vouchers on college enrollment and degree attainment. *Journal of Public Economics*, *122*, 1–12. <https://doi.org/10.1016/j.jpube.2014.11.013>
- Chubb, J. E., & Moe, T. M. (1990). *Politics, markets, & America's schools*. Brookings Institution Press.

- Cole, S. R., & Hernan, M. A. (2008). Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology*, *168*(6), 656–664. <https://doi.org/10.1093/aje/kwn164>
- Coleman, J. S., & Hoffer, T. (1987). *Public and private high schools: The impact of communities*. Basic Books.
- Cowen, J. M. (2008). School choice as a latent variable: estimating the “complier average causal effect” of vouchers in Charlotte. *Policy Studies Journal*, *36*(2), 301–315. <https://doi.org/10.1111/j.1541-0072.2008.00268.x>
- Cowen, J. M., Fleming, D. J., Witte, J. F., Wolf, P. J., & Kisida, B. (2013). School vouchers and student attainment: Evidence from a state-mandated study of Milwaukee’s Parental Choice Program. *Policy Studies Journal*, *41*(1), 147–167. <https://doi.org/10.1111/psj.12006>
- Dobbie, W. S., & Fryer, R. G. (2017). *Charter schools and labor market outcomes* (NBER Working Paper No. 22502). <http://www.nber.org/papers/w22502>
- Dynarski, M., Rui, N., Webber, A., Gutmann, B., & Bachman, M. (2019). *Evaluation of the D.C. Opportunity Scholarship Program: Impacts two years after students applied* (NCEE 2018-4010). U.S. Department of Education.
- Egalite, A. J., Porter, S. R., & Stallings, D. T. (2017). *A profile of applicants to North Carolina’s private school voucher program: Descriptive data on 2016-17 applicants to the Opportunity Scholarship Program* (OS Evaluation Report No. 3). <https://ced.ncsu.edu/elphd/wp-content/uploads/sites/2/2017/07/A-Profile-of-Applicants-to-North-Carolina%E2%80%99s-Opportunity-Scholarship-Program.pdf>
- Figlio, D. N., & Karbownik, K. (2016). *Evaluation of Ohio’s EdChoice Scholarship Program: Selection, competition, and performance effects*. Thomas B. Fordham Institute. https://edex.s3-us-west-2.amazonaws.com/publication/pdfs/FORDHAM%20Ed%20Choice%20Evaluation%20Report_online%20edition.pdf
- Fortson, K., Verbitsky-Savitz, N., Kopa, E., & Gleason, P. (2012). *Using an experimental evaluation of charter schools to test whether nonexperimental comparison group methods can replicate experimental impact estimates* (NCEE Technical Methods Report No. 2012-4019). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Friedman, M. (1962). *Capitalism and freedom*. University of Chicago Press.
- Friedman, M. (1955). The role of government in education. In R. A. Solo (Ed.), *Economics and the public interest* (pp. 123–144). Rutgers University Press.
- Fuller, H., & Page, L. F. (2014). *No struggle, no progress: A warrior’s life from black power to education reform*. Marquette University Press.
- Greene, J. P. (2001). Vouchers in Charlotte. *Education Next*, *1*(2), 55–60.
- Greene, J. P., Peterson, P. E., & Du, J. (1999). Effectiveness of school choice: The Milwaukee experiment. *Education and Urban Society*, *31*, 190–213.
- Hanushek, E. A., & Pace, R. R. (1995). Who chooses to teach (and why)? *Economics of Education Review*, *14*(2), 101–117. [https://doi.org/10.1016/0272-7757\(95\)90392-L](https://doi.org/10.1016/0272-7757(95)90392-L)
- Howell, W., Wolf, P. J., Campbell, D., & Peterson, P. E. (2002). School vouchers and academic performance: Results from three randomized field trials. *Journal of Policy Analysis and Management*, *21*(2), 191–217. <https://doi.org/10.1002/pam.10023>
- Jin, H., Barnard, J., & Rubin, D. B. (2010). A modified general location model for noncompliance with missing data: Revisiting the New York City School Choice Scholarship Program using principal stratification. *Journal of Educational and Behavioral Statistics*, *35*(2), 154–173. <https://doi.org/10.3102/1076998609346968>
- Kisida, B., & Wolf, P. J. (2015). Customer satisfaction and educational outcomes: Experimental impacts of the market-based delivery of public education. *International Public Management Journal*, *18*(2), 265–285. <https://doi.org/10.1080/10967494.2014.996629>
- Krueger, A., & Zhu, P. (2004). Another look at the New York City voucher experiment. *American Behavioral Scientist*, *47*(5), 658–698. <https://doi.org/10.1177/0002764203260152>
- Lee, B. K., Lessler, J., & Stuart, E. A. (2011). Weight trimming and propensity score weighting. *PLoS One*, *6*(3), 1–6. <https://doi.org/10.1371/journal.pone.0018174>
- Mill, J. S. (1869). *On liberty*. Longman, Roberts & Green.
- Mills, J. N., & Wolf, P. J. (2017). Vouchers in the Bayou: The effects of the Louisiana Scholarship Program on student achievement after two years. *Educational Evaluation and Policy Analysis*, *39*(3), 464–484. <https://doi.org/10.3102/016237371717693108>
- Mills, J. N. & Wolf, P. J. (2019). The effects of the Louisiana Scholarship Program on student achievement after four years. EDRE Working Paper 2019-10. <http://www.uaedreform.org/wp-content/uploads/Mills-Wolf-LSP-Achievement-After-4-Years-final.pdf>
- Neal, D. A., & Johnson, W. R. (1996). The role of pre-market factors in black–white differences. *Journal of Political Economy*, *104*(5), 869–895. <https://doi.org/10.1086/262045>
- Rivkin, S. G. (1995). Black/white differences in schooling and employment. *Journal of Human Resources*, *30*(4), 826–852. <https://doi.org/10.2307/146234>
- Rosenbaum, P. R., & Rubin, D. B. (1983b). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. <https://doi.org/10.1093/biomet/70.1.41>
- Rouse, C. E. (1998). Private school vouchers and student achievement: An evaluation of the Milwaukee parental choice program. *The Quarterly Journal of Economics*, *113*(2), 553–602.
- Waddington, R. J., & Berends, M. (2018). Impact of the Indiana Choice Scholarship Program: Achievement effects for students in upper elementary and middle school. *Journal of Policy Analysis and Management*, *37*(4), 783–808. <https://doi.org/10.1002/pam.22086>
- Webber, A., Rui, N., Garrison-Mogren, R., Olsen, R. B., Gutmann, B., & Bachman, M. (2019). *Evaluation of the D.C. Opportunity Scholarship Program: Impacts three years after students applied* (NCEE 2019-4006). U.S. Department of Education.
- Weinschrott, D. J., & Kilgore, S. B. (1998). Evidence from the Indianapolis Voucher Program. In P. E. Peterson, & B. C.

- Hassel (Eds.), *Learning from school choice* (pp. 307–333). Brookings Institution.
- Wettach, J. R. (2017). *School vouchers in North Carolina: The first three years*. The Children’s Law Clinic at Duke Law School. https://law.duke.edu/childedlaw/docs/School_Vouchers_NC.pdf
- White, J. (2016). *Give vouchers time: Low-income families need as many quality school options as possible*. Brookings Institution. <https://www.brookings.edu/research/give-vouchers-time-low-income-families-need-as-many-quality-school-options-as-possible/>
- Witte, J., Carlson, D., Cowen, J. M., Fleming, D. J., & Wolf, P. J. (2012). *MPCP longitudinal educational growth study: Fifth year report* (SCDP Milwaukee Evaluation Report No. 29). The School Choice. Demonstration Project. <http://www.uae-dreform.org/category/department-of-education-reform/scdp/milwaukee-evaluation/mpcp-final-reports/>
- Wolf, P. J., Kisida, B., Gutmann, B., Puma, M., Eissa, N., & Rizzo, L. (2013). School vouchers and student outcomes: Experimental evidence from Washington, D.C. *Journal of Policy Analysis and Management*, 32(2), 246–270. <https://doi.org/10.1002/pam.21691>
- Wolf, P. J., & McShane, M. (2013). Is the juice worth the squeeze? A benefit-cost analysis of the District of Columbia Opportunity Scholarship Program. *Education Finance and Policy*, 8(1), 74–99. https://doi.org/10.1162/EDFP_a_00083
- Wolf, P. J., Witte, J. F., & Kisida, B. (2019). *Do voucher students attain higher levels of education? Extended evidence from the Milwaukee Parental Choice Program* (EdWorkingPaper No. 19-115). <https://doi.org/10.26300/1tnp-hk76>

Authors

ANNA J. EGALITE is an assistant professor in the Department of Educational Leadership, Policy, and Human Development at North Carolina State University. Her research focuses on the evaluation of education policies and programs intended to close racial and economic achievement gaps.

D. T. STALLINGS is the director of policy research at the William and Ida Friday Institute for Educational Innovation in the College of Education at North Carolina State University. His research focuses on education policy evaluation and implementation.

STEPHEN R. PORTER is a professor of higher education in the Department of Educational Leadership, Policy, and Human Development at North Carolina State University. His research focuses on student success, with an emphasis on evaluation using quasi-experimental methods, and survey methods, particularly the validity of college student survey questions.