



UNIVERSITY OF
ARKANSAS

College of Education & Health Professions
Education Reform

WORKING PAPER SERIES

HETEROGENEOUS IMPACTS ACROSS SCHOOLS IN THE FIRST FOUR YEARS OF THE LOUISIANA SCHOLARSHIP PROGRAM

Matthew H. Lee, Jonathan N. Mills, and Patrick J. Wolf

April 23, 2019

EDRE Working Paper 2019-11

The University of Arkansas, Department of Education Reform (EDRE) working paper series is intended to widely disseminate and make easily accessible the results of EDRE faculty and students' latest findings. The Working Papers in this series have not undergone peer review or been edited by the University of Arkansas. The working papers are widely available, to encourage discussion and input from the research community before publication in a formal, peer reviewed journal. Unless otherwise indicated, working papers can be cited without permission of the author so long as the source is clearly referred to as an EDRE working paper.

**HETEROGENEOUS IMPACTS ACROSS SCHOOLS IN THE FIRST FOUR YEARS OF
THE LOUISIANA SCHOLARSHIP PROGRAM**

Matthew Lee*
Department of Education Reform
University of Arkansas
Fayetteville, AR 72701
mhl002@uark.edu

Jonathan N. Mills
Department of Education Reform
University of Arkansas
Fayetteville, AR 72701
jnm003@uark.edu

Patrick J. Wolf
Department of Education Reform
University of Arkansas
Fayetteville, AR 72701
pwolf@uark.edu

Louisiana Scholarship Program Evaluation Report #11

School Choice Demonstration Project, University of Arkansas, Fayetteville AR

Acknowledgements

We thank Gema Zamarro, Joe Waddington, Li Hao, Austin Nichols, Ben Scafidi, and others for their extensive feedback on earlier drafts of this work, the Louisiana Department of Education for their cooperation and assistance with providing the necessary data to conduct the analyses, and Yujie Sude for her help in collecting data beyond what was reported in the Private School Universe Survey. We thank the Smith Richardson Foundation for the grant support that made this research possible. We are grateful to Kathleen Wolf for editorial help. The content of this report is solely the responsibility of the authors and does not necessarily represent the views of the University of Arkansas, the Louisiana Department of Education, or the Smith Richardson Foundation.

* 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 initiative 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-based reforms such as the LSP should improve student outcomes, experimental evaluations of the program instead find significant negative effects of the program on math and reading scores after its first year. Those effects diminish to insignificant differences by the end of the third year before becoming negative again in the fourth year. Our study builds on previous work with an exploratory analysis of the variation in treatment effects across 13 school characteristics in the first four years of the program. In general, we do not observe effect heterogeneity across school characteristics, though we find evidence suggesting students who preferred larger schools, schools with higher tuition, and schools with longer school days experienced more favorable impacts from participating in the LSP relative to their peers who did not prefer such schools.

Keywords: school vouchers, school choice, student achievement, heterogeneous effects, mediators of education effects

HETEROGENEOUS IMPACTS ACROSS SCHOOLS IN THE FIRST FOUR YEARS 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 that school voucher or scholarship programs expand educational options available to families, thereby optimizing the possibility of a match between students' educational needs with school offerings, all while improving the education system as a whole through increased competition among schools for students (Chubb & Moe, 1990; Hoxby, 2003). Opponents argue choice programs harm traditional public schools and their students by stripping away funding and concentrating disadvantaged, non-choosing students within their walls (Gutmann, 2003; Lubienski & Lubienski, 2013).

However, the evidence on the achievement effects of private school voucher programs¹ is mixed, with slightly more studies reporting positive results than negative. Overall, the most rigorous empirical research indicates null or small positive impacts of vouchers on student achievement with noticeable, yet non-systematic variation in treatment effects across student subgroups as moderators (Egalite & Wolf, 2016; Shakeel, Anderson, & Wolf, 2016). Lesser known is the extent to which school characteristics mediate treatment effects. After all, school choice is a general policy; the particular form that school choice takes for students and families depends upon the specific schools chosen. In this paper, we address this question by examining how treatment effects vary across schools participating in one of the nation's first statewide school voucher initiatives, the Louisiana Scholarship Program (LSP).

¹ There are other forms of private school choice, including tax credit scholarships, education savings accounts, and tuition rebates. However, as the Louisiana Scholarship Program is a voucher program, our primary focus for the literature review will be on private school voucher programs. For more details on other forms of private school choice, we encourage the reader to consider EdChoice's 2019 volume, *The ABCs of School Choice*.

This paper builds on the private school choice literature by examining how LSP impacts vary across school environments. Specifically, we conduct an analysis focused on students experiencing the initial statewide expansion: the 2012-13 application cohort. Following the work of Mills (2015) and Mills and Wolf (2017a; 2017b; 2019), our analysis is restricted to eligible applicants who experienced an oversubscription lottery in order to calculate unbiased estimates of the impact of LSP voucher usage on student achievement. We then explore how these estimated treatment effects vary across 13 school characteristics, which include the school's religious affiliation, geographic location, tuition, enrollment, demographics, staffing, and instructional intensity. We also conduct tiered analysis for nine of those characteristics. Tiered analysis groups students based on key school characteristics that can be measured continuously as opposed to just categorically.

In general, we do not find much evidence of variation in LSP impacts on English Language Arts (ELA) and math outcomes. In terms of school setting, we find some evidence that students preferring schools in an urban setting experienced favorable math impacts relative to students preferring schools in a suburban, town, or rural setting. Tuition charged by the private school, the number of full-time teachers or equivalents employed by the school, and the total K-12 enrollment of the school tended to be positively associated with math outcomes. Schools enrolling a larger percentage of African American or LSP students had less favorable math outcomes in the first year relative to schools with the smallest proportions of these students. Finally, we find that math outcomes after the first year were positively associated with the number of instructional hours per school day or per school year, but not with the number of school days per year. These findings deepen our understanding of how the LSP impacted both students and the education system in Louisiana as a whole.

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, Illinois, Indiana, Nevada, and Ohio, in addition to Louisiana. Given this expansion, it is important to understand how these programs work on a state level, as we expect to see increased variation in school characteristics across an entire state when compared to programs concentrated in a single city. Understanding how this variation relates to voucher effects can provide guidance to future policymakers.

Second, our analysis builds on a highly rigorous causal design. Placement in an LSP school was determined by random lottery, so effect estimates for students applying to a first-choice school with a shared characteristic (e.g. Catholic or non-Catholic) support strong causal inferences. Thus, our study is the first analysis to provide scholars and policy makers with causal estimates of the impact of attending specific types of private schools with the aid of a school voucher. However, results regarding effect heterogeneity are merely exploratory, as such estimates are endogenous with unobserved characteristics that led students to sort themselves into particular first-choice school lotteries (e.g. Catholic relative to non-Catholic).

Our report proceeds as follows. First, we summarize the theories and existing literature on achievement impacts of school voucher programs in the United States in general and their heterogeneous effects in particular. Next, we describe the structure of and research on the Louisiana Scholarship Program. We then outline our methodology, describing our sample of students and analytical strategy. After presenting the results of our analysis, we conclude with a discussion of our findings and their implications for the program.

Relevant Theoretical and Empirical Literature on School Choice

An extensive scholarly literature exists on the theory behind and effects of private school choice programs. Here we highlight a few works that are most relevant to this specific study.

Theories of school choice

Political philosophers Thomas Paine (1791) and John Stuart Mill (1962 [1869]), as well as economist Milton Friedman (1955), developed the theory backing K-12 school choice interventions. A primary claim of school choice proponents is that, while government should offer funds in support of compulsory education mandates, it does not need to deliver the education itself (Friedman, 1955). School choice theory posits that choice will improve student academic outcomes by allowing families to seek out the schools that best meet their child's needs and by incentivizing schools to compete for students (Friedman, 1955; Moe, 2005).

Other scholars theorize that private school programs will have negative effects on student achievement. They claim that public schools have a comparative advantage over private schools in boosting test scores (Lubienski & Lubienski, 2013); that parents will make bad schooling choices for their children (Lauder & Hughes, 1999; Smith & Meier, 1995); or, that private schools are especially ineffective at educating the kinds of disadvantaged students targeted by voucher programs (Fuhrer 2013). These competing claims about the achievement effects of school vouchers amount to testable hypotheses.

Empirical evidence on the effects of voucher programs on student achievement

School voucher programs, in which students receive publicly-funded vouchers or scholarships to attend a participating private school of their choosing, represent one form of private school choice (Wolf, 2008). Egalite and Wolf (2016) note voucher programs can differ

along at least five dimensions: region served (cities or entire states), eligibility (means-tested or universal), level of regulation, voucher value, and number of vouchers available. As of January 2019, there were 56 private school choice programs in the United States (EdChoice, 2019).² A majority of these are means-tested and 19 are primarily or exclusively targeted to students with disabilities. Most of the programs open to general education students are operated at the local rather than the state level.

The most rigorous research focusing on the effects of voucher programs on student achievement reports mixed results, though results are more often positive than negative. A recent meta-analysis of all experimental evaluations of U.S. programs indicates small positive effects of vouchers on student math and reading achievement when averaged across all programs that are not statistically significant (Shakeel, Anderson & Wolf, 2016). This overall finding masks considerable heterogeneity across programs, with evaluations 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) finding statistically significant gains in ELA, math, or both subjects. In contrast, experimental or quasi-experimental evaluations of voucher programs in Louisiana (Abdulkadiroglu, Pathak, & Walters, 2018; Mills & Wolf, 2019), Washington, D.C. (Wolf et al., 2013; Dynarski, Rui, Webber, Guttman, & Bachman, 2017), Indiana (Waddington & Berends, 2018), and Ohio (Figlio & Karbownik, 2016) report statistically significant negative impacts of voucher programs on student achievement, particularly in math.

The achievement effects of voucher programs vary not only across programs, but also within programs. Several studies report differential effects by ethnicity (Howell, Wolf,

² We exclude personal tax deduction/credit programs from this count because they are individual tax benefits and not "programs."

Campbell, & Peterson, 2002; Barnard, Frangakis, Hill, & Rubin, 2003; Howell & Peterson, 2006; Jin, Barnard, & Rubin, 2010), ability (Bitler, Domina, Penner, & Hoynes, 2015), or quality of sending school (Wolf et al., 2013). Krueger and Zhu (2004) is the rare case of a voucher achievement study that finds no evidence of effect heterogeneity across any student subgroups. Effects can also vary over time (Wolf & Egalite, forthcoming), either diminishing (Howell & Peterson, 2006; Wolf et al., 2009; Wolf et al., 2013) or improving with time (Witte et al., 2014; Mills & Wolf, 2017b; Waddington & Berends, 2018).

Effect heterogeneity across school characteristics

The above studies all examine how student characteristics might “moderate” school choice achievement effects. Lesser known, however, is how voucher treatment effects vary across school characteristics defined as “mediators.” 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). Therefore, it seems plausible to expect increased variation in school characteristics in a system with school choice. The extent to which increased variation in school characteristics affects student achievement patterns in specific programs, as well as in general, is less clear.

Unfortunately, the research base examining how private school characteristics mediate student academic outcomes is limited. There is some evidence suggesting different types of private schools can impact student outcomes. Coleman, Hoffer, and Kilgore (1981) find that students attending Catholic schools were more likely to graduate or enroll in college than their public school peers, especially if they belonged to a racial minority group. 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 voucher effects. Waddington and Berends (2018) 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 only small differences.

Effects may also vary with school enrollment. Low performance, low enrollment, and greater probability of a private school closing tend to be correlated with each other (McShane, Kisida, Jensen, & Wolf, 2012). 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 schools that would have otherwise closed in the absence of the program. Declining enrollment prior to the start of the voucher program 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 been rarely examined. Our analysis adds to this scant research by examining if voucher effects vary systematically across types of schools participating in the LSP.

The Louisiana Scholarship Program Design and Prior Research

Initially created as a pilot program in New Orleans in 2008, the Louisiana State Legislature expanded the program statewide for the 2012-13 school year by passing Act 2 in 2012. The program is limited to students (1) with family income at or below 250 percent of the federal poverty line who also are (2) entering kindergarten or attending a public school that was graded C, D, or F for the prior school year according to the state's school accountability system or a school in the Recovery School District (RSD). During the years covering this report, 2012-

13 through 2015-16, the RSD included 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 governments provide in student funding to the local school system or the tuition charged by the student's chosen private school, whichever is less. Tuition at participating private schools ranged from \$2,200 to \$14,500, with a mean of \$5,494 and a median of \$5,045 in the first year, substantially lower than the average per pupil revenue of \$12,220 in Louisiana's traditional public schools. Participating private schools must accept the voucher as full value for tuition, even if the value of the voucher is less than the tuition charged.

Private schools must meet certain criteria to participate in the program involving admissions, 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 regulations have influenced schools' participation decisions (Kisida, Wolf, & Rhinesmith, 2015). Experimental studies in Florida (DeAngelis, Burke, & Wolf, 2018) and California and New York (DeAngelis, Burke, & Wolf, 2019) find that open-enrollment mandates and state standardized testing requirements reduce the likelihood that a private school leader is interested in participating in a voucher program. Taken together, these studies may explain why only a third of eligible private schools opted into the program in 2012-13. Private school participation in the LSP has increased slightly since that time.

Eligible applicants to the 2012-13 cohort could list 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 placement priorities.³

In cases of oversubscription to a specific school in the program,⁴ the LSP matching algorithm randomly assigns students to receive or not to receive an LSP private school placement to that particular school. Recent evaluations of the Louisiana Scholarship Program use these oversubscription lotteries for students' first-choice schools to estimate the impact of LSP scholarship usage on student achievement. Separate studies examining achievement impacts after one year report statistically significant negative impacts of voucher usage on student achievement in reading, math, science, and social studies (Abdulkadiroglu, Pathak & Walters, 2018; Mills, 2015). These negative effects diminish over time, with math effects roughly halved after two years (Mills & Wolf, 2017a) and not statistically significant after three years (Mills & Wolf, 2017b). However, effect estimates on both ELA and math become negative again in the fourth year (Mills & Wolf, 2019). The results are heterogeneous based on moderator characteristics, with negative effects persisting for younger students in math and statistically

³ The LSP scholarships are awarded according to the following guidelines. First, students with disabilities and "multiple birth siblings," siblings who are 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.

⁴ 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).

significant positive effects observed for students performing in the bottom third of the ELA distribution at baseline (Mills & Wolf, 2017b).

Methodology

We focus in this paper on describing how the impact on achievement after the first four years of using an LSP scholarship to attend a private school varied across different school settings for the 2012-13 cohort. In doing so, we build on two components of an ongoing evaluation of the LSP: studies examining the impact of the program on participant achievement (Abdulkadiroglu et al., 2018; Mills, 2015; Mills & Wolf 2017a; Mills & Wolf, 2017b; Mills & Wolf, 2019) and Sude, DeAngelis, and Wolf's (2018) 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), and subsequent partial recovery of that lost ground, it is important to determine if and how school characteristics mediated these outcomes.

Data

The data for this analysis come from several sources. The Louisiana Department of Education (LDE) provided student-level demographic data, testing data,⁵ and application information for all eligible LSP applicants.⁶ The LDE additionally provided information on LSP

⁵ This study uses student performance on the 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), 2012-13, and 2013-14 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). In 2014-15, Louisiana instead administered Partnership for Assessment of Readiness for College and Careers (PARCC), a criterion-referenced test aligned with the Common Core standards, in ELA and math. Our year three results represent student performance on these assessments. For more information, see Mills and Wolf (2017b).

⁶ For more information, see Mills and Wolf (2017a).

participating schools, which we have supplemented with data from the Private School Universe Survey (PSS) and reviews of school websites.⁷

Analytical strategy

Our analysis builds on the work of Mills (2015) and Mills and Wolf (2017a; 2017b; 2019) which leverage oversubscription lotteries occurring during the process of matching LSP applicants to schools for the 2012-13 school year. The analyses use the outcome of oversubscription lotteries as an instrument to predict actual enrollment in an LSP school, and use predicted enrollment to produce unbiased estimates of the program's Local Average Treatment Effect (LATE) on student achievement (Angrist & Pischke, 2009; Cowen, 2008; Mills & Wolf, 2017b; Mills & Wolf, 2019). Our analysis expands on this work by interacting predicted LSP enrollment with school characteristics in the original evaluation's two stage least squares (2SLS) model:

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

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

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

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;⁸

⁷ The PSS is a biennial survey intended to collect data on all private schools in the United States meeting the National Center for Education Statistics definition of private schools. While the intent is to be comprehensive, the survey does not include data for all private schools in Louisiana or participating in the LSP. We use data from the 2013-14 PSS to conduct our analysis. For more information, see Sude, DeAngelis, and Wolf (2018).

⁸ 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;⁹
- T_i indicates if a student received an LSP scholarship to their first-choice school;
- A_i is standardized student mathematics or ELA achievement in year one, two, three, or four of the program;¹⁰
- X_i' is a vector of student characteristics, including achievement, collected either at baseline (2011-12) or from the student's LSP application form; and
- 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 LATEs across school characteristics (2).¹¹ The estimate τ_1 indicates the experimental impact of using an LSP scholarship to attend a private school that does not have characteristic S_i . Adding τ_1 and τ_2 provides the experimental impact of using an LSP scholarship to attend a private school that does have characteristic S_i . τ_2 itself estimates the difference in the LSP effect between students randomly placed in a private school with

⁹ We include a fixed effect for first-choice school 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 first-choice school 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.

¹⁰ Student achievement scores are standardized using distributional parameters of outcomes from the control group.

¹¹ 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).

characteristic S_i and students randomly placed in a private school without characteristic S_i . The significance test on τ_2 determines if the impact of the LSP is heterogeneous or not with respect to characteristic S_i . We additionally account for nesting of students within lotteries with cluster-adjusted 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 school in 2012-13. Our analytical sample is restricted to students with baseline achievement data in grades three through seven in 2011-12 who additionally had outcome data in grades four through eight in 2012-13, 2013-14, 2014-15, or 2015-16. Table 1 presents descriptive statistics for the 899 LSP applicants meeting these criteria for our Year 4 analysis. Nearly 90 percent of students in our analytical sample are identified as African American. The overwhelming majority of students in our sample 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 at baseline, when they applied to the program. These patterns are similar in the first three years of our analysis.

The adjusted difference (column 4) 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) are very similar on nearly every characteristic included to those who do not (Control). This pattern gives us strong assurance that the LATEs underlying our mediator analysis are calculated with high internal validity.

Characteristics of schools participating in the LSP

This report examines how LSP achievement effects vary for this analytical sample across schools. Research by Sude, DeAngelis, and Wolf (2018) indicates private schools opting to participate in the LSP had lower enrollment, lower 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 for the 2012-13 school year respectively, is consistent with their analysis.

Overall, we observe several statistically significant differences between Louisiana private schools choosing to participate in the LSP and those choosing not to participate. In 2012-13, private schools that chose to participate in the LSP relative to private schools that chose not to participate were more likely to be religious (10.8 percentage points), Catholic (22.6), and coeducational (7.5). They charged almost \$850 less for tuition and staffed 7.5 fewer full-time equivalents on average. In terms of enrollment characteristics, LSP schools were smaller (87.6 students) and enrolled a greater percentage of African American students (32 percentage points). Their school year was slightly longer (1.7 days), but the length of the school day was roughly the same.

Table 1.

Characteristics of LSP applicants for Year 4 analysis

	<i>n</i>	Winners	Losers	Adjusted Difference (2) – (3)	Standard Error	<i>p</i> value
	(1)	(2)	(3)	(4)	(5)	(6)
Female	899	0.54	0.49	0.03	0.04	0.53
<i>Ethnicity</i>						
African American	899	0.87	0.89	-0.00	0.03	0.87
Hispanic	899	0.02	0.03	-0.01	0.01	0.62
White	899	0.06	0.06	0.00	0.02	0.97
Other	899	0.04	0.03	0.01	0.01	0.43
LEP	899	0.02	0.01	-0.00	0.01	0.86
FRL	730	0.94	0.95	-0.00	0.02	0.97
# of choices	899	1.89	2.34	-0.18*	0.10	0.07
<i>Baseline Tests^a</i>						
ELA	899	-0.33	-0.35	-0.01	0.08	0.93
Math	898	-0.34	-0.36	-0.03	0.08	0.66
Science	898	-0.45	-0.45	0.02	0.06	0.77
Social Studies	898	-0.37	-0.37	-0.01	0.07	0.92

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Scores are standardized within grade based on the observed distributions of scale scores across Louisiana. 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 four. *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.

Table 2.

Characteristics of LSP participating schools, 2012-13

	LSP		Non-LSP		Adjusted Difference (2) - (4)	Standard Error	p value
	N	Mean	N	Mean			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Religious	87	96.6%	197	85.8%	10.8***	0.04	0.007
Catholic	87	67.8%	197	45.2%	22.6***	0.06	0.000
Non-Catholic Religious	87	28.7%	197	40.6%	-11.9*	0.06	0.056
Coeducational	87	98.9%	197	91.4%	7.5**	0.03	0.017
Urban	87	49.4%	197	42.1%	7.3	0.06	0.256
Tuition	83	\$5,385.74	157	\$6,235.65	-\$849.91*	432.01	0.050
Full Time Equivalents	87	21.0	196	28.5	-7.5***	2.70	0.005
Student/Teacher Ratio	87	13.0	196	12.6	0.4	0.55	0.465
Total K-12 Enrollment	87	276	197	363	-87.6***	33.08	0.009
School Year Days	86	178.8	197	177.2	1.7**	0.77	0.030
School Day Hours	87	7.1	197	7.0	0.1	0.06	0.159
Total Instructional Hours	86	1268.3	197	1240.4	27.9**	12.87	0.031
<i>Demographics</i>							
% Native American	87	0.15%	194	0.96%	-0.01	0.73	0.272
% Asian	87	1.87%	194	2.49%	-0.01	0.61	0.313
% Hispanic	87	2.31%	194	4.08%	-1.77*	1.07	0.099
% African American	87	44.16%	194	12.16%	32.00***	3.61	0.000
% Pacific Islander	87	0.05%	194	0.25%	-0.21**	0.09	0.031
% Two or more	87	0.98%	194	1.97%	-1.00	0.78	0.202

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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 who taught full time, plus 0.875 times the number of teachers who taught between at least ¾ time but less than full-time, plus 0.625 times the number of teachers who taught at least ½ time but less than ¾ time, plus 0.375 times the number of teachers who taught at least ¼ time but less than ½ time, plus 0.125 times the number of teachers who taught less than ¼ time.

Characteristics of schools enrolling LSP students

Next, we look to see if there is any variation in school characteristics among LSP participating private schools. Tables 3-6 examine the distributions of school characteristics for LSP schools in which one or more LSP students enrolled in the first four years of statewide expansion. 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 outliers in the set of LSP schools. Therefore, we recommend caution when generalizing findings for these categories.

Over the course of four years, the schools in which LSP students enrolled did not change much for most characteristics, with some notable exceptions. LSP students tended to enroll in larger and more expensive schools and in schools in which African American or LSP students composed a smaller share of the student population over the course of the first four years of the program. This trend is not driven by a dramatic shift in one part of the distribution as values at the mean, median, 33rd percentile, and 66th percentile generally shifted upwards over the three years.

Table 3.

Measures of central tendency and variation in schools in which LSP students enrolled, 2012-13

	n	μ	σ	Min.	Max.	Median	33rd percentile	66th percentile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious	86	0.97	0.18					
Catholic	86	0.67	0.47					
Non-Catholic Religious	86	0.29	0.46					
Coeducational	86	0.99	0.11					
Urban	86	0.50	0.50					
Tuition	82	\$5,394	1,788	\$2,200	\$14,500	\$5,045	\$4,538	\$5,581
Full-Time Equivalents	86	21.0	15.8	3.0	86.1	15.6	12.2	20.9
Student/Teacher Ratio	86	13.0	3.7	3.8	22.9	13.3	11.1	14.6
Total K-12 Enrollment	86	275	206	26	912	205	159	272
% African American	86	44.60%	0.41	0.24%	100.00%	26.80%	9.05%	84.32%
% LSP	87	25.19%	0.27	0.15%	100.00%	11.11%	4.88%	30.75%
School Year Days	85	178.8	6.4	151	223	178	177	180
School Day Hours	86	7.1	0.5	5.5	9	7	7	7.1
Total Instructional Hours	85	1,268.0	118.9	974	2,007	1,260	1,241	1,274

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, including 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 who taught full time, plus 0.875 times the number of teachers who taught between at least $\frac{3}{4}$ time but less than full-time, plus 0.625 times the number of teachers who taught at least $\frac{1}{2}$ time but less than $\frac{3}{4}$ time, plus 0.375 times the number of teachers who taught at least $\frac{1}{4}$ time but less than $\frac{1}{2}$ time, plus 0.125 times the number of teachers who taught less than $\frac{1}{4}$ 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.

Table 4.

Measures of central tendency and variation in schools in which LSP students enrolled, 2013-14

	<i>n</i>	μ	σ	Min.	Max.	Median	33rd percentile	66th percentile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious	88	0.94	0.23					
Catholic	88	0.65	0.48					
Non-Catholic Religious	88	0.30	0.46					
Coeducational	88	0.99	0.11					
Urban	88	0.51	0.50					
Tuition	84	\$5,454	1,748	\$2,825	\$14,500	\$5,191	\$4,597	\$5,605
Full-Time Equivalents	88	21.1	15.9	3.0	86.1	15.8	12.0	21.5
Student/Teacher Ratio	88	12.8	3.6	3.8	22.9	13.3	11.0	14.5
Total K-12 Enrollment	88	276	210	26	912	209	157	282
% African American	88	44.29%	0.41	0.24%	100.00%	24.68%	9.44%	81.63%
% LSP	84	21.18%	0.22	0.15%	78.46%	10.75%	5.60%	33.43%
School Year Days	87	179.1	5.6	161	223	178	177	180
School Day Hours	88	7.1	0.5	5.5	9	7	7	7.1
Total Instructional Hours	87	1,267.8	116.7	974	2,007	1,260	1,239	1,274

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, including 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 who taught full time, plus 0.875 times the number of teachers who taught between at least ¾ time but less than full-time, plus 0.625 times the number of teachers who taught at least ½ time but less than ¾ time, plus 0.375 times the number of teachers who taught at least ¼ time but less than ½ time, plus 0.125 times the number of teachers who 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.

Table 5.

Measures of central tendency and variation in schools in which LSP students enrolled, 2014-15

	<i>n</i>	μ	σ	Min.	Max.	Median	33rd percentile	66th percentile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious	87	0.94	0.23					
Catholic	87	0.66	0.48					
Non-Catholic Religious	87	0.29	0.46					
Coeducational	87	0.99	0.11					
Urban	87	0.52	0.50					
Tuition	83	\$5,470	1,752	\$2,825	\$14,500	\$5,200	\$4,613	\$5,621
Full-Time Equivalents	87	21.3	15.9	3.0	86.1	15.9	12.1	22.0
Student/Teacher Ratio	87	12.9	3.6	3.8	22.9	13.3	11.0	14.6
Total K-12 Enrollment	87	279	209	26	912	211	161	290
% African American	87	43.65%	0.40	0.24%	100.00%	23.76%	9.28%	79.28%
% LSP	83	16.86%	0.17	0.13%	60.21%	8.91%	4.30%	21.54%
School Year Days	86	179.0	5.7	161	223	178	177	180
School Day Hours	87	7.1	0.5	5.5	9	7	7	7.1
Total Instructional Hours	86	1,270.4	115.0	974	2,007	1,260	1,242	1,274

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, including 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 who taught full time, plus 0.875 times the number of teachers who taught between at least ¾ time but less than full-time, plus 0.625 times the number of teachers who taught at least ½ time but less than ¾ time, plus 0.375 times the number of teachers who taught at least ¼ time but less than ½ time, plus 0.125 times the number of teachers who 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.

Table 6.

Measures of central tendency and variation in schools in which LSP students enrolled, 2015-16

	n	μ	σ	Min.	Max.	Median	33rd percentile	66th percentile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious	81	0.95	0.22					
Catholic	81	0.65	0.48					
Non-Catholic Religious	81	0.30	0.46					
Coeducational	81	0.99	0.11					
Urban	81	0.53	0.50					
Tuition	78	\$5,510	1,762	\$2,825	\$14,500	\$5,222	\$4,639	\$5,654
Full-Time Equivalents	81	21.5	16.2	3.0	86.1	15.9	12.1	22.0
Student/Teacher Ratio	81	13.0	3.6	3.8	22.9	13.3	11.2	14.6
Total K-12 Enrollment	81	280	208	26	912	211	170	291
% African American	81	42.95%	0.40	1.72%	100.00%	23.76%	9.29%	72.68%
% LSP	78	13.34%	0.14	0.13%	56.31%	6.62%	3.69%	14.80%
School Year Days	80	179.1	5.8	161	223	178	177	180
School Day Hours	81	7.1	0.5	5.5	9	7	7	7.2
Total Instructional Hours	80	1,275.5	117.2	974	2,007	1,260	1,246	1,275

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, including 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 who taught full time, plus 0.875 times the number of teachers who taught between at least $\frac{3}{4}$ time but less than full-time, plus 0.625 times the number of teachers who taught at least $\frac{1}{2}$ time but less than $\frac{3}{4}$ time, plus 0.375 times the number of teachers who taught at least $\frac{1}{4}$ time but less than $\frac{1}{2}$ time, plus 0.125 times the number of teachers who taught less than $\frac{1}{4}$ 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.

A note on causality

Some of the results we are about to present are causal while others are merely correlational and therefore exploratory. Our subgroup effect estimates represent the causal impact of a student using an LSP voucher to attend a school with a specific characteristic compared to a student seeking access to a private school with that characteristic but being denied access to it through the lottery's results. Our estimates of LSP impacts by school characteristic subgroups are causal because they focus on students who faced lotteries for placement in their first-choice private school, with the assistance of an LSP voucher. Students were seeking enrollment in a specific school with particular features. The lottery therefore determined which students among eligible LSP applicants to a specific school could actually experience those particular school characteristics and which applicants would be placed in the randomized control group. The LSP placement lottery is the first private school choice lottery with which we are familiar that randomly assigned students to experience or not experience particular private school characteristics. Therefore, the extent to which those school features systematically predicted positive or negative LSP achievement outcomes represents experimental, and therefore causal, associations.

On the other hand, our results regarding impact heterogeneity based on variation in school characteristics are not causal. They are, instead, exploratory because any differences between the causal impacts of the LSP on students due to differences in specific school characteristics could be caused by the school feature but also could be due to different populations of students preferring schools of choice with different characteristics. Since a mediation study like this one cannot untangle those two effects, we make clear throughout the study that our assessment of the extent to which LSP impacts vary by school characteristics is

correlational and therefore exploratory. Comparisons across lotteries, e.g. comparing the effects for students who preferred religious schools relative to those who preferred nonsectarian schools, do not necessarily support causal inferences.

Results

We examine effect heterogeneity in the first four years of the Louisiana Scholarship Program's statewide operation, exploring how various school characteristics could behave as mediators for student outcomes on Louisiana's state assessments in ELA and math. Our findings generally do not indicate that LATEs varied strongly across school settings over those four years. Parents and students of their own volition choose which schools to list as their preferences rather than being assigned schools randomly, and because of this selection into different lotteries, we emphasize again that this analysis has a causal component and a descriptive component. The former, which relies on the gold standard randomization used in Mills and Wolf's (2019) four-year evaluation of the LSP, is limited to the LATE impact estimates of schools with particular characteristics. The latter centers on the differences between LATE estimates by school characteristics, which could be endogenous with unobserved student characteristics that lead students to sort themselves into different lotteries.

By the third and fourth year, many of our instruments do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument, a first-stage regression with an F -statistic equal to or greater than 10. The main reason for this limitation is diminishing rates of compliance with lottery assignment by the third and fourth year of the program (Sude & Wolf, 2019).

Furthermore, effect estimates should be interpreted conservatively given the heteroskedastic structure of our error term due to lottery clusters. Since Staiger and Stock (1997) advised a rule of thumb for instrumental variables analysis, further literature has been developed

for both homoskedastic (Stock & Yogo, 2005) and heteroskedastic settings (Montiel Olea & Pflueger, 2013; Montiel Olea, Pflueger, & Wang, 2013; Andrews, Stock, & Sun, 2018). This research encourages the calculation of effective F -statistics (F^{EFF}), as even our bootstrapping of the standard errors does not sufficiently correct for problems introduced by multiple endogenous regressors. However, no industry-wide consensus has been formed on how to deal with multiple endogenous regressors in a heteroskedastic setting with panel data. Therefore, we continue to report which instruments fail to satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument and encourage the reader to interpret those results with caution.

Effect estimates that rely on instruments that fail to satisfy Staiger and Stock's (1997) rule are shaded in gray. Our preferred models control for a vector of student characteristics. Full results, including results from the simple model, can be found in the Appendix Tables A9-A24.

Outcomes by religious status, coeducational status, and urban setting

Table 7 provides our first look at the extent to which LSP impacts are differentiated by religious status, coeducational status, and urban setting. The table presents results for both ELA (columns 1, 3, 5, and 7) and math (columns 2, 4, 6, and 8) outcomes. The estimates in this table rely on our "Student Characteristics" model, which controls for student-level demographics.

We begin by considering a school's religious status (sectarian or secular), coeducational status (coeducational or single-gender), and urbanicity (whether located in an urban setting or in one of three alternatives: suburb, town, or rural) and find only one statistically significant difference in the achievement impacts of the LSP across four years. In the second year, students who preferred schools in an urban setting scored 0.61 standard deviations higher on the standardized math exam relative to their peers who preferred a school in a suburb, town, or rural setting ($p < 0.05$). We do not find statistically significant differences for students preferring

religious schools to non-religious schools or Catholic schools to non-Catholic schools (including religious non-Catholic schools) on either ELA or math impacts of the LSP.

By the third year, most instruments fail to satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument, thus providing unreliable descriptive estimates of any impact differences based on the religious status, coeducational status, or urbanicity of preferred schools. For full results by religious status, coeducational status, and urban setting, please see Appendix Tables A9, A13, A17, and A21.

Outcomes by school tuition and staffing

Next, we consider variation in LATE estimates by school tuition and staffing characteristics (Table 8). Effects on math outcomes were positively associated with the number of full-time equivalents staffed by a school. Students who preferred schools with an additional 10 full-time equivalents above the mean (approximately 21 FTE over the four years of analysis) experienced LSP math impacts that were 0.11 standard deviations higher at the end of the first year than their peers who preferred a school at the mean ($p < 0.01$). That association between higher staffing levels of the preferred private school and better (i.e. less negative) math impacts persisted into the second year (0.11 standard deviations, $p < 0.10$) and grew in the third year (0.29, $p < 0.10$). By year four, the lottery instrument used in the first stage of the model was too weak of a predictor of voucher use to yield a reliable result.

Other differences in LSP impacts by school characteristics were either not statistically significant, and thus indistinguishable from zero, or relied on instruments that failed to satisfy Staiger and Stock's (1997) rule of thumb for instrument relevance. For full results by school

Table 7.

Variation in LATE estimates by religious status, coeducational status, and urban setting

		2012-13		2013-14		2014-15		2015-16	
		ELA	Math	ELA	Math	ELA	Math	ELA	Math
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Religious, relative to non-religious</i>	<i>n</i>	1,772	1,772	1,551	1,550	1,172	1,171	715	703
	τ_2	0.13 (0.21)	0.14 (0.28)	-0.06 (0.45)	0.15 (0.48)	0.43 (1.47)	0.32 (2.14)	1.10 (21.55)	-0.66 (50.51)
	μ	0.84	0.84	0.86	0.86	0.89	0.89	0.88	0.88
<i>Catholic, relative to non-Catholic</i>	<i>n</i>	1,772	1,772	1,551	1,550	1,172	1,171	715	703
	τ_2	0.01 (0.13)	-0.23 (0.18)	-0.11 (0.23)	-0.34 (0.28)	-0.13 (0.48)	-1.04* (0.61)	0.43 (1.23)	-0.78 (1.47)
	μ	0.52	0.52	0.55	0.55	0.57	0.57	0.58	0.59
<i>Coeducational, relative to single-gender</i>	<i>n</i>	1,772	1,772	1,551	1,550	1,172	1,171	715	703
	τ_2	0.21 (0.22)	0.15 (0.28)	0.07 (0.44)	0.36 (0.52)	0.54 (0.94)	0.27 (1.19)	1.47 (23.05)	-0.21 (61.35)
	μ	0.83	0.83	0.85	0.85	0.88	0.88	0.88	0.89
<i>Urban, relative to suburban, town, or rural</i>	<i>n</i>	1,772	1,772	1,551	1,550	1,172	1,171	715	703
	τ_2	0.00 (0.14)	0.21 (0.18)	-0.02 (0.23)	0.61** (0.26)	-0.11 (0.43)	-0.25 (0.62)	-0.31 (1.35)	-0.11 (2.41)
	μ	0.49	0.49	0.51	0.51	0.52	0.52	0.53	0.53

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. All effect estimates rely on "Student Characteristics" model. "Religious" schools include all Catholic schools, Christian but non-Catholic schools, Muslim schools, and Jewish schools. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Tables A1, A3, A5, and A7). For full results, see Appendix Tables A9, A13, A17, and A21.

tuition, number of full-time equivalents, and student/teacher ratio, please see Appendix Tables A10, A14, A18, and A22.

Outcomes by school enrollment and demographics

When exploring how LATE estimates vary across enrollment characteristics, we find evidence that math impacts were positively associated with the total enrollment of a school (Table 9). Students who preferred schools with an additional 100 students above the mean ($\mu = 275$ -280 students over the four years of analysis) experienced LSP impacts 0.10 standard deviations higher in math after one year than their peers who preferred a school at the mean ($p < 0.01$), an association that persisted into the second year (0.11 standard deviations, $p < 0.05$). The association grew in the third year (0.21, $p < 0.10$), though the estimate relied on a weak instrument (Stock & Staiger, 1997). By the fourth year both the instrument and the association are too weak to draw conclusions about the relationship between school size and LSP achievement impacts. We do not find any systematic evidence of a relationship between the proportion of African American or LSP students enrolled in a school¹² and LSP impacts on student achievement.

Other differences in LSP impacts by school demographics relied on a weak instrument or were not statistically significant ($p > 0.10$) (see Appendix Tables A1-A8 for diagnostic measures of all instruments). For full results by total enrollment and enrollment demographics, please see Appendix Tables A11, A15, A19, and A23.

¹² We calculated the LSP share of school enrollments by dividing the number of LSP lottery winners enrolled in a given school by the total K-12 enrollment of the school as reported by the Private School Universe Survey.

Table 8.

Variation in LATE estimates by school tuition, number of full-time equivalents, and student/teacher ratio

		2012-13		2013-14		2014-15		2015-16	
		ELA	Math	ELA	Math	ELA	Math	ELA	Math
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Tuition, per \$1,000, relative to mean</i>	<i>n</i>	1,418	1,417	1,265	1,264	985	984	596	587
	τ_2	0.04 (0.06)	0.06 (0.08)	0.04 (0.11)	0.21 (0.19)	0.01 (0.32)	0.06 (0.26)	0.09 (11.49)	-0.01 (2.23)
	μ	\$5,394		\$5,454		\$5,470		\$5,510	
<i>Full-Time Equivalents, per 10 FTE, relative to mean</i>	<i>n</i>	1,504	1,503	1,351	1,350	1,049	1,048	632	623
	τ_2	0.00 (0.03)	0.11*** (0.04)	0.01 (0.05)	0.11* (0.06)	0.13 (0.10)	0.29* (0.15)	-0.07 (0.31)	0.15 (0.80)
	μ	21.0		21.1		21.3		21.5	
<i>Student/Teacher Ratio, relative to mean</i>	<i>n</i>	1,504	1,503	1,351	1,350	1,049	1,048	632	623
	τ_2	0.01 (0.02)	0.00 (0.03)	0.05 (0.04)	0.06 (0.06)	0.06 (0.08)	0.00 (0.11)	0.02 (0.22)	-0.09 (0.21)
	μ	12.95		12.80		12.86		12.96	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. "Religious" schools include all Catholic schools, Christian but non-Catholic schools, Muslim schools, and Jewish schools. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Tables A1, A3, A5, and A7). For full results, see Appendix Tables A10, A14, A18, and A22.

Table 9.

Variation in LATE estimates by total enrollment and enrollment demographics

		2012-13		2013-14		2014-15		2015-16	
		ELA	Math	ELA	Math	ELA	Math	ELA	Math
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total K-12 Enrollment, <i>per 100 students, relative to mean</i>	<i>n</i>	1,504	1,503	1,351	1,350	1,049	1,048	632	623
	τ_2	0.01 (0.02)	0.10*** (0.03)	0.03 (0.04)	0.11** (0.05)	0.10 (0.07)	0.21* (0.11)	-0.03 (0.23)	0.07 (1.32)
	μ	275.29		275.94		278.64		280.42	
% African American, <i>per 10pp, relative to mean</i>	<i>n</i>	1,504	1,503	1,351	1,350	1,049	1,048	632	623
	τ_2	-0.03 (0.02)	0.02 (0.03)	-0.10** (0.04)	-0.05 (0.05)	-0.10 (0.06)	-0.06 (0.08)	0.01 (0.16)	-0.08 (0.14)
	μ	44.60		44.29		43.65		42.95	
% LSP, <i>per 10 pp, relative to mean</i>	<i>n</i>	1,504	1,503	1,318	1,317	930	929	523	517
	τ_2	0.00 (0.03)	0.00 (0.04)	-0.17 (0.09)	-0.15 (0.12)	0.02 (0.23)	-0.33 (0.32)	0.38 (4.60)	-0.20 (4.41)
	μ	0.25		0.21		0.17		0.13	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. "Religious" schools include all Catholic schools, Christian but non-Catholic schools, Muslim schools, and Jewish schools. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Tables A1, A3, A5, and A7). For full results, see Appendix Tables A11, A15, A19, and A23.

Outcomes by instructional intensity

Next, we consider measures of instructional intensity including the number of days per school year, the number of hours per school day, and the total number of instructional hours in a school year, as reported in the PSS (Table 10). The LSP impacts on math outcomes are positively associated with the number of hours per school day or total instructional hours per school year¹³ of the preferred school. We find some evidence that LATE estimates varied with instructional hours, but not with the number of instructional days. Students who preferred schools with an additional hour per school day above the mean of 7.1 hours experienced LSP impacts over a fifth of a standard deviation higher in math than their peers who preferred a school at the mean (0.22, $p < 0.05$). Similarly, students who preferred schools with an additional 100 instructional hours per year above the mean of 1,268 hours experienced LSP impacts over a tenth of a standard deviation higher in math than their peers who preferred schools at the mean (0.10, $p < 0.05$). This pattern continued into the second year, when students who preferred schools with an additional hour per school day above the mean experienced LSP impacts 0.38 standard deviations higher in math and students who preferred schools with an additional 100 hours per year above the mean experienced LSP impacts 0.19 standard deviations higher in math, each relative to the LSP impacts on their peers who preferred a school at the mean. However, by the third year, all interaction estimates relied on weak instruments, suggesting possibly biased results. For full findings by instructional intensity, please reference Appendix Tables A12, A16, A20, and A24.

¹³ The total number of instructional hours is a simple interaction of days per year and hours per day, without adjusting for the number of partial days a school may have scheduled.

Table 10.

Variation in LATE estimates by instructional intensity

		2012-13		2013-14		2014-15		2015-16	
		ELA	Math	ELA	Math	ELA	Math	ELA	Math
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>School Year Days, per 10 days, relative to mean</i>	<i>n</i>	1,504	1,503	1,351	1,350	1,049	1,048	632	623
	τ_2	-0.02 (0.09)	0.03 (0.12)	0.12 (0.71)	0.31 (1.26)	0.17 (1.85)	0.35 (1.76)	0.34 (22.51)	0.13 (56.02)
	μ	178.8		179.1		179.0		179.1	
<i>School Day Hours, relative to mean</i>	<i>n</i>	1,504	1,503	1,351	1,350	1,049	1,048	632	623
	τ_2	0.03 (0.11)	0.22** (0.10)	0.25 (0.17)	0.38* (0.21)	0.32 (0.35)	0.32 (0.32)	0.69 (0.74)	0.15 (0.62)
	μ	7.09		7.08		7.09		7.12	
<i>Total Instructional Hours, per 100 hours, relative to mean</i>	<i>n</i>	1,504	1,503	1,351	1,350	1,049	1,048	632	623
	τ_2	0.01 (0.05)	0.10** (0.05)	0.11 (0.08)	0.19* (0.10)	0.13 (0.16)	0.16 (0.14)	0.24 (0.48)	0.06 (0.26)
	μ	1,268.0		1,267.8		1,270.4		1,275.5	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. "Religious" schools include all Catholic schools, Christian but non-Catholic schools, Muslim schools, and Jewish schools. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Tables A1, A3, A5, and A7). For full results, see Appendix Tables A12, A16, A20, and A24.

Results by tercile

Finally, we conduct a tiered analysis by dividing schools enrolling LSP students into three equally sized terciles according to the distribution of each school characteristic.¹⁴ Using our simple model, we restrict our analytic sample to only students who listed as their first preference a school in each respective tercile, comparing the outcomes of these students to the control group of students who also listed these schools as their first preference, but did not win an LSP scholarship. However, as in the previous section, we cannot claim that differences in LSP impacts necessarily were caused by the school characteristics captured by the terciles, as students were not randomly assigned within school characteristic terciles, but were randomized by first-choice school.

Evaluations of the LSP after one year report large negative effects in both ELA and math (Abdulkadiroglu et al., 2018; Mills, 2015) that dissipate for ELA after two years (Mills & Wolf, 2017a), become statistically insignificant for both subject domains after three years (Mills & Wolf, 2017b), and return to negative and statistically significant for both subject domains after four years (Mills & Wolf, 2019).

Variation in ELA and math LATE estimates by terciles of school tuition and staffing is summarized in Tables 11 and 12, respectively. Students who preferred schools that charged a higher tuition do not experience a negative effect of the LSP on math outcomes. In the first year of statewide expansion of the LSP, students who preferred a school in the bottom two terciles in terms of tuition charged experience math impacts of -0.47 and -0.67 standard deviations, respectively ($p < 0.01$). This negative math effect persists for students preferring a school in the bottom tercile of tuition in the second year (-0.69 standard deviations, $p < 0.05$). Similarly,

¹⁴ For the cutoff values for the bottom, middle, and upper terciles, refer to Tables 3-6.

Table 11.

Variation in ELA LATE estimates by terciles of school tuition, number of full-time equivalents, and student/teacher ratio

		2012-13		2013-14		2014-15		2015-16	
		<i>n</i>	β	<i>n</i>	β	<i>n</i>	β	<i>n</i>	β
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tuition	Top 33%	259	-0.01 (0.36)	226	0.54 (0.71)	202	-0.01 (6.07)	130	-0.55 (3.38)
	Middle 33%	564	-0.15 (0.14)	457	0.03 (0.22)	311	0.54 (0.40)	179	1.13 (4.66)
	Bottom 33%	599	-0.16 (0.14)	584	-0.30 (0.22)	475	-0.09 (0.39)	290	-0.32 (0.86)
Full-Time Equivalents	Top 33%	510	-0.03 (0.13)	445	0.16 (0.24)	309	0.41 (0.45)	168	0.04 (1.47)
	Middle 33%	640	-0.46*** (0.18)	564	-0.54** (0.22)	441	-0.34 (0.53)	277	-0.78 (11.16)
	Bottom 33%	375	0.13 (0.25)	361	0.13 (0.43)	319	0.23 (0.74)	207	0.46 (1.49)
Student/Teacher Ratio	Top 33%	628	0.00 (0.13)	580	0.29 (0.28)	488	0.44 (0.45)	297	0.01 (0.82)
	Middle 33%	532	-0.40** (0.20)	488	-0.61** (0.26)	360	-0.66 (0.90)	216	-0.26 (14.25)
	Bottom 33%	365	-0.18 (0.15)	302	-0.21 (0.29)	221	0.26 (0.44)	139	-0.33 (1.71)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within riskset. “ β ” represents the Local Average Treatment Effect, the estimated achievement effect for voucher usage, for this specified sample. Estimates that use instruments that do not satisfy Staiger and Stock’s (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Tables A2, A4, A6, and A8).

Table 12.

Variation in math LATE estimates by terciles of school tuition, number of full-time equivalents, and student/teacher ratio

		2012-13		2013-14		2014-15		2015-16	
		<i>n</i>	β	<i>n</i>	β	<i>n</i>	β	<i>n</i>	β
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tuition	Top 33%	259	-0.25 (0.34)	226	1.05 (0.81)	202	-0.59 (5.03)	130	-1.04 (2.37)
	Middle 33%	563	-0.47*** (0.14)	457	0.00 (0.17)	311	0.59 (0.39)	175	0.10 (4.07)
	Bottom 33%	599	-0.67*** (0.19)	583	-0.69** (0.30)	474	-0.38 (0.44)	285	-0.90 (0.89)
Full-Time Equivalents	Top 33%	509	-0.07 (0.11)	445	0.18 (0.19)	309	0.61 (0.39)	165	-0.82 (1.40)
	Middle 33%	641	-0.98*** (0.21)	563	-0.92*** (0.31)	441	-0.75 (0.46)	273	-1.91 (25.77)
	Bottom 33%	374	-0.41* (0.24)	361	0.15 (0.43)	318	-0.02 (0.80)	205	0.19 (1.15)
Student/Teacher Ratio	Top 33%	627	-0.42*** (0.13)	581	0.33 (0.21)	488	0.23 (0.37)	294	-0.66 (0.53)
	Middle 33%	533	-0.76 (0.24)	487	-1.12*** (0.33)	360	-0.93 (0.68)	211	-3.29 (115.91)
	Bottom 33%	364	-0.42** (0.20)	301	-0.14 (0.30)	220	0.31 (0.63)	138	0.09 (1.24)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within riskset. “ β ” represents the Local Average Treatment Effect, the estimated achievement effect for voucher usage, for this specified sample. Estimates that use instruments that do not satisfy Staiger and Stock’s (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Tables A2, A4, A6, and A8).

students who preferred a school in the bottom two terciles of the number of full-time teachers or equivalents employed by a school experience negative math impacts (-0.98 standard deviations, higher tuition do not experience a negative effect of the LSP on math outcomes. In the first year of statewide expansion of the LSP, students who preferred a school in the bottom two terciles in terms of tuition charged experience math impacts of -0.47 and -0.67 standard deviations, respectively ($p < 0.01$). This negative math effect persists for students preferring a school in the bottom tercile of tuition in the second year (-0.69 standard deviations, $p < 0.05$). Similarly, students who preferred a school in the bottom two terciles of the number of full-time teachers or equivalents employed by a school experience negative math impacts (-0.98 standard deviations, $p < 0.01$ and -0.41, $p < 0.10$, respectively). In contrast, students who preferred schools in the top tercile of tuition or full-time equivalents employed by the school do not experience a negative effect on math outcomes. The pattern is less clear in terms of a school's student/teacher ratio, which is associated with a negative math outcome in the top and bottom terciles in the first year, but only the middle tercile in the second year. There is no clear pattern for ELA outcomes across any of these characteristics.

Regarding total enrollment and enrollment demographic characteristics (Tables 13-14), students who preferred a school in the bottom tercile of total enrollment experience a negative LSP impact of almost three-quarters of a standard deviation in math achievement in the first year (-0.74, $p < 0.01$). Students who preferred a school in the middle tercile of total enrollment experience a negative LSP math impact of over half a standard deviation in the second year (-0.51, $p < 0.10$). Students who preferred a school in the top tercile of total enrollment do not experience a statistically significant effect on math outcomes in the first, second, and fourth years, and are positively impacted in the third year (0.68 standard deviations, $p < 0.10$).

Table 13.

Variation in ELA LATE estimates by terciles of total school enrollment and enrollment demographics

		2012-13		2013-14		2014-15		2015-16	
		<i>n</i>	β	<i>n</i>	β	<i>n</i>	β	<i>n</i>	β
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total K-12 Enrollment	Top 33%	503	-0.02 (0.12)	439	0.19 (0.27)	304	0.44 (0.47)	165	0.04 (1.44)
	Middle 33%	717	-0.28 (0.17)	652	-0.29 (0.23)	519	0.12 (0.47)	274	-0.47 (8.85)
	Bottom 33%	305	-0.17 (0.16)	279	-0.32 (0.42)	246	-0.67 (0.70)	213	-0.06 (0.72)
% African American	Top 33%	867	-0.18 (0.12)	813	-0.21 (0.24)	671	-0.30 (0.50)	403	0.08 (7.42)
	Middle 33%	608	-0.15 (0.16)	508	-0.05 (0.22)	363	0.48 (0.35)	226	-0.07 (2.70)
	Bottom 33%	50	-0.14 (0.52)	49	0.59 (0.70)	35	0.34 (0.73)	23	-1.54 (1.80)
% LSP	Top 33%	1,138	-0.18 (0.11)	1,012	-0.18 (0.19)	714	0.21 (0.30)	412	0.48 (0.95)
	Middle 33%	354	-0.19 (0.18)	302	0.23 (0.31)	221	-0.03 (0.44)	120	-0.69 (1.40)
	Bottom 33%	33	0.32 (0.62)	23	0.04 (0.60)	14	0.14 (1.38)	10	-2.52 (2.01)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within riskset. “ β ” represents the Local Average Treatment Effect, the estimated achievement effect for voucher usage, for this specified sample. Estimates that use instruments that do not satisfy Staiger and Stock’s (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Tables A2, A4, A6, and A8).

Table 14.

Variation in math LATE estimates by terciles of total school enrollment and enrollment demographics

		2012-13		2013-14		2014-15		2015-16	
		<i>n</i>	β	<i>n</i>	β	<i>n</i>	β	<i>n</i>	β
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total K-12 Enrollment	Top 33%	502	-0.04 (0.12)	439	0.22 (0.20)	304	0.68* (0.38)	162	-0.62 (1.05)
	Middle 33%	717	-0.79 (0.19)	652	-0.51* (0.30)	519	-0.35 (0.41)	272	-2.34 (9.79)
	Bottom 33%	305	-0.74*** (0.24)	278	-0.60 (0.47)	245	-0.86 (0.93)	209	-0.24 (0.60)
% African American	Top 33%	866	-0.52*** (0.14)	813	-0.22 (0.23)	670	-0.18 (0.47)	396	-1.04 (2.29)
	Middle 33%	608	-0.52*** (0.19)	507	-0.38 (0.30)	363	0.08 (0.45)	225	-0.58 (18.48)
	Bottom 33%	50	-0.56 (0.44)	49	0.55 (0.66)	35	-0.46 (0.68)	22	-1.65 (2.07)
% LSP	Top 33%	1,138	-0.54*** (0.13)	1,012	-0.31 (0.20)	714	-0.06 (0.28)	406	-0.76 (0.67)
	Middle 33%	353	-0.51** (0.26)	301	0.05 (0.31)	220	0.13 (0.62)	120	-0.69 (2.87)
	Bottom 33%	33	-0.10 (0.30)	23	0.52 (0.83)	14	-0.57 (0.58)	10	-1.00 (0.88)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within riskset. “ β ” represents the Local Average Treatment Effect, the estimated achievement effect for voucher usage, for this specified sample. Estimates that use instruments that do not satisfy Staiger and Stock’s (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Tables A2, A4, A6, and A8).

In terms of demographic composition, students who preferred schools with the greatest proportion of African American or LSP students experience negative effects of the program on math outcomes after the first year. Students who preferred a school in the top two terciles by proportion of African American or LSP students experience a negative math impact of more than half a standard deviation. In contrast, students who preferred a school in the bottom tercile of these characteristics do not experience an effect. All effect estimates for math by enrollment demographics become statistically insignificant in the second year and continue to be so after four years (Table 14). No statistically significant associations of enrollment demographics are detected on ELA outcomes (Table 13).

Finally, no clear pattern emerges in terms of a school's instructional intensity and the achievement impacts of the LSP (Tables 15-16). In the first year, students who preferred a school in the top tercile of the length of school year experience a negative effect on ELA outcomes of -0.66 standard deviations ($p < 0.10$). Students who preferred a school in the middle tercile of total instructional hours experience a negative effect on ELA outcomes of almost half a standard deviation (-0.49, $p < 0.05$) and students who preferred a school in the middle tercile of days per school year experience a negative math effect of half a standard deviation ($p < 0.01$). Despite failing to detect any statistically significant associations on ELA outcomes in terms of hours per school day, students who preferred a school in the bottom tercile of the length of school day experience a negative effect on ELA outcomes of 0.57 standard deviations ($p < 0.10$). Students in all terciles of instructional intensity experience negative effects on math outcomes after the first year, but all other effect estimates became statistically insignificant by the second year.

Table 15.

Variation in ELA LATE estimates by terciles of instructional intensity

		2012-13		2013-14		2014-15		2015-16	
		<i>n</i>	β	<i>n</i>	β	<i>n</i>	β	<i>n</i>	β
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
School Year Days	Top 33%	227	-0.66*	172	-0.55	98	-0.26	62	-0.62
			(0.35)		(0.39)		(0.78)		(5.13)
	Middle 33%	1,167	-0.14	1,086	-0.13	883	0.04	548	-0.12
			(0.11)		(0.17)		(0.33)		(2.31)
	Bottom 33%	131	0.22	112	0.57	88	0.51	42	0.74
			(0.28)		(0.52)		(0.90)		(5.21)
School Day Hours	Top 33%	752	-0.10	670	0.03	510	0.21	267	0.80
			(0.14)		(0.24)		(0.38)		(0.96)
	Middle 33%	460	-0.22	395	-0.01	300	0.06	202	-1.00
			(0.17)		(0.29)		(0.57)		(3.05)
	Bottom 33%	313	-0.25	305	-0.57*	259	-0.41	183	-1.70
			(0.24)		(0.34)		(2.11)		(117.32)
Total Instructional Hours	Top 33%	704	-0.05	631	0.09	474	0.37	249	1.04
			(0.14)		(0.24)		(0.33)		(0.97)
	Middle 33%	332	-0.49**	291	-0.27	208	-0.70	146	-1.70
			(0.21)		(0.31)		(0.90)		(1.12)
	Bottom 33%	489	-0.13	448	-0.35	387	0.00	257	-1.14
			(0.18)		(0.31)		(0.83)		(78.38)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within riskset. “ β ” represents the Local Average Treatment Effect, the estimated achievement effect for voucher usage, for this specified sample. Estimates that use instruments that do not satisfy Staiger and Stock’s (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Tables A2, A4, A6, and A8).

Table 16.

Variation in math LATE estimates by terciles of instructional intensity

		2012-13		2013-14		2014-15		2015-16	
		<i>n</i>	β	<i>n</i>	β	<i>n</i>	β	<i>n</i>	β
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
School Year Days	Top 33%	227	-0.73 (0.50)	172	-0.29 (0.55)	98	0.80 (1.15)	62	-0.21 (5.52)
	Middle 33%	1,165	-0.50*** (0.11)	1,085	-0.31 (0.20)	882	-0.20 (0.33)	541	-1.20 (1.30)
	Bottom 33%	132	-0.48 (0.29)	112	0.36 (0.60)	88	-0.37 (0.76)	40	0.29 (3.91)
School Day Hours	Top 33%	752	-0.38*** (0.13)	669	-0.12 (0.26)	510	0.09 (0.31)	260	-1.16 (0.86)
	Middle 33%	460	-0.73*** (0.24)	395	-0.15 (0.30)	300	-0.35 (0.69)	202	-0.24 (15.61)
	Bottom 33%	312	-0.58*** (0.22)	305	-0.65 (0.41)	258	-0.26 (6.45)	181	-1.29 (210.11)
Total Instructional Hours	Top 33%	703	-0.33** (0.14)	630	-0.05 (0.26)	474	0.22 (0.28)	244	-0.98 (0.67)
	Middle 33%	332	-0.92*** (0.34)	291	-0.37 (0.35)	208	-0.57 (1.06)	146	-0.79 (1.56)
	Bottom 33%	489	-0.54*** (0.17)	448	-0.50 (0.35)	386	-0.38 (0.76)	253	-0.79 (73.62)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within riskset. “ β ” represents the Local Average Treatment Effect, the estimated achievement effect for voucher usage, for this specified sample. Estimates that use instruments that do not satisfy Staiger and Stock’s (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Tables A2, A4, A6, and A8).

Summary of differential effects

The clearest results of our exploration of heterogeneous effects emerged from the tercile analysis of school tuition, total K-12 enrollment, and school day hours. Many of the negative and statistically significant math test score effects are concentrated in the bottom tercile schools regarding those three factors (Figure 1). LSP students whose first-choice school was in the top tercile of total K-12 enrollment actually experienced a positive, statistically significant math impact in 2014-15. LSP students whose first-choice schools were in the upper third of the distribution on school tuition experienced a large positive math impact in 2013-14 but it was not statistically significant.

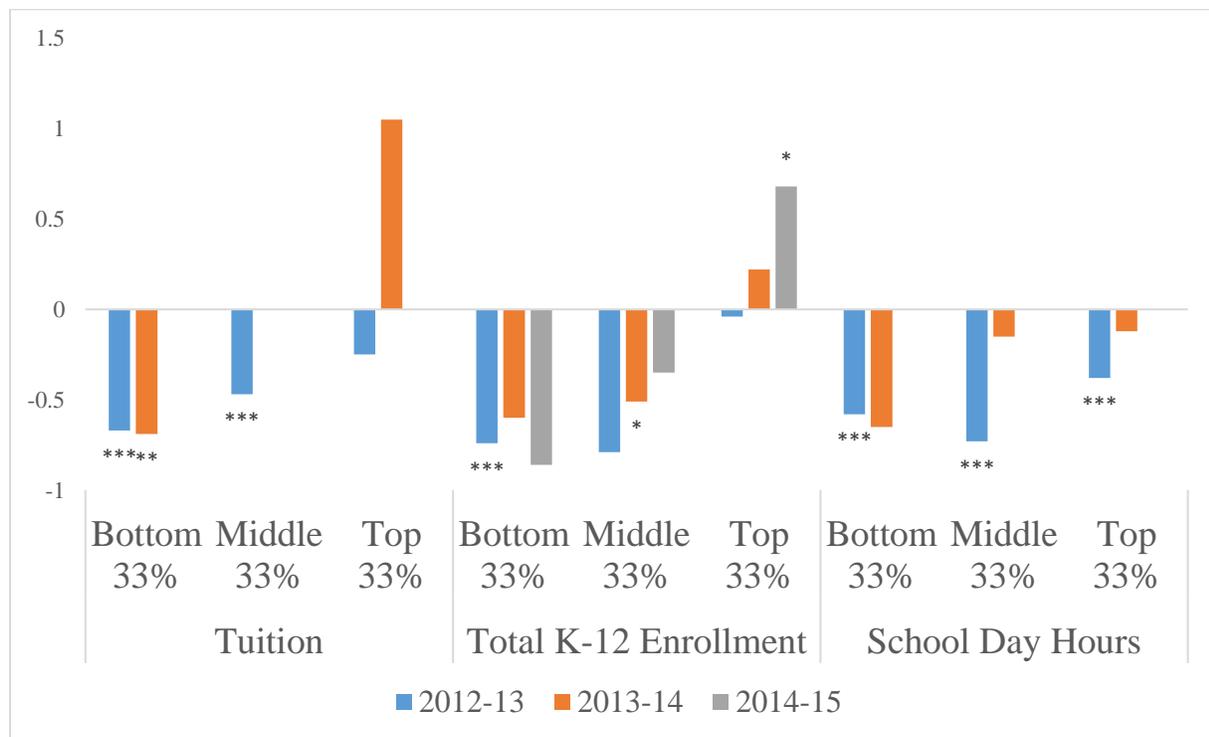


Figure 1. Variation in LSP math impacts across terciles of school tuition charged, total school enrollment, and school day hours, 2012-13 through 2014-15.

Notes. Figure presents point estimates from regression models for 2012-13 through 2014-15 for standardized math impacts of ever using an LSP voucher placement. Estimates in 2014-15 for tuition and school day hours, as well as 2015-16 generally, failed a reliability test and therefore are not presented. *** p<.01, ** p<.05, * p<.10.

Robustness checks

The analysis sample in the preceding section is restricted to students with baseline achievement to allow us to control for confounding influences of baseline achievement in our models. As a robustness check, we replicate our analysis without restricting our sample to those with baseline test scores. We find our results to be highly robust across both specifications, with only a few exceptions.

One advantage of using an unrestricted sample is that the increased sample size should increase statistical power and the relevance of our instruments, as captured by larger first stage F -statistics. Most of the effect estimates that relied on weak instruments in our baseline analysis remained statistically insignificant even with strong instruments in our unrestricted analysis, with only two exceptions. In our third year of analysis, the effect estimation for the association between FTEs and both ELA and math outcomes relied on a strong instrument. The association is not statistically significant for ELA, but preferring a school with an additional 10 FTEs above the mean is associated with experiencing an LSP impact that is 0.20 standard deviations higher in math in the third year ($p < 0.10$).

In our main analysis, if a statistically significant effect or association was detected in our baseline achievement sample, an effect estimate was consistently calculated in the same direction in the unrestricted sample. Overall, eight statistically significant interactions were detected in our baseline sample and six statistically significant interactions were detected in our unrestricted sample, with six statistically significant interactions overlapping between the two analyses. Five statistically significant interactions were detected in the original analysis on the baseline test score sample that were not detected in the analysis on the unrestricted sample. Those findings were for the associations of LSP share on ELA in 2013-14, hours per school day on math in

2012-13 and 2013-14, and total instructional hours on math in 2012-13 and 2013-14. Two statistically significant interactions were detected in the analysis on the unrestricted sample that were not detected in the analysis of the baseline achievement sample. Those findings were for the associations of student/teacher ratio on math in 2013-14 and proportion of students who identify as African American on ELA in 2014-15. The average difference between corresponding effect estimates across the two samples was relatively small, at 0.12 standard deviations, with a standard deviation in the differences in estimates of 0.39.

In our tiered analysis, 30 statistically significant effects were detected in the original analysis on the baseline test score sample and 84 statistically significant effects were detected in the analysis on the unrestricted sample. These estimates consistently pointed in the same direction. For the analysis on the unrestricted sample, in the second year, students who preferred a school in the middle tercile of LSP share experience a decline of 0.40 standard deviations on math outcomes relative to their non-LSP peers ($p < 0.10$). For the same unrestricted sample, in the third year of analysis, preferring a school in the middle tercile of proportion of students who identify as African American (-0.68 standard deviations, $p < 0.01$) or in the middle tercile of LSP share (-0.56, $p < 0.10$) leads to negative LSP math impacts. Also for the unrestricted sample, in the fourth year of analysis, LSP students who preferred a school in the middle tercile of tuition experience math outcomes that are 0.57 standard deviations lower relative to their control group ($p < 0.05$). In the analysis on the baseline sample, in contrast, all four of these corresponding LSP effect estimates were positive, though statistically insignificant. The average difference between corresponding effect estimates across the two samples for the tiered analysis is 0.02 standard deviations and the standard deviation of the differences is 0.67.

Overall, the pattern of findings is robust across our two samples. We generally do not find consistent evidence that effects varied across school characteristics, with the possible exceptions of school size, tuition, and length of school day. These findings are summarized in Tables 17-20. We caution the reader that these tables are a selected list of statistically significant findings that are not necessarily representative of our overall effect estimates, most of which were not statistically significant at conventional levels. For the sake of brevity, full results are not included in this report but will be made available online and by request.

Conclusion

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

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, notable exceptions. Our study shows that outcomes are positively associated with size, tuition, and length of school day. Moreover, it is possible that the generally insignificant differences regarding the associations of other school features with the LSP impacts partially reflect low statistical power, as several differences in impact estimates statistically insignificant are nonetheless large in magnitude. This condition could be due to

limited variation in certain characteristics (see Tables 3-6), small sample size, particularly for the tiered analysis (see Tables 11-16), or the limited predictive power and relevance of instruments, especially by the third and fourth years (see Appendix Tables A1-A8).

Table 17.

Statistically significant findings in both samples, in baseline achievement sample, and in unrestricted sample, 2012-13

Statistically Significant Finding	In Both	Only in Baseline	Only in Unrestricted
	(1)	(2)	(3)
FTE positively associated with math	X		
Total K-12 enrollment positively associated with math	X		
Negative ELA outcomes for middle FTE tercile	X		
Negative ELA outcomes for middle student/teacher ratio tercile	X		
Negative math outcomes for middle tuition tercile	X		
Negative math outcomes for bottom tuition tercile	X		
Negative math outcomes for middle FTE tercile	X		
Negative math outcomes for bottom FTE tercile	X		
Negative math outcomes for top student/teacher ratio tercile	X		
Negative math outcomes for bottom student/teacher ratio tercile	X		
Negative math outcomes for bottom total K-12 enrollment tercile	X		
Negative math outcomes for top % African American tercile	X		
Negative math outcomes for middle % African American tercile	X		
Negative math outcomes for top % LSP tercile	X		
Negative math outcomes for middle % LSP tercile	X		
Negative ELA outcomes for top school year days tercile	X		
Negative math outcomes for middle school year days tercile	X		
Negative math outcomes for top school day hours tercile	X		
Negative math outcomes for middle school day hours tercile	X		
Negative math outcomes for bottom school day hours tercile	X		
Negative math outcomes for top total instructional hours tercile	X		
Negative math outcomes for middle total instructional hours tercile	X		
Negative math outcomes for bottom total instructional hours tercile	X		
School day hours positively associated with math		X	
Total instructional hours positively associated with math		X	
Negative ELA outcomes for middle total instructional hours tercile		X	
Negative ELA outcomes for bottom tuition tercile			X
Negative ELA outcomes for bottom student/teacher ratio tercile			X
Negative math outcomes for middle student/teacher ratio tercile			X
Negative ELA outcomes for middle total K-12 enrollment tercile			X
Negative ELA outcomes for bottom total K-12 enrollment tercile			X
Negative math outcomes for middle total K-12 enrollment tercile			X
Negative ELA outcomes for top % African American tercile			X
Negative ELA outcomes for top % LSP tercile			X
Negative ELA outcomes for middle school year days tercile			X
Negative math outcomes for top school year days tercile			X
Negative math outcomes for bottom school year days tercile			X
Negative ELA outcomes for top school day hours tercile			X
Negative ELA outcomes for middle total instructional hours tercile			X

Table 18.

Statistically significant findings in both samples, in baseline achievement sample, and in unrestricted sample, 2013-14

Statistically Significant Finding	In Both	Only in Baseline	Only in Unrestricted
	(1)	(2)	(3)
Urban schools positive relative to non-urban schools on math	X		
FTE positively associated with math	X		
Total K-12 enrollment positively associated with math	X		
Negative math outcomes for bottom tuition tercile	X		
Negative math outcomes for middle FTE tercile	X		
Negative ELA outcomes for middle student/teacher ratio tercile	X		
Negative math outcomes for middle student/teacher ratio tercile	X		
Negative math outcomes for middle total K-12 enrollment tercile	X		
Negative ELA outcomes for middle FTE tercile		X	
Negative ELA outcomes for bottom school day hours tercile		X	
Negative ELA outcomes for bottom tuition tercile			X
Negative ELA outcomes for middle FTE tercile			X
Negative ELA outcomes for middle total K-12 enrollment tercile			X
Negative math outcomes for bottom total K-12 enrollment tercile			X
Negative ELA outcomes for middle % African American tercile			X
Negative math outcomes for top % African American tercile			X
Negative math outcomes for middle % African American tercile			X
Negative math outcomes for top % LSP tercile			X
Negative math outcomes for middle % LSP tercile			X
Negative math outcomes for middle school year days tercile			X
Negative math outcomes for top school day hours tercile			X
Negative math outcomes for middle school day hours tercile			X
Negative math outcomes for bottom school day hours tercile			X
Negative ELA outcomes for middle total instructional hours tercile			X
Negative math outcomes for middle total instructional hours tercile			X
Negative math outcomes for bottom total instructional hours tercile			X

Table 19.

Statistically significant findings in both samples, in baseline achievement sample, and in unrestricted sample, 2014-15

Statistically Significant Finding	In Both	Only in Baseline	Only in Unrestricted
	(1)	(2)	(3)
FTE positively associated with math	X		
Positive math outcomes for top total K-12 enrollment tercile		X	
Negative math outcomes for bottom tuition tercile			X
Negative ELA outcomes for middle FTE tercile			X
Negative math outcomes for middle FTE tercile			X
Negative ELA outcomes for middle student/teacher ratio tercile			X
Negative math outcomes for middle student/teacher ratio tercile			X
Negative math outcomes for middle total K-12 enrollment tercile			X
Negative math outcomes for middle % African American tercile			X
Negative math outcomes for middle % LSP tercile			X
Negative math outcomes for middle school year days tercile			X
Negative math outcomes for middle school day hours tercile			X
Negative math outcomes for bottom total instructional hours tercile			X

Table 20.

Statistically significant findings in both samples, in baseline achievement sample, and in unrestricted sample, 2015-16

Statistically Significant Finding	In Both	Only in Baseline	Only in Unrestricted
	(1)	(2)	(3)
Negative math outcomes for middle tuition tercile			X
Negative math outcomes for bottom tuition tercile			X
Negative ELA outcomes for middle FTE tercile			X
Negative math outcomes for top FTE tercile			X
Negative math outcomes for middle FTE tercile			X
Negative math outcomes for middle student/teacher ratio tercile			X
Negative ELA outcomes for middle total K-12 enrollment tercile			X
Negative math outcomes for top total K-12 enrollment tercile			X
Negative math outcomes for middle total K-12 enrollment tercile			X
Negative math outcomes for middle % African American tercile			X
Negative ELA outcomes for middle % LSP tercile			X
Negative math outcomes for top % LSP tercile			X
Negative math outcomes for middle % LSP tercile			X
Negative math outcomes for middle school year days tercile			X
Negative math outcomes for top school day hours tercile			X
Negative ELA outcomes for middle total instructional hours tercile			X
Negative math outcomes for top total instructional hours tercile			X
Negative math outcomes for middle total instructional hours tercile			X

Our finding that LSP-participating schools with higher enrollments generated significantly better (or less negative) initial voucher impacts than LSP-participating schools with lower enrollments is consequential. That finding implies that a movement 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. Given the limitations of our research design, that connection is a correlation and not an experimental impact, but it is consistent with market theory. Our finding that higher school tuitions are associated with better LSP impacts suggests that tuition levels may be serving as proxies for private school quality in the LSP. Finally, our finding that LSP schools with longer school days were correlated with better LSP achievement impacts suggest that private schools in the LSP that adopt the conditions of “academic press” (Berends, Goldring, Stein, & Cravens, 2010) tend to be better able to reproduce the average gains achieved by public school students on the state accountability test. As it stands, our analysis indicates LSP treatment effects on student achievement do not generally differentiate across school characteristics, with these few exceptions.

Though our study builds upon a highly rigorous experimental design, our study suffers from three important limitations. First, weak instruments often limit our analysis. There is no scholarly consensus on calculating effective F -statistics for analyses with multiple endogenous regressors. That problem afflicts our study, as students sort themselves into first-preference lotteries, affecting both LSP enrollment and the school characteristic interactions. Therefore, we depend on an established, but imperfect, rule of thumb to determine which instruments do and do not minimize bias in our estimates. Second, limited variation in school characteristics in our sample of schools may limit our ability to detect or generalize any effects or associations we

find. For example, while we find that LSP test score impacts are positively associated with the number of hours in the school day, the schools in our sample do not vary widely in this regard.

Finally, while our study does have strong causal inference within lotteries, we cannot make causal comparisons across lotteries. For students who experienced a first-choice lottery for a school with a particular characteristic (e.g. a school in the top tercile of tuition), LSP scholarships were awarded independently of any observed or unobserved characteristics. Thus, the estimates of the impact of the LSP for the subgroup of students who preferred a school in the top tercile of tuition support strong causal inferences. However, students may sort themselves into different lotteries (e.g. preferring a school in the middle or bottom tercile of tuition) for reasons observed or unobserved that may also be related to their outcomes. Thus, comparisons across subgroups of students who preferred schools with different characteristics – the main focus of this mediator analysis – do not support strong causal inferences and are merely descriptive of associations in the first four years of the LSP.

Sources Cited

- 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. JSTOR, www.jstor.org/stable/4132848.
- 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. DOI: 10.1257/app.20160634.
- Andrews, I., Stock, J., & Sun, L. (2018). *Weak instruments in IV regression: Theory and practice*. Working paper. Accessed at: <https://scholar.harvard.edu/iandrews/publications/weak-instruments-iv-regression-theory-and-practice>.
- Angrist, J. D. & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.
- 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), 299-323. DOI 10.1198/016214503000071.
- Berends, M., Goldring, E., Stein, M., & Cravens, X. (2010). Instructional conditions in charter schools and students' mathematics achievement gains. *American Journal of Education*, 116(3), 303-335. DOI: [10.1086/651411](https://doi.org/10.1086/651411).
- 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. DOI: 10.1080/19345747.2014.921259.
- Bloom, H. S. & Unterman, R. (2014). Can small schools of choice improve educational prospects for disadvantaged students? *Journal of Policy Analysis and Management*, 33(2): 290-319. DOI: 10.1002/pam.21748.
- Chubb, J. E. & Moe, T. M. (1990). *Politics, markets, & America's schools*. Washington, DC: Brookings Institution Press.
- 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(2), 301-315. DOI: [10.1111/j.1541-0072.2008.00268.x](https://doi.org/10.1111/j.1541-0072.2008.00268.x).

- DeAngelis, C. A., Burke, L. M., & Wolf, P. J. (2018). *The effects of regulations on private school choice program participation: Experimental evidence from Florida*. EDRE working paper 2018-08. Social Science Research Network. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3273428#.
- DeAngelis, C. A., Burke, L. M., & Wolf, P. J. (2019). *The effects of regulations on private school choice program participation: Experimental evidence from California and New York*. EDRE working paper 2019-07. Social Science Research Network. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3349453.
- 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.
- EdChoice (2019). *The ABCs of school choice: A comprehensive guide to every private school choice program in America*. Indianapolis, IN: EdChoice.
- 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. DOI: 10.1080/0161956X.2016.1207436.
- 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%20online%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/>
- Fuhrer, R. (2013). Know the facts on parental school choice. *NEA-Alaska President's Blog*, available at www.neaalaska.org
- 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. Available at: http://media.hoover.org/sites/default/files/documents/ednext20012_46b.pdf.
- Greene, J. P., Peterson, P. E., & Du, J. (1999). Effectiveness of school choice: The Milwaukee experiment. *Education and Urban Society*, 31(2), 191–213. DOI: [10.1177/0013124599031002005](https://doi.org/10.1177/0013124599031002005).

- Gutmann, A. (2003). Assessing arguments for school choice: Pluralism, parental rights, or educational results? In A. Wolfe (Ed.), *School choice: The moral debate* (pp. 126-148). Princeton, NJ: Princeton University Press.
- 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(2), 191–217. DOI: [10.1002/pam.10023](https://doi.org/10.1002/pam.10023).
- Hoxby, C. M. (2003). School choice and school productivity (or could school choice be a tide that lifts all boats?). In C. M. Hoxby (Ed.), *The economics of school choice* (pp. 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. DOI: [10.3102/1076998609346968](https://doi.org/10.3102/1076998609346968).
- 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. Available at: <http://www.aei.org/publication/views-private-schools-attitudes-school-choice-programs-three-states/>.
- Krueger, A. B. & Zhu, P. (2004). Another look at the New York City school voucher experiment. *American Behavioral Scientist*, 47(5), 658–698. DOI: [10.1177/0002764203260152](https://doi.org/10.1177/0002764203260152).
- Lauder, H. & Hughes, D. (1999). *Trading in futures: Why markets in education don't work*. Berkshire, United Kingdom: Open University Press.
- Lubienski, C. A. & Lubienski, S. T. (2013). *The public school advantage: Why public schools outperform private schools*. Chicago, IL: University of Chicago Press.
- McShane, M. Q., Kisida, B., Jensen, L. I., & Wolf, P. J. (2012, February). *Milwaukee Parental Choice Program: Descriptive report on participating schools 2010-11* (Milwaukee Evaluation Report #33). School Choice Demonstration Project, University of Arkansas, Fayetteville, AR. <http://www.uaedreform.org/milwaukee-parental-choice-program-descriptive-report-on-participating-schools-2010-11/>
- Mill, J. S. (1962). In M. Warnock (Ed.), *Utilitarianism, on liberty, essay on Bentham*. New York: Meridian. (Original work published 1869).
- 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*, 39(3), 464-484. DOI: [10.3102/0162373717693108](https://doi.org/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>
- Mills, J. N. & Wolf, P. J. (2019, April). *The effects of the Louisiana Scholarship Program on student achievement after four years*. EDRE working paper no. 2019-10. Social Science Research Network, April 23.
- Montiel Olea, J. L. & Pflueger, C. E. (2013). A robust test for weak instruments. *Journal of Business and Economic Statistics*, 31(3), 358-369. DOI: 10.1080/00401706.2013.806694.
- Montiel Olea, J. L., Pflueger, C. E., & Wang, S. (2013). A robust test for weak instruments in Stata. *Stata Journal*. Accessed at: <https://www.carolinpflueger.com/MontielPfluegerWang20131104.pdf>.
- 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(2), 553-602. Retrieved from <http://www.jstor.org/stable/2586913>.
- Shakeel, M. D., Anderson, K. P., & Wolf, P. J. (2016, May). *The participant effects of private school vouchers across the globe: A meta-analytic and systematic review*. EDRE working paper 2016-07. Social Science Research Network, May 10. Available at: <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>.
- Smith, K.B., & Meier, K. (1995). *The case against school choice: Politics, markets and fools*. New York: Routledge.
- Staiger, D. & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3), 557-586. Accessed at: https://scholar.harvard.edu/files/stock/files/instrumental_variables_regression_with_weak_instruments.pdf.
- Stock, J. H. & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D W. K. Andrews & J. H. Stock (Eds.), *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, (Chapter 5, pp. 80-108). Cambridge, UK: Cambridge University Press.

- Sude, Y., DeAngelis, C. A., & Wolf, P. J. (2018). Supplying choice: An analysis of school participation decisions in voucher programs in Washington, DC, Indiana, and Louisiana. *Journal of School Choice*, 12(1), 8-33. DOI: [10.1080/15582159.2017.1345232](https://doi.org/10.1080/15582159.2017.1345232).
- Sude, Y. & Wolf, P. J. (2019, April). *Do you get cream with your choice? Characteristics of students who moved into or out of the Louisiana Scholarship Program*. EDRE working paper no. 2019-13. Social Science Research Network, April 23.
- 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), pp. 783-808. DOI: [10.1002/pam.22086](https://doi.org/10.1002/pam.22086).
- 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., & Egalite, A. J. (forthcoming). Does private school choice improve student achievement? A review of the evidence. In M. Berends, R. J. Waddington, & J. A. Schoenig (Eds.), *School choice at the crossroads: Research perspectives*. New York: Routledge.
- 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. & Hoople, D. S. (2006). Looking inside the black box: What school factors explain voucher gains in Washington, DC? *Peabody Journal of Education*, 81(1), 7-26. DOI: [10.1207/S15327930pje8101_2](https://doi.org/10.1207/S15327930pje8101_2).
- 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. DOI: [10.1002/pam.21691](https://doi.org/10.1002/pam.21691).

Appendix

Table A1.

First stage F-statistics for primary analysis, one-year impacts for the 2012-13 cohort

	Simple Model				Student Characteristics Model			
	<i>n</i>	<i>F</i> -statistic for lottery (<i>mediator</i>)	<i>F</i> -statistic for interaction	<i>F</i> -statistic for lottery (non- <i>mediator</i>)	<i>n</i>	<i>F</i> -statistic for lottery (<i>mediator</i>)	<i>F</i> -statistic for interaction	<i>F</i> -statistic for lottery (non- <i>mediator</i>)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ELA								
Religious	1,795	668.61	568.52	115.91	1,772	120.56	51.05	15.07
Catholic	1,795	668.61	284.31	442.42	1,772	120.56	31.22	49.66
Coeducational	1,795	668.61	597.60	96.86	1,772	120.56	54.44	14.08
Urban	1,795	668.61	296.35	399.28	1,772	120.56	28.46	44.15
Tuition	1,422	303.59	90.07		1,418	54.02	23.31	
Full-Time Equivalents	1,525	317.92	90.54		1,504	57.22	25.19	
Student/Teacher Ratio	1,525	302.49	56.53		1,504	52.62	19.01	
Total Enrollment	1,525	342.85	207.43		1,504	60.91	51.90	
% African American	1,525	303.68	163.71		1,504	51.96	31.71	
% LSP	1,525	292.11	108.40		1,504	49.49	28.05	
School Year Days	1,525	296.59	47.00		1,504	49.54	14.38	
School Day Hours	1,525	294.33	87.22		1,504	49.90	25.33	
Total Instructional Hours	1,525	297.05	71.52		1,504	50.46	19.14	
Math								
Religious	1,795	665.07	564.17	116.36	1,772	120.13	50.97	15.88
Catholic	1,795	665.07	284.31	436.48	1,772	120.13	32.14	48.40
Coeducational	1,795	665.07	592.96	97.27	1,772	120.13	54.44	14.06
Urban	1,795	665.07	295.92	395.57	1,772	120.13	28.30	44.17
Tuition	1,421	301.31	87.86		1,417	54.06	22.73	
Full-Time Equivalents	1,524	315.90	89.62		1,503	58.39	25.25	
Student/Teacher Ratio	1,524	299.46	55.80		1,503	52.45	16.65	
Total Enrollment	1,524	341.49	204.47		1,503	61.91	52.22	
% African American	1,524	301.85	162.09		1,503	52.24	31.56	
% LSP	1,524	289.63	108.58		1,503	49.47	30.70	
School Year Days	1,524	294.14	49.37		1,503	49.41	14.07	
School Day Hours	1,524	292.36	87.03		1,503	49.91	22.48	
Total Instructional Hours	1,524	294.81	71.36		1,503	50.39	17.98	

Table A2.

First stage F-statistics for tiered analysis, one-year impacts for the 2012-13 cohort

	Bottom Tercile		Middle Tercile		Top Tercile	
	<i>n</i>	<i>F</i> -statistic	<i>n</i>	<i>F</i> -statistic	<i>n</i>	<i>F</i> -statistic
	(1)	(2)	(3)	(4)	(5)	(6)
ELA						
Tuition	599	236.41	564	436.49	259	53.14
FTE	375	89.19	640	275.17	510	350.59
Student/Teacher Ratio	365	246.93	532	150.84	628	245.04
Total Enrollment	305	116.00	717	232.07	503	327.16
% African American	50	35.20	608	262.31	867	280.48
% LSP	33	18.67	354	137.14	1,138	419.90
School Year Days	131	92.68	1,167	444.95	227	52.51
School Day Hours	313	112.02	460	157.77	752	308.54
Total Instructional Hours	489	205.16	332	94.81	704	283.35
Math						
Tuition	599	232.54	563	435.42	259	53.14
FTE	374	86.83	641	275.85	509	349.56
Student/Teacher Ratio	364	237.93	533	151.28	627	244.51
Total Enrollment	305	112.45	717	232.07	502	326.17
% African American	50	35.20	608	262.31	866	276.52
% LSP	33	18.67	353	132.31	1,138	420.25
School Year Days	132	94.05	1,165	439.53	227	52.51
School Day Hours	312	108.53	460	157.77	752	309.11
Total Instructional Hours	489	201.39	332	94.81	703	282.81

Table A3.

First stage F-statistics for primary analysis, two-year impacts for the 2012-13 cohort

	Simple Model				Student Characteristics Model			
	<i>n</i>	<i>F</i> -statistic for lottery (<i>mediator</i>)	<i>F</i> -statistic for interaction	<i>F</i> -statistic for lottery (non- <i>mediator</i>)	<i>n</i>	<i>F</i> -statistic for lottery (<i>mediator</i>)	<i>F</i> -statistic for interaction	<i>F</i> -statistic for lottery (non- <i>mediator</i>)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ELA								
Religious	1,572	201.47	178.68	30.90	1,551	30.79	17.33	4.56
Catholic	1,572	201.47	119.26	84.24	1,550	30.79	14.34	9.39
Coeducational	1,572	201.47	176.93	30.74	1,551	30.79	16.83	4.55
Urban	1,572	201.47	97.64	116.54	1,551	30.79	11.35	12.87
Tuition	1,267	95.25	11.13		1,265	16.81	5.16	
Full-Time Equivalents	1,370	121.81	54.95		1,351	25.91	14.05	
Student/Teacher Ratio	1,370	94.22	13.18		1,351	16.97	4.81	
Total Enrollment	1,370	109.82	74.25		1,351	22.77	22.55	
% African American	1,370	120.94	45.83		1,351	21.85	7.35	
% LSP	1,337	117.42	26.28		1,318	22.18	5.19	
School Year Days	1,370	102.66	1.53		1,351	18.61	1.34	
School Day Hours	1,370	92.28	25.74		1,351	16.70	7.44	
Total Instructional Hours	1,370	93.62	17.25		1,351	17.15	5.20	
Math								
Religious	1,571	199.21	176.40	30.89	1,551	31.29	17.46	4.30
Catholic	1,571	199.21	116.51	84.44	1,550	31.29	14.53	8.54
Coeducational	1,571	199.21	174.66	30.74	1,550	31.29	16.98	3.79
Urban	1,571	199.21	95.37	116.73	1,550	31.29	11.03	12.88
Tuition	1,266	93.61	11.20		1,264	16.96	4.88	
Full-Time Equivalents	1,369	120.70	54.61		1,350	26.22	14.55	
Student/Teacher Ratio	1,369	93.05	12.17		1,350	17.28	3.62	
Total Enrollment	1,369	108.92	74.25		1,350	23.56	24.33	
% African American	1,369	120.24	44.82		1,350	22.12	6.96	
% LSP	1,336	116.77	25.49		1,317	22.58	5.15	
School Year Days	1,369	101.32	1.53		1,350	18.63	1.40	
School Day Hours	1,369	91.10	24.61		1,350	16.85	7.04	
Total Instructional Hours	1,369	92.20	16.65		1,350	17.13	5.64	

Table A4.

First stage F-statistics for tiered analysis, two-year impacts for the 2012-13 cohort

	Bottom Tercile		Middle Tercile		Top Tercile	
	<i>n</i>	<i>F</i> -statistic	<i>n</i>	<i>F</i> -statistic	<i>n</i>	<i>F</i> -statistic
	(1)	(2)	(3)	(4)	(5)	(6)
ELA						
Tuition	584	67.85	457	126.52	226	12.03
FTE	361	23.53	564	80.40	445	111.99
Student/Teacher Ratio	302	43.08	488	76.49	580	70.62
Total Enrollment	279	23.12	652	74.68	439	110.16
% African American	49	17.88	508	132.11	813	73.23
% LSP	23	19.87	302	58.30	1,012	129.68
School Year Days	112	17.32	1,086	134.67	172	37.16
School Day Hours	305	33.00	395	48.86	670	103.46
Total Instructional Hours	448	38.06	291	40.19	631	119.04
Math						
Tuition	583	68.02	457	126.52	226	11.10
FTE	361	22.47	563	80.34	445	111.99
Student/Teacher Ratio	301	43.28	487	76.42	581	68.52
Total Enrollment	278	23.16	652	72.35	439	110.16
% African American	49	17.88	507	132.05	813	71.67
% LSP	23	19.87	301	58.81	1,012	127.17
School Year Days	112	17.32	1,085	132.45	172	37.16
School Day Hours	305	33.15	395	48.86	669	100.63
Total Instructional Hours	448	38.18	291	40.19	630	115.72

Table A5.

First stage F-statistics for primary analysis, three-year impacts for the 2012-13 cohort

	Simple Model				Student Characteristics Model			
	<i>n</i>	<i>F</i> -statistic for lottery (<i>mediator</i>)	<i>F</i> -statistic for interaction	<i>F</i> -statistic for lottery (non- <i>mediator</i>)	<i>n</i>	<i>F</i> -statistic for lottery (<i>mediator</i>)	<i>F</i> -statistic for interaction	<i>F</i> -statistic for lottery (non- <i>mediator</i>)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ELA								
Religious	1,192	71.50	65.76	6.66	1,172	11.34	8.55	1.72
Catholic	1,192	71.50	48.06	25.46	1,171	11.34	8.50	3.81
Coeducational	1,192	71.50	62.52	9.13	1,172	11.34	8.34	1.61
Urban	1,192	71.50	34.74	36.77	1,172	11.34	5.71	4.68
Tuition	988	40.02	6.46		985	8.81	4.06	
Full-Time Equivalents	1,069	51.59	41.09		1,049	9.88	10.18	
Student/Teacher Ratio	1,069	33.09	7.34		1,049	8.45	3.01	
Total Enrollment	1,069	40.86	29.53		1,049	8.33	7.78	
% African American	1,069	45.52	18.02		1,049	10.13	5.49	
% LSP	949	44.18	8.52		930	10.23	3.33	
School Year Days	1,069	37.85	1.62		1,049	8.68	1.50	
School Day Hours	1,069	35.27	11.81		1,049	8.03	4.12	
Total Instructional Hours	1,069	36.38	5.57		1,049	7.99	2.91	
Math								
Religious	1,191	71.55	65.82	6.66	1,184	10.97	8.29	1.75
Catholic	1,191	71.55	48.06	25.49	1,171	10.97	8.48	3.65
Coeducational	1,191	71.55	62.57	9.13	1,171	10.97	8.13	1.64
Urban	1,191	71.55	34.73	36.82	1,171	10.97	5.49	4.99
Tuition	987	40.00	6.48		984	8.40	4.78	
Full-Time Equivalents	1,068	51.64	41.21		1,048	9.66	10.70	
Student/Teacher Ratio	1,068	33.10	7.33		1,048	8.23	2.51	
Total Enrollment	1,068	40.87	29.55		1,048	7.98	8.76	
% African American	1,068	45.53	18.03		1,048	9.92	5.18	
% LSP	948	44.37	8.52		929	9.87	3.25	
School Year Days	1,068	37.90	1.62		1,048	8.38	1.50	
School Day Hours	1,068	35.29	11.81		1,048	7.71	4.16	
Total Instructional Hours	1,068	36.40	5.57		1,048	7.65	2.96	

Table A6.

First stage F-statistics for tiered analysis, three-year impacts for the 2012-13 cohort

	Bottom Tercile		Middle Tercile		Top Tercile	
	<i>n</i>	<i>F</i> -statistic	<i>n</i>	<i>F</i> -statistic	<i>n</i>	<i>F</i> -statistic
	(1)	(2)	(3)	(4)	(5)	(6)
ELA						
Tuition	475	29.48	311	41.01	202	5.50
FTE	319	10.64	441	25.95	309	39.16
Student/Teacher Ratio	221	17.64	360	17.32	488	30.31
Total Enrollment	246	10.49	519	23.27	304	38.70
% African American	35	7.47	363	41.13	671	27.18
% LSP	14	8.57	221	26.37	714	46.35
School Year Days	88	15.72	883	42.44	98	14.30
School Day Hours	259	6.99	300	19.65	510	49.49
Total Instructional Hours	387	11.69	208	13.94	474	53.01
Math						
Tuition	474	29.54	311	41.01	202	5.50
FTE	318	10.65	441	25.95	309	39.16
Student/Teacher Ratio	220	17.68	360	17.32	488	30.31
Total Enrollment	245	10.49	519	23.27	304	38.70
% African American	35	7.47	363	41.13	670	27.21
% LSP	14	8.57	220	26.54	714	46.35
School Year Days	88	15.72	882	42.49	98	14.30
School Day Hours	258	7.00	300	19.65	510	49.49
Total Instructional Hours	386	11.70	208	13.94	474	53.01

Table A7.

First stage F-statistics for primary analysis, four-year impacts for the 2012-13 cohort

	Simple Model				Student Characteristics Model			
	<i>n</i>	<i>F</i> -statistic for lottery (<i>mediator</i>)	<i>F</i> -statistic for interaction	<i>F</i> -statistic for lottery (non- <i>mediator</i>)	<i>n</i>	<i>F</i> -statistic for lottery (<i>mediator</i>)	<i>F</i> -statistic for interaction	<i>F</i> -statistic for lottery (non- <i>mediator</i>)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ELA								
Religious	735	16.16	13.73	2.62	715	5.27	3.13	1.26
Catholic	735	16.16	7.03	15.00	715	5.27	2.11	2.84
Coeducational	735	16.16	14.09	2.15	715	5.27	3.23	1.13
Urban	735	16.16	9.30	7.06	715	5.27	2.19	2.20
Tuition	599	12.94	5.39		596	4.25	2.03	
Full-Time Equivalents	652	19.02	35.24		632	4.84	9.79	
Student/Teacher Ratio	652	7.39	7.79		632	3.29	4.02	
Total Enrollment	652	12.93	13.09		632	4.55	4.88	
% African American	652	9.93	5.80		632	3.14	3.00	
% LSP	542	10.70	1.74		523	3.39	1.52	
School Year Days	652	14.76	3.30		632	5.01	1.18	
School Day Hours	652	7.34	4.28		632	3.32	3.10	
Total Instructional Hours	652	7.81	3.30		632	3.47	3.17	
Math								
Religious	723	13.55	12.08	1.50	706	4.32	2.83	1.12
Catholic	723	13.55	5.84	13.03	703	4.32	2.07	2.59
Coeducational	723	13.55	12.43	1.15	703	4.32	2.88	1.03
Urban	723	13.55	7.91	5.82	703	4.32	2.03	2.23
Tuition	590	11.38	6.06		587	3.81	2.19	
Full-Time Equivalents	643	17.72	30.36		623	4.80	10.28	
Student/Teacher Ratio	643	6.40	6.80		623	2.89	3.82	
Total Enrollment	643	10.92	10.66		623	3.78	5.58	
% African American	643	8.58	4.61		623	2.96	2.66	
% LSP	536	9.35	1.49		517	2.90	1.64	
School Year Days	643	15.43	4.99		623	5.09	1.41	
School Day Hours	643	6.42	3.96		623	2.86	3.12	
Total Instructional Hours	643	6.75	3.23		623	2.95	3.40	

Table A8.

First stage F-statistics for tiered analysis, four-year impacts for the 2012-13 cohort

	Bottom Tercile		Middle Tercile		Top Tercile	
	<i>n</i>	<i>F</i> -statistic	<i>n</i>	<i>F</i> -statistic	<i>n</i>	<i>F</i> -statistic
	(1)	(2)	(3)	(4)	(5)	(6)
ELA						
Tuition	290	7.76	179	6.66	130	5.79
FTE	207	5.70	277	2.94	168	15.73
Student/Teacher Ratio	139	6.63	216	0.66	297	14.67
Total Enrollment	213	8.22	274	1.55	165	15.53
% African American	23	4.97	226	5.14	403	6.82
% LSP	10		120	13.05	412	8.51
School Year Days	42	1.90	548	9.73	62	3.29
School Day Hours	183	0.86	202	9.14	267	8.14
Total Instructional Hours	257	0.94	146	13.24	249	8.36
Math						
Tuition	285	7.32	175	6.53	130	4.20
FTE	205	5.74	273	2.59	165	13.70
Student/Teacher Ratio	138	6.64	211	0.66	294	11.65
Total Enrollment	209	8.38	272	1.28	162	13.71
% African American	22	2.96	225	4.95	396	5.88
% LSP	10		120	10.02	406	7.54
School Year Days	40	1.99	541	8.16	62	3.29
School Day Hours	181	0.87	202	7.71	260	7.09
Total Instructional Hours	253	0.94	146	10.80	244	7.22

Table A9.

Variation in LATE estimates by religious status, coeducational status, and urban setting, one-year impacts for the 2012-13 cohort

		ELA		Math	
		Simple	Student Characteristics	Simple	Student Characteristics
		(1)	(2)	(3)	(4)
Religious	<i>n</i>	1,516	1,495	1,515	1,494
	τ_1	-0.17*	-0.17**	-0.53***	-0.60***
		(0.10)	(0.07)	(0.10)	(0.10)
Non-Religious	<i>n</i>	279	277	280	278
	τ_1	-0.42	-0.30	-0.75**	-0.73***
		(0.28)	(0.21)	(0.31)	(0.27)
Difference	τ_2	0.25	0.13	0.23	0.14
		(0.29)	(0.21)	(0.32)	(0.28)
Catholic	<i>n</i>	946	927	946	927
	τ_1	-0.21	-0.19*	-0.61***	-0.74***
		(0.14)	(0.11)	(0.16)	(0.14)
Non-Catholic	<i>n</i>	849	845	849	845
	τ_1	-0.21*	-0.19**	-0.51***	-0.51***
		(0.12)	(0.08)	(0.13)	(0.12)
Difference	τ_2	0.00	0.01	-0.10	-0.23
		(0.18)	(0.13)	(0.21)	(0.18)
Coeducational	<i>n</i>	1,499	1,478	1,498	1,477
	τ_1	-0.16*	-0.16*	-0.52***	-0.60***
		(0.09)	(0.07)	(0.10)	(0.10)
Single-Gender	<i>n</i>	296	294	297	295
	τ_1	-0.52*	-0.37*	-0.78**	-0.74***
		(0.29)	(0.21)	(0.31)	(0.27)
Difference	τ_2	0.36	0.21	0.25	0.15
		(0.30)	(0.22)	(0.32)	(0.28)
Urban	<i>n</i>	895	875	894	874
	τ_1	-0.14	-.19*	-0.38***	-0.51***
		(0.13)	(0.10)	(0.13)	(0.13)
Suburban, Town, or Rural	<i>n</i>	900	897	901	898
	τ_1	-0.28**	-0.19*	-0.73***	-0.72***
		(0.13)	(0.10)	(0.15)	(0.13)
Difference	τ_2	0.14	0.00	0.35*	0.21
		(0.18)	(0.14)	(0.20)	(0.18)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A1).

Table A10.

Variation in LATE estimates by school tuition, number of full-time equivalents, and student/teacher ratio, one-year impacts for the 2012-13 cohort

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Tuition (demeaned, average = \$5,394)				
<i>n</i>	1,422	1,418	1,421	1,417
Estimated effect at average	-0.12 (0.10)	0.13* (0.07)	-0.48*** (0.11)	-0.59*** (0.10)
Interaction (per \$1,000)	0.04 (0.08)	0.03 (0.06)	0.12 (0.09)	0.06 (0.08)
Full-Time Equivalents (demeaned, average = 21.0)				
<i>n</i>	1,525	1,504	1,524	1,503
Estimated effect at average	-0.17* (0.10)	-0.17** (0.08)	-0.57*** (0.10)	-0.65*** (0.09)
Interaction (per 10 FTE)	0.01 (0.05)	0.00 (0.03)	0.13*** (0.04)	0.11*** (0.04)
Student/Teacher Ratio (demeaned, average = 13.0)				
<i>n</i>	1,525	1,504	1,524	1,503
Estimated effect at average	-0.21** (0.09)	-0.18** (0.07)	-0.57*** (0.11)	-0.61*** (0.10)
Interaction	0.04 (0.03)	0.01 (0.02)	0.04 (0.03)	0.00 (0.03)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A1).

Table A11.

Variation in LATE estimates by total enrollment and enrollment demographics, one-year impacts for the 2012-13 cohort

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Total Enrollment (demeaned, average = 275)				
<i>n</i>	1,525	1,504	1,524	1,503
Estimated effect at average	-0.18*	-0.17**	-0.61***	-0.68***
	(0.10)	(0.08)	(0.11)	(0.10)
Interaction (per 100 students)	0.03	0.01	0.12	0.10
	(0.03)	(0.02)	(0.03)	(0.03)
% African American (demeaned, average = 44.6%)				
<i>n</i>	1,525	1,504	1,524	1,503
Estimated effect at average	-0.08	-0.09	-0.49***	-0.65***
	(0.12)	(0.10)	(0.14)	(0.11)
Interaction (per 10pp)	-0.03	-0.03	-0.01	0.02
	(0.03)	(0.02)	(0.04)	(0.03)
% LSP (demeaned, average = 25.2%)				
<i>n</i>	1,525	1,504	1,524	1,503
Estimated effect at average	-0.21*	-0.18*	-0.49***	-0.60***
	(0.12)	(0.09)	(0.13)	(0.11)
Interaction (per 10pp)	0.02	0.00	-0.02	0.00
	(0.04)	(0.03)	(0.04)	(0.04)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A1).

Table A12.

Variation in LATE estimates by instructional intensity, one-year impacts for the 2012-13 cohort

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Days per school year (demeaned, average = 178.8)				
<i>n</i>	1,525	1,504	1,524	1,503
Estimated effect at average	-0.17*	-0.17**	-0.52***	-0.60***
	(0.10)	(0.07)	(0.10)	(0.09)
Interaction (per 10 days)	0.00	-0.02	0.11	0.03
	(0.17)	(0.09)	(0.13)	(0.12)
Hours per school day (demeaned, average = 7.1)				
<i>n</i>	1,525	1,504	1,524	1,503
Estimated effect at average	-0.17*	-0.17**	-0.54***	-0.62***
	(0.09)	(0.07)	(0.10)	(0.09)
Interaction	0.03	0.03	0.20	0.22
	(0.15)	(0.11)	(0.14)	(0.10)
Total instructional hours (demeaned, average = 1,268.0)				
<i>n</i>	1,525	1,504	1,524	1,503
Estimated effect at average	-0.17*	-0.17*	-0.54***	-0.62***
	(0.09)	(0.07)	(0.10)	(0.09)
Interaction (per 100 hours)	0.01	0.01	0.10*	0.10**
	(0.07)	(0.05)	(0.06)	(0.05)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A1).

Table A13.

Variation in LATE estimates by religious status, coeducational status, and urban setting, two-year impacts for the 2012-13 cohort

		ELA		Math	
		Simple	Student Characteristics	Simple	Student Characteristics
		(1)	(2)	(3)	(4)
Religious	<i>n</i>	1,360	1,341	1,359	1,340
	τ_1	-0.12 (0.16)	-0.13 (0.11)	-0.25 (0.18)	-0.33** (0.15)
Non-Religious	<i>n</i>	212	210	212	210
	τ_1	-0.16 (0.52)	-0.06 (0.43)	-0.40 (0.63)	-0.47 (0.46)
Difference	τ_2	0.04 (0.54)	-0.06 (0.45)	0.15 (0.64)	0.15 (0.48)
Catholic	<i>n</i>	866	848	866	848
	τ_1	-0.19 (0.20)	-0.17 (0.15)	-0.40 (0.25)	-0.49*** (0.17)
Non-Catholic	<i>n</i>	706	703	705	702
	τ_1	-0.04 (0.24)	-0.06 (0.18)	-0.11 (0.24)	-0.15 (0.24)
Difference	τ_2	-0.15 (0.30)	-0.11 (0.23)	-0.29 (0.33)	-0.34 (0.28)
Coeducational	<i>n</i>	1,344	1,325	1,343	1,324
	τ_1	-0.09 (0.16)	-0.11 (0.12)	-0.21 (0.18)	-0.30** (0.15)
Single-Gender	<i>n</i>	228	226	228	226
	τ_1	-0.36 (0.57)	-0.18 (0.42)	-0.64 (0.75)	-0.65 (0.51)
Difference	τ_2	0.27 (0.60)	0.07 (0.44)	0.43 (0.75)	0.36 (0.52)
Urban	<i>n</i>	815	797	816	798
	τ_1	-0.02 (0.22)	-0.13 (0.15)	0.15 (0.18)	-0.05 (0.16)
Suburban, Town, or Rural	<i>n</i>	757	754	755	752
	τ_1	-0.24 (0.23)	-0.11 (0.18)	-0.74*** (0.28)	-0.66*** (0.21)
Difference	τ_2	0.22 (0.32)	-0.02 (0.23)	0.88*** (0.32)	0.61** (0.26)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A3).

Table A14.

Variation in LATE estimates by school tuition, number of full-time equivalents, and student/teacher ratio, two-year impacts for the 2012-13 cohort

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Tuition (demeaned, average = \$5,454)				
<i>n</i>	1,267	1,265	1,266	1,264
Estimated effect at average	-0.02 (0.18)	-0.07 (0.13)	-0.08 (0.21)	-0.19 (0.17)
Interaction (per \$1,000)	0.07 (0.17)	0.04 (0.11)	0.26 (0.26)	0.21 (0.19)
Full-Time Equivalents (demeaned, average = 21.1)				
<i>n</i>	1,370	1,351	1,369	1,350
Estimated effect at average	-0.12 (0.18)	-0.14 (0.12)	-0.30 (0.19)	-0.37** (0.15)
Interaction (per 10 FTE)	0.02 (0.07)	0.01 (0.05)	0.13* (0.07)	0.11* (0.06)
Student/Teacher Ratio (demeaned, average = 12.8)				
<i>n</i>	1,370	1,351	1,369	1,350
Estimated effect at average	-0.25 (0.17)	-0.19 (0.13)	-0.41* (0.21)	-0.40** (0.17)
Interaction	0.10 (0.07)	0.05 (0.04)	0.13* (0.07)	0.06 (0.06)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A3).

Table A15.

Variation in LATE estimates by total enrollment and enrollment demographics, two-year impacts for the 2012-13 cohort

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Total Enrollment (demeaned, average = 276)				
<i>n</i>	1,370	1,351	1,369	1,350
Estimated effect at average	-0.16 (0.18)	-0.16 (0.13)	-0.36* (0.20)	-0.41*** (0.15)
Interaction (per 100 students)	0.06 (0.05)	0.03 (0.04)	0.13** (0.05)	0.11** (0.05)
% African American (demeaned, average = 44.3%)				
<i>n</i>	1,370	1,351	1,369	1,350
Estimated effect at average	0.18 (0.19)	0.13 (0.15)	-0.06 (0.25)	-0.19 (0.18)
Interaction (per 10pp)	-0.11** (0.05)	-0.10** (0.04)	-0.07 (0.06)	-0.05 (0.05)
% LSP (demeaned, average = 21.2%)				
<i>n</i>	1,337	1,318	1,336	1,317
Estimated effect at average	0.09 (0.20)	0.11 (0.16)	-0.01 (0.22)	-0.09 (0.18)
Interaction (per 10pp)	-0.12 (0.12)	-0.17* (0.09)	-0.15 (0.13)	-0.15 (0.12)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A3).

Table A16.

Variation in LATE estimates by instructional intensity, two-year impacts for the 2012-13 cohort

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Days per school year (demeaned, average = 179.1)				
<i>n</i>	1,370	1,351	1,369	1,350
Estimated effect at average	-0.11 (0.17)	-0.13 (0.12)	-0.26 (0.19)	-0.33** (0.15)
Interaction (per 10 days)	0.13 (1.22)	0.12 (0.71)	0.38 (1.62)	0.31 (1.26)
Hours per school day (demeaned, average = 7.1)				
<i>n</i>	1,370	1,351	1,369	1,350
Estimated effect at average	-0.13 (0.17)	-0.15 (0.12)	-0.27 (0.20)	-0.35** (0.15)
Interaction	0.27 (0.26)	0.25 (0.17)	0.39 (0.29)	0.38* (0.21)
Total instructional hours (demeaned, average = 1,267.8)				
<i>n</i>	1,370	1,351	1,369	1,350
Estimated effect at average	-0.14 (0.17)	-0.16 (0.12)	-0.29 (0.20)	-0.36** (0.15)
Interaction (per 100 hours)	0.12 (0.12)	0.11 (0.08)	0.20 (0.14)	0.19* (0.10)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A3).

Table A17.

Variation in LATE estimates by religious status, coeducational status, and urban setting, three-year impacts for the 2012-13 cohort

		ELA		Math	
		Simple	Student Characteristics	Simple	Student Characteristics
		(1)	(2)	(3)	(4)
Religious	<i>n</i>	1,062	1,042	1,061	1,041
	τ_1	0.07	0.06	-0.08	-0.17
		(0.25)	(0.20)	(0.26)	(0.26)
Non-Religious	<i>n</i>	130	130	130	130
	τ_1	-0.46	-0.36	-0.41	-0.49
		(1.25)	(1.45)	(2.48)	(2.10)
Difference	τ_2	0.53	0.43	0.33	0.32
		(1.27)	(1.47)	(2.51)	(2.14)
Catholic	<i>n</i>	691	673	691	673
	τ_1	0.09	-0.04	-0.39	-0.62***
		(0.29)	(0.25)	(0.25)	(0.22)
Non-Catholic	<i>n</i>	501	499	500	498
	τ_1	-0.10	0.09	0.26	0.42
		(0.48)	(0.39)	(0.50)	(0.53)
Difference	τ_2	0.18	-0.13	-0.65	-1.04*
		(0.57)	(0.48)	(0.57)	(0.61)
Coeducational	<i>n</i>	1,047	1,027	1,046	1,026
	τ_1	0.11	0.08	-0.06	-0.17
		(0.26)	(0.21)	(0.27)	(0.26)
Single-Gender	<i>n</i>	145	145	145	145
	τ_1	-0.61	-0.45	-0.50	-0.44
		(0.90)	(0.93)	(1.02)	(1.15)
Difference	τ_2	0.72	0.54	0.44	0.27
		(0.94)	(0.94)	(1.05)	(1.19)
Urban	<i>n</i>	628	609	628	609
	τ_1	0.01	-0.04	-0.08	-0.33
		(0.41)	(0.34)	(0.41)	(0.47)
Suburban, Town, or Rural	<i>n</i>	564	563	563	562
	τ_1	0.02	0.07	-0.16	-0.08
		(0.33)	(0.26)	(0.36)	(0.35)
Difference	τ_2	-0.01	-0.11	0.08	-0.25
		(0.54)	(0.43)	(0.55)	(0.62)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A5).

Table A18.

Variation in LATE estimates by school tuition, number of full-time equivalents, and student/teacher ratio, three-year impacts for the 2012-13 cohort

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Tuition (demeaned, average = \$5,470)				
<i>n</i>	988	985	987	984
Estimated effect at average	0.18 (0.38)	0.14 (0.27)	0.00 (0.35)	-0.12 (0.31)
Interaction (per \$1,000)	0.00 (0.49)	0.01 (0.32)	0.07 (0.41)	0.06 (0.26)
Full-Time Equivalents (demeaned, average = 21.3)				
<i>n</i>	1,069	1,049	1,068	1,048
Estimated effect at average	0.03 (0.27)	0.02 (0.22)	-0.19 (0.25)	-0.29 (0.24)
Interaction (per 10 FTE)	0.11 (0.13)	0.13 (0.10)	0.27* (0.15)	0.29* (0.15)
Student/Teacher Ratio (demeaned, average = 12.9)				
<i>n</i>	1,069	1,049	1,068	1,048
Estimated effect at average	-0.09 (0.28)	0.01 (0.22)	-0.19 (0.32)	-0.19 (0.29)
Interaction	0.16 (0.11)	0.06 (0.08)	0.10 (0.10)	0.00 (0.11)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A5).

Table A19.

Variation in LATE estimates by total enrollment and enrollment demographics, three-year impacts for the 2012-13 cohort

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Total Enrollment (demeaned, average = 279)				
<i>n</i>	1,069	1,049	1,068	1,048
Estimated effect at average	-0.02 (0.28)	0.00 (0.22)	-0.26 (0.27)	-0.33 (0.25)
Interaction (per 100 students)	0.12 (0.09)	0.10 (0.07)	0.22* (0.11)	0.21** (0.11)
% African American (demeaned, average = 43.7%)				
<i>n</i>	1,069	1,049	1,068	1,048
Estimated effect at average	0.45* (0.25)	0.32 (0.24)	0.11 (0.28)	-0.05 (0.25)
Interaction (per 10pp)	-0.15* (0.08)	-0.10 (0.06)	-0.08 (0.08)	-0.06 (0.08)
% LSP (demeaned, average = 16.9%)				
<i>n</i>	949	930	948	929
Estimated effect at average	-0.02 (0.38)	0.10 (0.32)	0.12 (0.42)	0.18 (0.36)
Interaction (per 10pp)	0.17 (0.33)	0.02 (0.23)	-0.16 (0.39)	-0.33 (0.32)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A5).

Table A20.

Variation in LATE estimates by instructional intensity, three-year impacts for the 2012-13 cohort

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Days per school year (demeaned, average = 179.0)				
<i>n</i>	1,069	1,049	1,068	1,048
Estimated effect at average	0.05 (0.30)	0.05 (0.22)	-0.13 (0.28)	-0.22 (0.26)
Interaction (per 10 days)	0.18 (2.68)	0.17 (1.85)	0.39 (2.44)	0.35 (1.76)
Hours per school day (demeaned, average = 7.1)				
<i>n</i>	1,069	1,049	1,068	1,048
Estimated effect at average	0.03 (0.28)	0.03 (0.24)	-0.12 (0.29)	-0.23 (0.26)
Interaction	0.28 (0.38)	0.32 (0.35)	0.24 (0.32)	0.32 (0.32)
Total instructional hours (demeaned, average = 1,275.5)				
<i>n</i>	1,069	1,049	1,068	1,048
Estimated effect at average	0.03 (0.28)	0.03 (0.23)	-0.14 (0.28)	-0.24 (0.25)
Interaction (per 100 hours)	0.12 (0.18)	0.13 (0.16)	0.14 (0.14)	0.16 (0.14)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A5).

Table A21.

Variation in LATE estimates by religious status, coeducational status, and urban setting, four-year impacts for the 2012-13 cohort

		ELA		Math	
		Simple	Student Characteristics	Simple	Student Characteristics
		(1)	(2)	(3)	(4)
Religious	<i>n</i>	648	628	639	619
	τ_1	-0.14 (0.71)	-0.19 (0.57)	-0.93 (0.57)	-0.94 (0.84)
Non-Religious	<i>n</i>	87	87	84	84
	τ_1	-1.95 (59.16)	-1.29 (21.51)	-0.65 (9.74)	-0.28 (50.01)
Difference	τ_2	1.80 (59.12)	1.10 (21.55)	-0.27 (9.76)	-0.66 (50.51)
Catholic	<i>n</i>	436	418	433	415
	τ_1	-0.16 (0.82)	-0.16 (0.60)	-1.06 (2.08)	-1.17 (0.98)
Non-Catholic	<i>n</i>	299	297	290	288
	τ_1	-0.66 (1.29)	-0.60 (1.06)	-0.66 (0.91)	-0.39 (0.98)
Difference	τ_2	0.50 (1.48)	0.43 (1.23)	-0.39 (2.31)	-0.78 (1.47)
Coeducational	<i>n</i>	652	632	643	623
	τ_1	-0.14 (0.69)	-0.16 (0.61)	-0.89 (0.56)	-0.89 (1.38)
Single-Gender	<i>n</i>	83	83	80	80
	τ_1	-2.17 (55.47)	-1.62 (22.91)	-0.99 (9.84)	-0.68 (60.39)
Difference	τ_2	2.04 (55.43)	1.47 (23.05)	0.10 (9.88)	-0.21 (61.35)
Urban	<i>n</i>	397	378	394	375
	τ_1	-0.24 (1.88)	-0.48 (0.91)	-0.60 (2.15)	-0.92 (0.98)
Suburban, Town, or Rural	<i>n</i>	338	337	329	328
	τ_1	-0.50 (1.37)	-0.17 (0.90)	-1.22 (14.06)	-0.81 (2.04)
Difference	τ_2	0.26 (2.31)	-0.31 (1.35)	0.61 (14.18)	-0.11 (2.41)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A7).

Table A22.

Variation in LATE estimates by school tuition, number of full-time equivalents, and student/teacher ratio, four-year impacts for the 2012-13 cohort

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Tuition (demeaned, average = \$5,510)				
<i>n</i>	599	596	590	587
Estimated effect at average	-0.05 (5.07)	-0.13 (11.32)	-0.68 (2.65)	-0.73 (2.08)
Interaction (per \$1,000)	0.00 (0.01)	0.00 (0.01)	0.00 (0.00)	0.00 (0.00)
Full-Time Equivalents (demeaned, average = 21.5)				
<i>n</i>	652	632	643	623
Estimated effect at average	-0.08 (0.66)	-0.16 (0.54)	-0.93 (0.88)	-0.93 (0.92)
Interaction (per 10 FTE)	-0.16 (0.37)	-0.07 (0.31)	0.15 (0.64)	0.15 (0.80)
Student/Teacher Ratio (demeaned, average = 13.0)				
<i>n</i>	652	632	643	623
Estimated effect at average	-0.24 (0.72)	-0.20 (0.65)	-0.89 (0.61)	-0.84 (0.65)
Interaction	0.12 (0.24)	0.02 (0.22)	-0.01 (0.16)	-0.09 (0.21)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A7).

Table A23.

Variation in LATE estimates by total enrollment and enrollment demographics, four-year impacts for the 2012-13 cohort

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Total Enrollment (demeaned, average = 280)				
<i>n</i>	652	632	643	623
Estimated effect at average	-0.11 (1.15)	-0.17 (0.53)	-0.92 (1.41)	-0.92 (1.90)
Interaction (per 100 students)	-0.06 (0.32)	-0.03 (0.23)	0.09 (0.33)	0.07 (1.32)
% African American (demeaned, average = 43.0%)				
<i>n</i>	652	632	643	623
Estimated effect at average	-0.22 (0.62)	-0.21 (0.59)	-0.71 (0.60)	-0.73 (0.48)
Interaction (per 10pp)	0.04 (0.24)	0.01 (0.16)	-0.09 (0.43)	-0.08 (0.14)
% LSP (demeaned, average = 13.3%)				
<i>n</i>	542	523	536	517
Estimated effect at average	-0.50 (4.51)	-0.44 (3.42)	-0.80 (2.70)	-0.79 (1.49)
Interaction (per 10pp)	0.68 (5.77)	0.38 (4.60)	0.07 (3.06)	-0.20 (4.41)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A7).

Table A24.

Variation in LATE estimates by instructional intensity, four-year impacts for the 2012-13 cohort

	ELA		Math	
	Simple	Student Characteristics	Simple	Student Characteristics
	(1)	(2)	(3)	(4)
Days per school year (demeaned, average = 179.1)				
<i>n</i>	652	632	643	623
Estimated effect at average	-0.25 (25.13)	-0.31 (1.03)	-0.95 (13.41)	-0.94 (3.39)
Interaction (per 10 days)	0.29 (298.17)	0.34 (22.51)	0.14 (169.98)	0.13 (56.02)
Hours per school day (demeaned, average = 7.1)				
<i>n</i>	652	632	643	623
Estimated effect at average	-0.16 (15.71)	-0.19 (0.60)	-0.89 (2.41)	-0.89 (0.59)
Interaction	0.55 (7.84)	0.69 (0.74)	-0.09 (1.56)	0.15 (0.62)
Total instructional hours (demeaned, average = 1,275.5)				
<i>n</i>	652	632	643	623
Estimated effect at average	-0.21 (4.32)	-0.26 (0.65)	-0.89 (1.30)	-0.91 (0.58)
Interaction (per 100 hours)	0.20 (0.80)	0.24 (0.48)	0.00 (0.34)	0.06 (0.26)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes. Bootstrapped standard errors account for clustering within risk set. Estimates that use instruments that do not satisfy Staiger and Stock's (1997) rule of thumb for a strong instrument are shaded in gray (see Appendix Table A7).