Heterogeneous Effects across Schools in the First Year of the Louisiana Scholarship Program

Matthew Lee*

Jonathan N. Mills

Patrick J. Wolf

School Choice Demonstration Project
Department of Education Reform
University of Arkansas
201 Graduate Education Building
Fayetteville, AR 72701
Tel: (479) 575-6345

Paper prepared for the 43rd Annual Conference of the
Association for Education Finance and Policy
Portland, OR

March, 2018

* Corresponding author. Address 201 Graduate Education Building, Fayetteville, Arkansas, 72701;
Tel.:1-479-575-6345; *E-mail address*: mhl002@uark.edu

Abstract

The Louisiana Scholarship Program (LSP) is a school voucher program that offers publicly-funded scholarships to students from economically-disadvantaged families to attend a participating private school of their choice. While school choice theory suggests that market pressures should improve quality, experimental evaluations of the program find significant negative effects in math and reading scores after their first year; however, effect estimates appear to diminish towards insignificant differences by the end of the third year. These effects are generally homogeneous across student characteristics with some important exceptions, including positive and significant effects in ELA for the lowest performers at baseline and negative effects that persist into Year 3 for applicants to earlier grades. This study builds on previous work with an exploratory analysis of the variation in treatment effects across school characteristics in the first year. In general, we do not observe effect heterogeneity across school characteristics, though we find less negative results for math outcomes in non-Catholic religious schools; better impacts for schools with higher enrollments, hours per school day, total instructional hours, and employees; and no statistically significant differences between treatment and control students for schools at the top of the tuition distribution.

Heterogeneous Effects across Schools in the First Year of the Louisiana Scholarship Program

Private school choice programs, which provide families with public funds to attend private schools of their choosing, are among the most controversial education reform initiatives in the United States. Proponents argue such programs, often called school vouchers or scholarships, expand educational options available to families, thereby better facilitating the matching of student education needs with school offerings and improving the education system as a whole through increased competition among schools for students. Opponents argue choice programs harm traditional public schools and the students who remain in them by stripping away funding and concentrating disadvantaged non-choosing students within their walls.

The extent to which these concerns play out is, however, an empirical question, and the evidence to date is somewhat mixed but leans positive. Taken as a whole, the most rigorous empirical research indicates null or small positive impacts of vouchers on student achievement across programs with noticeable, yet non-systematic, variation in treatment effects across student subgroups as moderators (Egalite & Wolf, 2016; Shakeel, Anderson, Wolf, 2016). Lesser known, however, is the extent to which treatment effects vary across school settings as mediators. After all, parental school choice is a general policy. The form that school choice takes for individual and small groups of students will depend upon the specific schools chosen. This paper examines how treatment effects vary across schools participating in one of the nation's first statewide school voucher programs: the Louisiana Scholarship Program (LSP).

Initially created as a pilot program in New Orleans in 2008, the LSP was expanded statewide for the 2012-13 school year. A total of 9,736 eligible students applied to the means-tested program for the 2012-13 cohort and 5,296 received LSP vouchers. Both initial and

ongoing experimental evaluations of the program report large negative effects on English Language Arts (ELA) and mathematics achievement after one year of voucher usage (Mills, 2015; Abdulkadiroglu, Pathak & Walters, 2018) that diminish over time (Mills & Wolf, 2017a; Mills & Wolf, 2017b). Moreover, while voucher usage does not appear to be associated with differential effects by student gender or ethnicity, the evidence suggests persistent negative effects of the program in math for younger applicants and positive effects in ELA for the lowest initial performers. This paper builds on this work by examining how LSP effects vary across school environments.

Specifically, we conduct an exploratory analysis focused on students experiencing the initial statewide expansion: the 2012-13 application cohort. Following the work of Mills & Wolf (2017a), our analysis is restricted to eligible applicants who experienced an oversubscription lottery in order to first calculate unbiased estimates of the impact of LSP voucher usage on student achievement. We then explore how these estimated treatment effects vary across school characteristics, including religious affiliation, geographic location, school tuition, enrollment, student-teacher ratio, instructional hours, and student demographics. Our results suggest that effects can vary across school characteristics in some important ways. The most favorable effect estimates were found in religious (particularly non-Catholic) schools, urban schools, and schools in the top tercile of tuition, K-12 enrollment, and instructional hours. Broadly speaking, however, we observe limited variation in treatment effects across school characteristics. It is possible that this is partially the result of low power, as several differences in effect estimates that are not statistically significant are nonetheless large in magnitude.

This work offers two contributions to the literature on private school choice programs. First, our analysis examines a statewide program. While the majority of choice programs in the

U.S. have been concentrated in urban locales, the number of programs serving entire states has grown in recent years, and now includes programs in Florida, Indiana, and Ohio. Given this expansion in popularity, it is important to understand how these programs work on a broader scale. One would expect, for example, to see increased variation in school characteristics across an entire state as compared to programs concentrated in a single city. Understanding how this variation relates to treatment effects can provide guidance to future policymakers considering such programs.

Second, our analysis builds on a highly rigorous casual design. While this paper's results are purely exploratory—and should not be interpreted causally—the underlying treatment effects we seek to examine stem from a gold standard experimental research design. While the characteristics of one's most preferred school are not determined at random, our focus on estimated treatment effects with strong internal validity gives us strong assurance that we are describing their correlations with school features accurately.

These factors combine to provide a much deeper understanding of how the LSP impacted both students and the education system as a whole. We proceed by summarizing the existing literature on school voucher programs in the US and describing the Louisiana Scholarship Program. We then outline our methodology, including a discussion of our sample and analytical strategy. After presenting the results of our analysis, we conclude with a discussion of our findings and their implications for the program.

The impacts of K-12 voucher programs on achievement in the U.S.

School vouchers are a type of private school choice program that gives families publicly financed scholarships to attend private K-12 schools of their own choosing (Wolf, 2008). The theory backing school choice interventions was developed by political philosophers like Thomas

Paine (1791) and John Stuart Mill (1962 [1869]), as well as by economists like Milton Friedman (1955). A primary claim of school choice programs like vouchers is that, while government should work to offer funds in support of compulsory education mandates, it does not necessarily need to directly deliver the education itself (Friedman, 1955). In contrast, school choice theory posits that choice will improve student academic outcomes by more efficiently allowing families to seek out the schools that best meet their child's needs and by directly incentivizing schools to compete for students (Friedman, 1955; Moe, 2005).

Egalite and Wolf (2016) note voucher programs can differ along five dimensions: region served (locales or states), eligibility (means-tested or universal), level of regulation, value of their voucher allotments, and number of vouchers available. As of January 2017, Schultz et al. (2017) note that there were 54 voucher programs in the US, over half of which were means-tested with an additional 19 restricted to students with disabilities. Further, the majority of programs not exclusively restricted to students with disabilities tend to be operated at the local—rather than state—level. The empirical evidence supporting these programs is mixed, a point to which we next turn.

Empirical evidence on the effects of voucher programs on student achievement

The most rigorous research examining the effects of voucher programs on student achievement broadly suggests mixed results with a slight positive lean. A recent meta-analysis of all experimental evaluations of U.S. programs reports statistically small positive, yet not statistically significant, effects of private school choice programs on student math and reading achievement when averaged across all programs (Shakeel, Anderson & Wolf, 2016). This overall finding masks considerable heterogeneity in effects across programs. For example, two studies of Charlotte's Children's Scholarship Fund (Greene, 2001; Cowen, 2008) as well as early

experimental evaluations of the Milwaukee Parental Choice Program (Greene, Peterson, & Du, 1999; Rouse, 1998) report statistically significant gains in either reading, math, or both subjects. On the other hand, recent experimental evaluations in Louisiana (Abdulkadiroglu, Pathak, & Walters, 2018; Mills & Wolf, 2017b) and Washington, D.C. (Dynarski et al., 2017), as well as rigorous quasi-experimental studies in Indiana (Waddington & Berends, 2017) and Ohio (Figlio & Karbownik, 2016), report statistically significant negative impacts of voucher programs on student achievement, particularly in math. While these latter findings have raised considerable concerns about the feasibility of voucher programs for improving student academics, a broad look at the literature suggests considerable variation across programs.

In addition to heterogeneity in general effects across programs, there is noticeable heterogeneity within programs. Several studies report differential effects by gender, ethnicity, and socioeconomic status; however, there is limited evidence of strong systemic variation in these patterns. For example, Wolf et al. (2013) find that students with higher previous performance, students applying from public schools not classified as "in need of improvement," and females disproportionately benefitted from voucher receipt. Other programs appear to have strong positive effects for African American students, particularly males. Such is the case for the privately-funded Parents Advancing Choice in Education Scholarships in Dayton, OH (Howell et al. 2002) as well as for the New York City voucher program (Barnard, Frangakis, Hill, & Rubin, 2003; Howell & Peterson, 2006; Jin, Barnard, & Rubin, 2010). On the other hand, another evaluation of the New York City voucher program by Krueger and Zhu (2004), which uses a unique method for classifying students as African American, finds no evidence of significant achievement gains, overall or for any participant subgroup. Moreover a fifth study concludes the

program had no clear effects for subgroups along the achievement distribution (Bitler, Domina, Penner, & Hoynes, 2015).

Moreover, there is evidence suggesting effects vary over time. An evaluation of the privately-funded Washington Scholarship Fund in DC found that initial achievement gains disappeared in the third and final years of the study (Howell & Peterson, 2006). An evaluation of a separate DC-based voucher program, the DC Opportunity Scholarship Program, reported significant positive impacts in reading after three years (Wolf et al., 2009, p. 36) that were significant only at a 94 percent confidence level in the fourth and final year of the study (Wolf et al., 2013). Other studies, however, report findings suggesting effects may grow over time. A recent evaluation of the Milwaukee voucher program concluded that a combination of the choice program and a high-stakes testing policy generated test score gains in reading only in the study's fourth and final year (Witte et al. 2014). Similarly, evaluations of the Louisiana Scholarship Program (Mills & Wolf, 2017) and Indiana Choice Scholarship Program (Waddington & Berends, 2017) report large negative impacts early on that slowly return to null after 3 years. Evidence from the Indiana program suggests that voucher students who persisted in the program over four years actually benefited academically compared to matched public school students (Waddington & Berends, 2017).

Lesser known, however, is how voucher treatment effects vary across schools. Choice proponents often argue school choice programs will promote a diverse and innovative education system as schools attempt to cater to family educational desires (Hoxby, 2003). It seems plausible therefore to expect increased variation in school characteristics in a system with school choice.

The research base examining this question is scant, especially regarding the extent to which different school characteristics mediate voucher effects. Previous studies suggest that effects can vary across school characteristics. James Coleman and colleagues (1981) found that students attending Catholic schools were more likely to graduate or enroll in college than their peers, and that these differences were particularly pronounced for minority students. Correlational evidence from Wolf and Hoople's (2006) study of the Washington (DC) Scholarship Fund suggests less extensive school facilities, more homework, and more involved teachers are all related to positive effects. Waddington and Berends (2017) find positive effects in ELA for students using Indiana's Choice scholarship Program to attend Catholic schools compared to negative effects for other voucher users. They also examine if treatment effects vary across urban, suburban, and rural schools, reporting little differences in treatment effects.

Effects may also vary with school enrollment. Low performance, low enrollment, and greater probability of closing tend to be correlated with each other. In addition, effects may depend on the extent to which a given school is reliant on the LSP to maintain enrollment. School choice theory predicts families will respond to low-quality schools by moving to other options (Friedman, 1955). One would expect larger negative effects for LSP students accessing low-quality private schools that would have otherwise closed in the absence of the program. Declining pre-LSP enrollment may serve as a helpful proxy in identifying these schools.

Nevertheless, the extent to which voucher effects are mediated by school characteristics is an empirical question that, to date, has not been examined exhaustively. Our analysis attempts to add to this research by examining if voucher effects vary across schools participating in the Louisiana Scholarship Program.

The Louisiana Scholarship Program and its impact on student achievement

The Louisiana Scholarship Program (LSP) was expanded statewide by Act 2 of the 2012 Regular Session of the Louisiana Legislature and Senate. The program is means-tested: it is limited to students (1) with family income at or below 250 percent of the federal poverty line attending a public school that was graded C, D, or F for the prior school year according to the state's school accountability system, (2) entering kindergarten, or (3) enrolled in the Recovery School District, which includes most of the public schools in the city of New Orleans, several in Baton Rouge, and a single school in Shreveport, Louisiana. In the program's first year, 9,736 students were eligible applicants, a majority of them outside New Orleans.

The LSP voucher is worth 90% of the amount the state and local government provides in student funding to the local school system or the tuition charged by the student's chosen private school, whichever is less. Average tuition at participating private schools ranges from $2,966 to $8,999, with a median of $4,925, compared to average per pupil spending of $8,500 in Louisiana's public schools. LSP regulation stipulates that participating private schools must accept the voucher as full value for tuition, even if the value of the voucher is less.

Private schools must meet certain criteria to participate in the program involving enrollment, financial practice, student mobility, and the health, safety and welfare of students. A survey of participating and non-participating private schools in Louisiana suggests that concern about present and future program regulatory requirements have influenced schools' choices to participate (Kisida, Wolf, & Rhinesmith, 2013), potentially explaining why only a third of eligible private schools opted into the program in 2012-13, although school participation in the LSP has increased slightly since that time.

Eligible applicants to the 2012-13 cohort could submit up to five private school preferences when applying to the program. The Louisiana Department of Education then used a matching algorithm similar to the deferred acceptance lottery used in New York City (Abdulkadiroglu, Pathak, & Roth, 2005) to allocate LSP scholarships to students. The algorithm prevents gaming, incentivizing families to reveal their true school preference rankings. It attempts to place students into their top ranked school while accounting for certain placement priorities, with higher preference given to students participating in the New Orleans pilot program and, among new applicants, to students who had previously attended lower performing public schools.[1]

In cases of oversubscription,[2] the LSP matching algorithm randomly assigned students to receive or not receive an LSP scholarship. Two recent evaluations of the Louisiana Scholarship Program use these oversubscription lotteries for students' first choice school to estimate the impact of LSP scholarship usage on student achievement. Both studies report statistically significant negative impacts of voucher usage on student achievement in reading, math, science, and social studies after one year of scholarship usage (Abdulkadiroglu, Pathak & Walters, 2018; Mills, 2015). These negative effects appear to diminish over time, with math effects roughly halved after two years (Mills & Wolf, 2017a) and not statistically significant after three years (Mills & Wolf, 2017b). The results are heterogeneous, with negative effects persisting for

[1] The LSP priority categories are the following. First, students with disabilities and "multiple birth siblings" – siblings with the same birthdate such as twins, triplets, etc. – are manually awarded LSP scholarships if there is available space at their preferred school. Remaining students are assigned one of six priorities: Priority 1 – Students who received LSP scholarships in the prior school year who are applying to the same school; Priority 2 – Non-multiple birth siblings of Priority 1 awardees in the current round; Priority 3 – Students who received LSP scholarships in the prior school year who are applying to a different school; Priority 4 – New applicants who attended public schools that received a "D" or "F" grade in Louisiana's school accountability system at baseline; Priority 5 – New applicants who attended public schools that received a "C" grade; Priority 6 – New applicants who are applying to kindergarten. See Mills & Wolf (2017a) for further information on the LSP matching process.

[2] Oversubscription lotteries occurred when there were more students applying to a given grade in a given school who were members of the same priority category than seats available (Mills & Wolf, 2017a).

younger students in math, but not ELA, and statistically significant positive effects observed for students performing in the bottom third of the ELA distribution at baseline (Mills & Wolf, 2017b).

The following section outlines our methodology for exploring how characteristics of participating LSP schools mediated these observed effects.

Methodology

We focus in this paper on describing how the impact on achievement after 1 year of using an LSP scholarship to attend a private school varied across different school settings. In doing so, we build on two set portions of an ongoing evaluation of the LSP: studies examining the impact of the program on participant achievement (Mills, 2015; Mills & Wolf 2017a; Mills & Wolf, 2017b) and Sude, DeAngelis, and Wolf's (2017) examination of the types of private schools that opted to participate in the LSP. Given the strikingly negative estimates of the program's effect on student achievement after one year (Abdulkadiroglu et al., 2018; Mills, 2015), it is important to determine how school characteristics mediated these outcomes.

Data

The data for this analysis come from several sources. Louisiana Department of Education provided student-level demographic data, testing data,[3] and application information for all eligible LSP applicants.[4] The LDOE additionally provided information on LSP participating

[3] This study uses student performance on the 2012-13 Louisiana state assessments in grades three through eight as our primary outcome measure of interest. The Louisiana program of assessments offers two alternative assessments for students with disabilities. Performance on these assessments is excluded from our analysis. All students participating in the LSP are required to be tested by their private schools, using the state accountability assessments, for any grade in which the public school system also tests its students. The 2011-12 (baseline) and 2012-13 assessment data in our study contain student scores on the LEAP and iLEAP exams, criterion-referenced tests aligned to Louisiana state education standards. For more information, see Mills (2015).

[4] For more information, see Mills and Wolf (2017a).

schools, which we have supplemented with data from the Private School Universe Survey and reviews of school websites.[5]

Analytical strategy

Our analysis builds on the work of Mills (2015) and Mills and Wolf (2017b) which leverage oversubscription lotteries occurring during the process of matching LSP applicants to schools for the 2012-13 school year. The analyses instrument for actual enrollment in an LSP school using the outcome of oversubscription lotteries for first-choice school preferences to produce unbiased estimates of the program's Local Average Treatment Effect (LATE) on student achievement (Angrist & Pischke, 2009, Cowen, 2008; Mills & Wolf, 2017b). Our analysis expands on this work by adding an interaction of predicted LSP enrollment with school characteristics in the original evaluation's two stage least squares (2SLS) model:

$$(1a) \quad U_i = \sum \pi_j^1 R_{ji} + \delta_1^1 T_i + \delta_2^1 (T_i \times S_i) + \boldsymbol{X_i \beta^1} + u_i^1$$

$$(1b) \quad (U_i \times S_i) = \sum \pi_j^1 R_{ji} + \delta_1^2 T_i + \delta_2^2 (T_i \times S_i) + \boldsymbol{X_i \beta^2} + u_i^2$$

$$(2) \quad A_i = \sum \alpha_j R_{ji} + \tau_1 \widehat{E_i} + \tau_2 \widehat{(E_i \times S_i)} + \boldsymbol{X_i \gamma} + \epsilon_i$$

Where *i* denotes student, *j* denotes first-choice school lottery, and:

- U_i indicates if a student used an LSP scholarship to enroll in an LSP-participating private school in the 2012-13 school year[6]

[5] For more information, see Sude, DeAngelis, and Wolf (2017).

[6] Prior evaluations of school voucher programs have examined enrollment effects in several ways. For example, Mayer et al. (2002) define enrollment as being "consistently enrolled in a private school", while Rouse (1998) defines enrollment as the number of years enrolled in an attempt to capture potential dosage effects. By defining enrollment as "ever attending a private school" our study falls in line with the Wolf et al. (2013) evaluation of the DC Opportunity Scholarship Program.

- R_i is a fixed effect for a student's first choice school lottery[7]
- T_i indicates if a student received an LSP scholarship to their first choice school
- A_i is standardized student mathematics or English Language Arts achievement in Year 1 of the program[8]
- X_i is a vector of student characteristics – including achievement – collected either at baseline (2011-12) or from the student's LSP application form
- S_i is a particular school characteristic of interest, including school's religious affiliation, geographic location, school tuition, enrollment, student-teacher ratio, instructional hours, and student demographics.

The 2SLS procedure first estimates two equations to generate one's predicted likelihood of using a scholarship to attend an LSP school (1a) and an interaction of this prediction with the school characteristic of interest (1b). These predicted values are then used to produce unbiased estimates of the distribution of local average treatment effects across school characteristics (2).[9] We additionally account for nesting of students within lotteries by calculating bootstrapped standard errors (Angrist & Pischke, 2009).

Analytical sample

We examine variation in LSP achievement effects across schools by focusing on a subsample of LSP applicants who experienced oversubscription lotteries for their first-choice

[7] We include a fixed effect for first school choice lottery to account for differing probabilities of success across lotteries (Gerber & Green, 2012). By using fixed effects, we are essentially comparing lottery winners and losers within the same strata to calculate unbiased estimates of the effect of being randomly offered an LSP scholarship. The approach is comparable to analyzing the impact of hundreds of "mini-experiments" and aggregating the results across them.

[8] Student achievement scores are standardized using distributional parameters of outcomes from the control group.

[9] The 2SLS procedure effectively treats students who lose their first choice lottery but go on to win an LSP to a lower school preference as control-group crossovers. The result is an unbiased estimate of the effect of using a LSP scholarship to attend one's first choice school for those who both faced and complied with their lottery assignment for placement in their first-choice school (Bloom & Unterman, 2014).

school in 2012-13. Our analytical sample is restricted to students with baseline achievement data in grades 3 through 7 in 2011-12 who additionally had outcome data in grades 3-7 in 2012-13. Table 1 presents descriptive statistics for LSP applicants meeting these criteria. Nearly 90 percent of students in our analytical sample are African American, the overwhelming majority are free- or reduced-price lunch eligible and are performing at least 35 percent of a standard deviation below the state average on the Louisiana assessments across all subjects.

Table 1.[10]

Characteristics of LSP applicants

	N	Winners	Losers	Adjusted Difference (2) – (3)	Standard Error	*P*-Value
	(1)	(2)	(3)	(4)	(5)	(6)
Female	1,359	0.51	0.51	0.00	0.03	0.93
Ethnicity						
Black	1,359	0.88	0.89	-0.02	0.19	0.25
Hispanic	1,359	0.03	0.02	0.01	0.12	0.29
White	1,359	0.07	0.07	0.00	0.14	0.82
Other	1,359	0.03	0.02	0.01	0.01	0.50
LEP	1,360	0.01	0.01	0.00	0.01	0.65
FRL	1,339	0.81	0.92	-0.01	0.02	0.79
# of choices	1,360	1.97	2.43	-0.25***	0.08	0.00
Baseline Tests[a]						
ELA	1,360	-0.37	-0.32	-0.05	0.06	0.47
Math	1,359	-0.41	-0.39	0.01	0.06	0.87
Science	1,358	-0.50	-0.49	0.02	0.06	0.79
Social Studies	1,358	-0.44	-0.39	-0.04	0.07	0.55

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

[10] Scores are standardized within grade based on the observed distributions of scale scores across Louisiana. *Notes.* Analysis sample excludes students with disabilities and multiple birth siblings. The analysis sample represents LSP applicants to grades four through eight in 2012-13 who did not list a special education exclusion on their LSP application and were not multiple birth siblings. The analysis sample is additionally restricted to students with baseline in grades three through five. *Treatment* refers to students receiving LSP scholarships to their first choice private school. All other students comprise the control group. Demographics are drawn from the 2011-12 testing data. *Adjusted Diff* is the difference between Treatment and Control group students, controlling for first choice school lottery fixed effects. "Standard error" indicates standard error of the difference, which accounts for clustering within lotteries. *Source*: Authors' calculations.

The adjusted difference column compares the characteristics of LSP applicants experiencing the same lottery for their most preferred private school. The general lack of significant differences in these more refined comparisons indicates that students receiving an LSP scholarship by lottery (Treatment) and those who do not (Control) are very similar on nearly every characteristic included. This pattern gives us strong assurance that the local average treatment effects underlying our mediator analysis are calculated with strong internal validity.

School characteristics examined

This paper examines how LSP voucher use effects vary for this analytical sample across schools. Research by Sude, DeAngelis, and Wolf (2017) indicates private schools opting to participate in the LSP had lower enrollment, higher tuition, tended to serve more minority students, and were more likely to be Catholic schools than Louisiana private schools that opted not to participate. Table 2, which compares the characteristics of participating and non-participating LSP schools, confirms these findings.

Table 3 examines the distributions of school characteristics for LSP participating private schools. Fortunately for our analysis, several of the characteristics examined here demonstrate meaningful variation. The range for tuition is quite large, with a low of $2,200 and a high of $14,500. Enrollment, student/teacher ratios, and the number of days in a school year also vary strongly across schools. On the other hand, several variables demonstrate limited variation. For example, only 3 percent of all participating private schools are non-religious and, among religious schools, the overwhelming majority are Catholic. Similarly, nearly all participating private schools are coeducational. The limited variation for these categories suggests that non-

religious and single-gender schools are likely outliers. We therefore recommend caution when generalizing findings for these categories.

Table 2.[11]

Characteristics of Louisiana private schools by LSP participation

	LSP		Non-LSP		Difference	Standard	
	N	Mean	N	Mean	(2) – (4)	Error	*P*-Value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tuition	83	5,385.74	157	6,235.65	-849.91*	432.01	0.0503
Student/Teacher Ratio	87	12.99	196	12.60	0.40	0.55	0.4654
Full-Time Equivalent	87	20.97	196	28.46	-7.49***	2.66	0.0052
K-12 Enrollment	87	275.72	197	363.34	-87.62***	33.08	0.0085
School Day Hours	87	7.09	197	7.00	0.09	0.03	0.1594
School Year Days	86	178.83	197	177.16	1.67**	0.77	0.0302
Total Instructional Hours	86	1,268.27	197	1,240.40	27.88**	12.87	0.0312
Religious	87	0.97	197	0.86	0.11***	0.04	0.0072
Catholic	87	0.68	197	0.45	0.23***	0.06	0.0004
Religious Non-Catholic	87	0.29	197	0.41	-0.12*	0.06	0.0563
Urban	87	0.49	197	0.42	0.07	0.06	0.2557
Coed	87	0.99	197	0.91	0.08**	0.03	0.0170
Demographics							
% Indian	87	0.15	194	0.96	-0.81	0.73	0.1360
% Asian	87	1.87	194	2.49	-0.61	0.61	0.1567
% Hispanic	87	2.31	194	4.08	-1.77*	1.07	0.0988
% Black	87	44.16	194	12.16	32.00***	3.61	0.0000
% Pacific Islander	87	0.05	194	0.26	-0.21**	0.10	0.0309
% ≥ 2 Ethnicities	87	0.98	194	1.97	-1.00	0.78	0.2022

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

[11] *Notes.* "Standard error" means standard error of the difference. "LSP" refers to private schools that chose to participate in the Louisiana Scholarship Program. "Non-LSP" refers to private schools that chose not to participate in the Louisiana scholarship Program. The number of full-time equivalent teachers is the sum of all teachers that taught full time, plus 0.875 times the number of teachers that taught between at least ¾ time but less than full-time, plus 0.625 times the number of teachers that taught at least ½ time but less than ¾ time, plus 0.375 times the number of teachers that taught at least ¼ time but less than ½ time, plus 0.125 times the number of teachers that taught less than ¼ time.

Table 3.[12]

Measures of central tendency and variation of LSP participating school characteristics

	N	Mean	Std. Dev.	Min	Max
	(1)	(2)	(3)	(4)	(5)
Religious	87	0.97	0.18	0	1
Catholic	87	0.68	0.47	0	1
Religious Non-Catholic	87	0.29	0.05	0	1
Coed	87	0.99	0.11	0	1
Urban	87	0.49	0.50	0	1
Tuition	83	$5,385.74	$1,778.34	$2,200	$14,500
K-12 Enrollment	87	275.72	204.37	26	912
Student/Teacher Ratio	87	12.99	3.70	3.8	22.92
Full-Time Equivalent	87	20.97	1.68	3	86.1
School Year Days	86	178.83	6.39	151	223
School Day Hours	87	7.09	0.51	5.5	9
Total Instructional Hours	86	1268.27	12.74	973.5	2007
% Black	87	44.16	40.99	0.24	1.00

Results

This section presents our primary results examining how LSP Local Average Treatment Effects (LATE) vary across schools. Our findings generally do not indicate year 1 LATEs varied strongly across school settings. Notable exceptions exist, however, with results suggesting less negative results for math outcomes in non-Catholic religious schools; better impacts for schools with higher enrollments, hours per school day, total instructional hours, and employees; and no statistically significant differences for schools at the top of the tuition distribution.

Table 4 examines the extent to which LSP LATEs vary by a school's religious nature, coeducational status, and urban setting. Results are presented for performance on Louisiana's state assessments in English Language Arts (ELA) and math. Columns 1 and 3 present results for

[12] *Notes.* Mean of binary variables indicates the proportion of schools that identify with that characteristic. "Religious Non-Catholic" schools include all religious schools that do not identify as Catholic, include Muslim, Jewish, and several Protestant denominations. A "0" for "Religious Non-Catholic" would indicate that the school is either Catholic or non-religious. The number of full-time equivalents is the sum of all teachers that taught full time, plus 0.875 times the number of teachers that taught between at least ¾ time but less than full-time, plus 0.625 times the number of teachers that taught at least ½ time but less than ¾ time, plus 0.375 times the number of teachers that taught at least ¼ time but less than ½ time, plus 0.125 times the number of teachers that taught less than ¼ time. Total instructional hours is the total number of hours per day multiplied by the number of days per year and does not adjust for partial days.

regression models controlling only for lottery fixed effects. Columns 2 and 4 present our preferred models, which additionally control for student demographics, baseline achievement, and the number of school preferences listed on the student's application.

Broadly speaking, for both ELA and math outcomes, we do not find statistically significant differences based on the religious affiliation or secular nature of the school. Regardless of the religious affiliation or secular nature of the school attended by an LSP student, effect estimates for math outcomes are consistently negative and statistically significant ($p < 0.01$). This pattern also generally holds for ELA outcomes; however, the observed negative effect of the LSP on test scores is not statistically significant for students attending non-religious schools.[13]

While effects are generally negative and statistically significant for both test score outcomes and across all religious categorizations, the estimated differences in effects across categories are, themselves, not statistically significant. The lone exception is for Catholic schools: LATE estimates for students using an LSP to attend a Catholic school are nearly .30 percent of a standard deviation worse in math compared to students attending non-Catholic schools ($p < 0.05$). This is a strikingly large difference that contrasts with expectations from correlational research suggesting an academic advantage of Catholic schools (Coleman et al., 1981).

Table 4 also presents results comparing LATE estimates for students attending coeducational and single-gender schools as well as urban versus non-urban schools. No statistically significant differences are detected for LSP students attending single-gender schools when compared with those attending coeducational schools on any outcome. This result is also

[13] This is likely the result of a low statistical power: only 3 percent of LSP private schools are non-religious (see Table 2).

Table 4.[14]

Variation in LATE estimates by school religiosity, coeducational status, and urban location

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Religious	-0.173*	-0.170**	-0.527***	-0.560***
	(0.095)	(0.070)	(0.107)	(0.098)
N	1,517	1,496	1,516	1,495
Non-Religious	-0.422	-0.300	-0.753**	-0.734***
	(0.290)	(0.209)	(0.310)	(0.268)
N	279	277	280	278
Difference	0.249	0.130	0.225	0.136
	(0.304)	(0.220)	(0.324)	(0.282)
Catholic	-0.214	-0.187*	-0.614***	-0.737***
	(0.145)	(0.108)	(0.150)	(0.131)
N	947	928	947	928
Non-Catholic	-0.209*	-0.194**	-0.511***	-0.507***
	(0.125)	(0.083)	(0.139)	(0.120)
N	849	845	849	845
Difference	-0.005	0.007	-0.102	-0.230
	(0.198)	(0.134)	(0.206)	(0.172)
Religious Non-Catholic	-0.114	-0.147**	-0.403***	-0.403***
	(0.116)	(0.074)	(0.147)	(0.131)
N	570	568	569	567
Catholic or Non-religious	-0.263**	-0.214*	-0.647***	-0.736***
	(0.127)	(0.094)	(0.134)	(0.117)
N	1,226	1,205	1,227	1,206
Difference	0.149	0.068	0.244*	0.334**
	(0.172)	(0.115)	(0.199)	(0.170)
Coeducational	-0.156*	-0.157**	-0.523***	-0.597***
	(0.093)	(0.070)	(0.105)	(0.098)
N	1,500	1,479	1,499	1,478
Single-Gender	-0.516*	-0.370*	-0.777**	-0.745***
	(0.296)	(0.217)	(0.318)	(0.276)
N	296	294	297	295
Difference	0.360	0.213	0.254	0.148
	(0.307)	(0.226)	(0.330)	(0.289)
Urban	-0.140	-0.192**	-0.384***	-0.512***
	(0.130)	(0.097)	(0.134)	(0.126)
N	896	876	895	875
Non-Urban	-0.280**	-0.190**	-0.733***	-0.723***
	(0.131)	(0.094)	(0.146)	(0.126)
N	900	897	901	898
Difference	0.140	-0.002	0.349*	0.211
	(0.183)	(0.133)	(0.196)	(0.172)

*Bootstrapped standard errors calculated. *** p < 0.01, ** p < 0.05, * p < 0.1*

[14] *Notes.* Bootstrapped standard errors account for clustering within riskset. "Religious" schools include all Catholic schools, Christian but non-Catholic schools, Muslim schools, and Jewish schools. "Religious Non-Catholic" schools include all religious schools that are not Catholic. "Non-Urban" schools include all schools located in Town, Suburb, or Rural settings, according to the Private School Survey. All first-stage regressions satisfy Staiger and Stock's rule of thumb for a strong instrument (See Appendix Table A1).

generally the case when we compare LATE estimates between students using a voucher to attend an urban school compared with those attending a non-urban school, which includes schools in town, suburban, and rural settings.[15] Interestingly, math LATE estimates are noticeably less negative for urban schools in our simple models which do not control for student demographics or baseline achievement. This difference disappears, however, in the preferred fully specified models.

Next, we explore how LATE estimates vary across school factors such as tuition, enrollment, instructional staffing, student demographics, and instructional days. Table 5 presents results from models examining if there are differential effects across schools. The characteristics in these models have been demeaned, which allows us to determine both the general LATE for students attending an average school in our sample as well as how LATEs vary for students in environments that deviate from the average. Table 6 offers a different view by presenting LATE estimates separately for students attending schools in the top, middle, and bottom thirds of the distribution for a given school category.[16] Combined, the evidence presented in Tables 5 and 6 offer a helpful view of how LSP LATEs vary across schools.

Looking first at Table 5, we find that LATE estimates do not generally vary across schools with different tuitions. The LATE estimates for the average school largely correspond to the overall estimates of the LSP for both ELA and math (Abdulkadiroglu et al., 2018; Mills, 2015; Mills & Wolf, 2017a); and we observe little deviations in the estimated effect when moving away from the average tuition. Table 6, however, presents a slightly different picture

[15] Interestingly, math LATE estimates are noticeably less negative for urban schools than non-urban schools, with point estimates suggesting a difference between 20 and 35 percent of a standard deviation. Results are likely driven by schools in Shreveport and Baton Rouge, as a larger proportion of students in New Orleans were not subject to a randomized lottery, due to the fact that many pilot program students were guaranteed a voucher.
[16] See appendix Table A2 for tercile calculations.

Table 5.[17] *Variation in LATE estimates by school tuition, enrollment, teaching staff, demographics, and instructional days*

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Tuition (demeaned, average = $5,385.74)				
N	1,423	1,419	1,419	1,422
Estimated effect at average	-0.126	-0.141**	-0.504***	-0.598***
	(0.096)	(0.071)	(0.101)	(0.086)
Interaction (per $100)	0.004	0.003	-0.012	0.006
	(0.009)	(0.006)	(0.010)	(0.007)
K-12 Enrollment (demeaned, average = 275.72)				
N	1,526	1,505	1,525	1,504
Estimated effect at average	-0.171*	-0.170**	-0.550***	-0.626***
	(0.099)	(0.075)	(0.113)	(0.098)
Interaction	.0003	.0001	.0012***	.0010***
	(.0003)	(.0002)	(.0003)	(.0003)
% Black (demeaned, average = 44.16%)				
N	1,526	1,505	1,525	1,504
Estimated effect at average	-0.173*	-0.174**	-0.525***	-0.598***
	(0.099)	(0.073)	(0.118)	(0.106)
Interaction	-0.003	-0.003	-0.001	0.002
	(0.004)	(0.003)	(0.004)	(0.003)
Student-Teacher Ratio (demeaned, average = 13.96)				
N	1,526	1,505	1,525	1,504
Estimated effect at average	-0.166*	-0.168**	-0.523***	-0.602***
	(0.099)	(0.074)	(0.116)	(0.106)
Interaction	0.042	0.013	0.044	0.004
	(0.036)	(0.023)	(0.033)	(0.029)
Full-Time Equivalent (demeaned, average = 20.97)				
N	1,526	1,505	1,525	1,504
Estimated effect at average	-0.166*	-0.168**	-0.537***	-0.617***
	(0.098)	(0.074)	(0.110)	(0.100)
Interaction	-0.001	-0.0001	0.013***	0.011**
	(0.005)	(0.003)	(0.005)	(0.005)
School Year Days (demeaned, average = 178.83)				
N	1,526	1,505	1,525	1,504
Estimated effect at average	-0.166*	-0.167**	-0.530***	-0.604***
	(0.097)	(0.074)	(0.114)	(0.103)
Interaction	0.001	-0.002	0.011	0.003
	(0.018)	(0.012)	(0.014)	(0.012)
School Day Hours (demeaned, average = 7.089)				
N	1,526	1,505	1,525	1,504
Estimated effect at average	-0.165*	-0.167**	-0.518***	-0.596***
	(0.099)	(0.072)	(0.116)	(0.100)
Interaction	0.028	0.030	0.201	0.217*
	(0.148)	(0.113)	(0.154)	(0.119)
Total Instructional Hours (demeaned, average = 1268.27)				
N	1,526	1,505	1,525	1,504
Estimated effect at average	-0.166*	-0.168**	-0.523***	-0.601***
	(0.098)	(0.070)	(0.120)	(0.100)
Interaction	.0001	.0001	0.001	0.001*
	(0.001)	(0.001)	(0.001)	(0.001)

*Bootstrapped standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1*

[17] *Notes.* Bootstrapped standard errors account for clustering within riskset. All variables are demeaned. The interaction represents the effect estimate moving away from the mean. See Table 3 for note on how Full-Time Equivalent and Total Instructional Hours were calculated. All first-stage regressions satisfy Staiger and Stock's rule of thumb for a strong instrument (See Appendix).

Table 6.[18]

Variation in LATE estimates by school tuition, enrollment, teaching staff, demographics, and instructional days, by tercile

	ELA				Math			
	N	Effect Estimate	Standard Error	*P*-Value	N	Effect Estimate	Standard Error	*P*-Value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tuition								
Top 33%	238	0.145	0.351	0.679	238	0.026	0.325	0.935
Middle 33%	604	-0.213	0.133	0.110	603	-0.578***	0.151	0.000
Bottom 33%	581	-0.123	0.157	0.413	581	-0.646***	0.194	0.001
K-12 Enrollment								
Top 33%	607	-0.059	0.119	0.620	606	-0.291**	0.130	0.025
Middle 33%	671	-0.337*	0.175	0.054	672	-0.725***	0.172	0.000
Bottom 33%	248	-0.013	0.192	0.947	247	-0.601**	-0.265	0.023
% Black								
Top 33%	447	-0.403*	0.220	0.067	448	-0.750***	0.253	0.003
Middle 33%	943	-0.113	0.105	0.281	941	-0.414***	0.119	0.000
Bottom 33%	136	0.059	0.268	0.825	136	-0.606**	0.302	0.045
Student-Teacher Ratio								
Top 33%	633	0.022	0.152	0.885	632	-0.430***	0.157	0.006
Middle 33%	534	-0.464**	0.217	0.032	535	-0.781***	0.221	0.000
Bottom 33%	359	-0.141	0.139	0.313	358	-0.377**	0.188	0.045
Full-Time Equivalents								
Top 33%	603	-0.194	0.141	0.169	602	-0.360**	0.175	0.040
Middle 33%	677	-0.293**	0.144	0.041	678	-0.750***	0.143	0.000
Bottom 33%	346	0.283	0.379	0.456	245	-0.306	0.352	0.384
School Year Days								
Top 33%	131	0.219	0.279	0.433	227	-0.727	0.498	0.144
Middle 33%	1,168	-0.136	0.106	0.197	1,166	-0.495***	0.122	0.000
Bottom 33%	227	-0.663*	0.353	0.060	132	-0.477	0.294	0.105
School Day Hours								
Top 33%	525	-0.176	0.164	0.282	525	-0.365**	0.172	0.034
Middle 33%	688	-0.114	0.138	0.409	688	-0.639***	0.196	0.001
Bottom 33%	313	-0.253	0.237	0.285	312	-0.577***	0.218	0.008
Total Instructional Hours								
Top 33%	691	-0.082	0.142	0.564	690	-0.350**	0.149	0.018
Middle 33%	346	-0.406**	0.205	0.048	346	-0.867***	0.292	0.003
Bottom 33%	489	-0.127	0.182	0.484	489	-0.543***	0.173	0.002

*Bootstrapped standard errors calculated. *** p < 0.01, ** p < 0.05, * p < 0.1.*

[18] *Notes.* Because terciles are calculated at the student level, terciles are not equally weighted with observations. Thus, statistically non-significant, yet substantial point estimates can be observed. Note that the variance of the number of days per school year is particularly small. Effect estimates were calculated by running each model exclusively for students within a tercile. All first-stage regressions satisfy Staiger and Stock's rule of thumb for a strong instrument (See Appendix).

whereby students attending private schools in the top tuition tercile appear to have experienced impacts on math that are both statistically and substantively miniscule, in contrast with the large negative effects for schools in the lower two terciles.

Total K-12 enrollment is positively associated with math outcomes, as using a voucher to attend a school with an additional 100 students is associated with about a tenth of a standard deviation gain on math outcomes compared to LSP students attending a school of average enrollment. Given the range of total K-12 enrollment in our sample (see Table 3), the difference in expectation of a school at the top of the distribution compared to one at the bottom can be substantial, similar to our findings for full-time equivalent teachers. No statistically significant effects are detected for ELA outcomes in either of our models. Similarly, outcomes did not vary to a significant degree according to the proportion of the student population that is African-American; however there is evidence suggesting larger negative effects for students attending schools in the top tercile on that student demographic factor.

Generally speaking, we find that both ELA and math outcomes are positively associated with the size of the school, as measured by total K-12 enrollment and number of full-time equivalent teachers on staff. The relationship between student outcomes and a school's student/teacher ratio is less clear. As staffing is correlated with total K-12 enrollment, the wide range of number of teachers per school in our sample similarly suggests that the difference in expectation of a school at the top of the distribution compared to one at the bottom can be quite large. No statistically significant differences are detected on ELA outcomes, though more school employees are associated with better math impacts in our preferred model.

In examining the intensity of schooling, we consider the number of days in the school year, the number of hours in the school day, and total instructional hours—a function of the

former two characteristics. With two exceptions, no statistically significant effects are detected for either ELA or math outcomes across all of our models. Both the number of hours in a school day and total instructional hours in a school year were estimated to have a positive and statistically significant effect on math outcomes in our school characteristics model. Assuming a 180-day school year, LSP students attending a school with an additional hour of instruction per day experience math impacts that are roughly a fifth of a standard deviation less negative than students attending other LSP schools, holding all else equal ($p < 0.1$). Using a voucher to attend a school with an additional hour in the school day above the mean is associated with a 0.217 standard deviation gain in math outcomes, holding all else equal ($p < 0.1$). No consistent pattern emerges according to the number of school days in the year, but this may be due to lack of variation, as the vast majority of LSP students in our sample attended a school in the middle tercile of the distribution.

Conclusion

This paper examines how impacts of the Louisiana Scholarship Program (LSP), a statewide school voucher program, vary across participating private schools. This examination is especially important in the case of the LSP, in which early evaluations of the program revealed startlingly large negative effects of the program on student achievement after only one year of participation (Abdulkadiroglu et al., 2018; Mills, 2015). This paper helps provide a comprehensive understanding of the LSP's impacts on participating students by determining the extent to which these effects vary across private school types.

Specifically, we examine how treatment effects vary across school characteristics such as religious identity, urban setting, tuition, enrollment, staffing, and student demographics. Broadly speaking, we find limited evidence of variation in treatment effects related to school

characteristics. There are, however, noticeable exceptions. The most favorable effect estimates were found in religious (particularly non-Catholic) schools, urban schools, and schools in the top tercile of tuition, K-12 enrollment, and instructional hours. Moreover, it is possible that the generally insignificant differences partially reflect low statistical power, as several differences in effect estimates that are not statistically significant are nonetheless large in magnitude.

We stress, however, that this analysis is purely exploratory. Families do not select schools at random, thereby implying school characteristics are not exogenously determined. In fact, our finding that LSP-participating schools with higher enrollments generated significantly better (or less bad) initial voucher impacts than LSP-participating schools with lower enrollments implies that a stampede to higher-quality (or at least away from lower-quality) by choosing parents is producing a connection between more popular and more effective private schools. Evaluating the causal impact of variation in school characteristics on student outcomes would require random assignment of school characteristics to students, an action that is both implausible and likely at odds with the purported benefits of school choice.

Nevertheless, the treatment effects underlying our analysis are exogenously determined, thereby giving us strong assurance that we are describing their correlations with school features accurately. As it stands, our analysis indicates treatment effects do not generally differentiate across school characteristics, with a few tantalizing exceptions.

References

Abdulkadiroglu, A., Pathak, P. A., & Roth, A. E. (2005). The New York City high school match. *American Economic Review, Papers and Proceedings*, *95*, 364-367.

Abdulkadiroglu, A., Pathak, P. A., & Walter, C. R. (2016). Free to choose: Can school choice reduce student achievement? NBER. Accessed at: http://www.nber.org/papers/w21839.pdf.

Abdulkadiroglu, A., Pathak, P. A. & Walters, C. R. (2018). Free to choose: Can school choice reduce student achievement? *American Economic Journal: Applied Economics*, *10*(1): 175-206.

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Angrist, J.R., Imbens, G., and Rubin, D. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91(434):444-455.

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*, 299-323.

Bitler, M. P., Domina, T., Penner, E. K., & Hoynes, H. W. (2015). Distributional effects of a school voucher program: Evidence from New York City (*Journal of Research on Educational Effectiveness*, *8*(3), 419-450.

Bloom, H. S. & Unterman, R. (2014). Can small schools of choice improve educational prospects for disadvantages students? *Journal of Policy Analysis and Management*, *33*(2): 290-319.

Coleman, J. S., Hoffer, T., & Kilgore, S. (1981). Public and private schools. An analysis of high school and beyond: A National Longitudinal Study for the 1980's. Report to the National Center for Education Statistics. National Opinion Research Center. Chicago: National Opinion Research Center. Accessed at: https://files.eric.ed.gov/fulltext/ED214314.pdf.

Cowen, J. M. (2008). School choice as a latent variable: Estimating the complier average causal effect of vouchers in Charlotte. *Policy Studies Journal*, 36, 301–315.Egalite, A. J. & Wolf, P. J. (2016). A review of the empirical research on private school choice, *Peabody Journal of Education*, 91:4, 441-454.

Dynarski, M., Rui, N., Webber, A., Guttman, B., & Bachman, M. (2017). Evaluation of the DC Opportunity Scholarship Program: Impacts after one year. U.S. Department of Education, Institute for Education Sciences, National Center for Education Evaluation and Regional Assistance, Washington, DC: U.S. Government Printing Office, NCEE 2017-4022.

Engberg, J., Gill, B., Zamarro, G., & Zimmer, R. (2011). Closing schools in a shrinking district: Do student outcomes depend on which schools are closed? *Journal of Urban Economics* 71 (2012), 189-203.

Figlio, D. & Karbownik, K. (2016). Evaluation of Ohio's EdChoice Scholarship Program: Selection, competition, and performance effects. Washington D.C.: Thomas B. Fordham Foundation. Available at http://edex.s3-us-west-2.amazonaws.com/publication/pdfs/FORDHAM%20Ed%20Choice%20Evaluation%20Report_online%20edition.pdf.

Friedman, M. (1955). The Role of Government in Education. In R. A. Solo (Ed.), *Economics and the Public Interest* (pp. 123–144). New Brunswick, NJ: Rutgers University Press.

Friedman Foundation for Educational Choice. (2015). School Choice: Louisiana. Retrieved December 27, 2015, from http://www.edchoice.org/school-choice/state/louisiana/

Frendeway, M., Sawatka, K., Marcavage, W., Carney, K., Martinez, K., & Dauphin, P. (2015). *School choice yearbook 2014–15: Breaking down barriers to school choice*. Washington, DC: Alliance for School Choice.

Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. New York, NY: W. W. Norton & Company.

Greene, J. P. (2001). Vouchers in Charlotte. *Education Matters*, *1*, 55–60.

Greene, J. P., Peterson, P. E., & Du, J. (1999). Effectiveness of school choice: The Milwaukee experiment. *Education and Urban Society*, 31(2), 191–213.

Howell, W. G., & Peterson, P. E. (with Wolf, P. J., & Campbell, D. E.) (2006). *The educational gap: Vouchers and urban schools* (Rev. ed.). Washington, DC: Brookings.

Howell, W. G., Wolf, P. J., Campbell, D. E., & Peterson, P. E. (2002). School vouchers and academic performance: Results from three randomized field trials. *Journal of Policy Analysis and Management*, 21, 191–217

Hoxby, C. M. (2003). School choice and school productivity could school choice be a tide that lifts all boats? In C. M. Hoxby (Ed.), *The economics of school choice* (287-341). Chicago, IL: University of Chicago Press.

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.

Kisida, B., Wolf, P. J., & Rhinesmith, E. (2015). *Views from private schools: Attitudes about school choice programs in three states*. Washington, DC: American Enterprise Institute.

Krueger, A. B., & Zhu, P. (2004). Another look at the New York City school voucher experiment. *American Behavioral Scientist*, 47, 658–698

Mill, J. S. (1962). Utilitarianism, on Liberty, Essay on Bentham. (Warnock, M. ed.) New York: Meridian.

Mills, J. N. (2015). *The effectiveness of cash transfers as a policy instrument in K-16 education* (Doctoral Dissertation). University of Arkansas, Fayetteville, AR.

Mills, J. N. & Wolf, P. J. (2017a). Vouchers in the bayou: The effects of the Louisiana Scholarship Program on student achievement after two years. *Educational Evaluation and Policy Analysis*. Available at http://journals.sagepub.com/doi/abs/10.3102/0162373717693108

Mills, J. N. & Wolf, P. J. (2017b). *The effects of the Louisiana Scholarship Program on student achievement after three years* (Louisiana Scholarship Program Evaluation Report #7). New Orleans, Louisiana: Education Research Alliance for New Orleans. Available at: http://educationresearchalliancenola.org/publications/the-effects-of-the-louisiana-scholarship-program-on-student-achievement-after-three-years

Paine, T. (1791). The Rights of Man: Answer to Mr. Burke's Attack on the French Revolution. London: J. S. Jordan.

Rouse, C. E. (1998). Private school vouchers and student achievement: An evaluation of the Milwaukee Parental Choice Program. *Quarterly Journal of Economics*, 113, 553–602.

Schultz, T., Carney, K., Marcavage, W., Jackson, N., Clements, E., Dauphin, P., Martinez, K. (2017). *School choice yearbook 2016–17.* Washington, DC: Alliance for School Choice.

Shakeel, M.D., Anderson, K. P., & Wolf, P. J. (2016). The participant effects of private school vouchers across the globe: A meta-analytic and systematic review. Social Science Research Network, EDRE Working Paper 2016-07, May 10, retrieved from: http://www.uaedreform.org/downloads/2016/05/the-participant-effects-of-private-school-vouchers-across-the-globe-a-meta-analytic-and-systematic-review-2.pdf.

Staiger, D. & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, *65*(3): 557-586.

Sude, Y., DeAngelis, C.A., & Wolf, P.J. (2017). Supplying choice: An analysis of school participation decisions in voucher programs in Washington, DC, Indiana, and Louisiana. *Journal of School Choice*. Accessed at: http://www.tandfonline.com/doi/abs/10.1080/15582159.2017.1345232.

Waddington, R. J. & Berends, M. (2017). Impact of the Indiana Choice Scholarship Program: Achievement effects for students in upper elementary and middle school. South Bend, IN: Center for Research on Educational Opportunity, Notre Dame. Available at http://creo.nd.edu/images/people/Waddington__Berends_Indiana_Voucher_Impacts_06.24.17.pdf

Wolf, P. J. (2008). School voucher programs: What the research says about parental school choice. *Brigham Young University Law Review*, 2008, 415–446.

Wolf, P. J. & Hoople, D. S. (2006). What school factors explain voucher gains in Washington, DC? *Peabody Journal of Education*, 81(1), 7-26.

Wolf, P. J., Gutmann, B., Puma, M, Kisida, B., Rizzo, L., & Eissa, N. O. (2009) Evaluation of the DC Opportunity Scholarship Program: Impacts after three years, U.S. Department of Education, Institute for Education Sciences, National Center for Education Evaluation and Regional Assistance, Washington, DC: U.S. Government Printing Office, NCEE 2009-4050, March. http://ies.ed.gov/ncee/pubs/20094050/

Wolf, P. J., Kisida, B., Gutmann, B., Puma, M, Eissa, N. O., & Rizzo, L., (2013) School vouchers and student outcomes: Experimental evidence from Washington, DC. *Journal of Policy Analysis and Management*, 32, 246-270.

Appendix

Table 1

First stage F-statistics

	Simple Model		School Characteristics Model	
	N	*F*-statistic	N	*F*-statistic
	(1)	(2)	(3)	(4)
ELA				
Religious	1,796	341.82	1,773	56.28
Catholic	1,796	362.94	1,773	60.75
Religious Non-Catholic	1,796	412.79	1,773	74.41
Coed	1,796	346.82	1,773	57.35
Urban	1,796	347.41	1,773	57.81
Tuition	1,423	54.78	1,419	13.76
K-12 Enrollment	1,526	238.34	1,505	53.33
Student/Teacher Ratio	1,526	79.26	1,505	21.36
Full-Time Equivalent	1,526	317.72	1,505	57.19
School Year Days	1,526	47.99	1,505	15.19
School Day Hours	1,526	83.29	1,505	24.2
Total Instructional Hours	1,526	296.86	1,505	50.43
% Black	1,526	184.49	1,505	38.21
Math				
Religious	1,796	339.87	1,773	55.95
Catholic	1,796	359.97	1,773	60.15
Religious Non-Catholic	1,796	407.02	1,773	73.58
Coed	1,796	344.72	1,773	57.07
Urban	1,796	345.34	1,773	57.33
Tuition	1,422	53.55	1,418	13.25
K-12 Enrollment	1,525	234.58	1,504	54.00
Student/Teacher Ratio	1,525	77.57	1,504	17.57
Full-Time Equivalent	1,525	315.7	1,504	58.35
School Year Days	1,525	50.15	1,504	18.01
School Day Hours	1,525	83.02	1,504	21.51
Total Instructional Hours	1,525	294.62	1,504	50.36
% Black	1,525	183.62	1,504	35.55

Notes. Staiger and Stock (1997) recommend a threshold for instrument relevance of joint-F statistics of 10 or higher.
Source. Authors' calculations.

Table A2.

Distribution of school characteristics for analytical sample

	Range	Bottom 33%	Mean	Top 33%
	(1)	(2)	(3)	(4)
Tuition	$2,200-$14,500	$5,074	$5,712	$6,236
K-12 Enrollment	26-912	27.70	203.67	98.89
Student/Teacher Ratio	3.80-22.92	10.99	13.96	14.35
Full-Time Equivalent	3-86.1	8.35	20.97	14.41
School Year Days	151-223	177	178	180
School Day Hours	5.50-9.00	7.01	7.21	7.34
Total Instructional Hours	973.50-2007.00	1,243.54	1,268.27	1,278.44
% Black	0.24-100.00	18.19	34.19	58.39

Notes. Terciles calculated at the school level. Mean calculated at the school level. See Table 3 for note on how Full-Time Equivalent and Total Instructional Hours were calculated.