

More findings about school vouchers and test scores, and they are still negative

Mark Dynarski and Austin Nichols

Executive Summary

Vouchers to pay for students to attend private schools continue to command public attention. The current administration has proposed vouchers in its budget, and more than half of states are operating or have proposed voucher programs.

Four recent rigorous studies—in the District of Columbia, Louisiana, Indiana, and Ohio—used different research designs and reached the same result: on average, students that use vouchers to attend private schools do less well on tests than similar students that do not attend private schools. The Louisiana and Indiana studies offer some hints that negative effects may diminish over time. Whether effects ever will become positive is unclear.

Test scores are not the only education outcome and some observers have downplayed them, citing older evidence that voucher programs increase high school graduation and college-going. We lack evidence that the current generation of voucher programs will yield these longer-term outcomes. We also lack evidence of how public schools and private schools differ in their instructional and teaching strategies that would explain negative effects on test scores. Both questions should be high on the research agenda.

Vouchers to pay for students to attend private schools continue to command public attention. The current administration has proposed vouchers in its budget, and more than half of states are operating or have proposed voucher programs.

[Dynarski wrote in this forum](#) last year about recent studies that had shown negative effects of vouchers on test scores in Louisiana and Indiana. Since that time, new studies of vouchers in DC and Ohio have been released, and the Louisiana and Indiana studies released findings from an additional year.

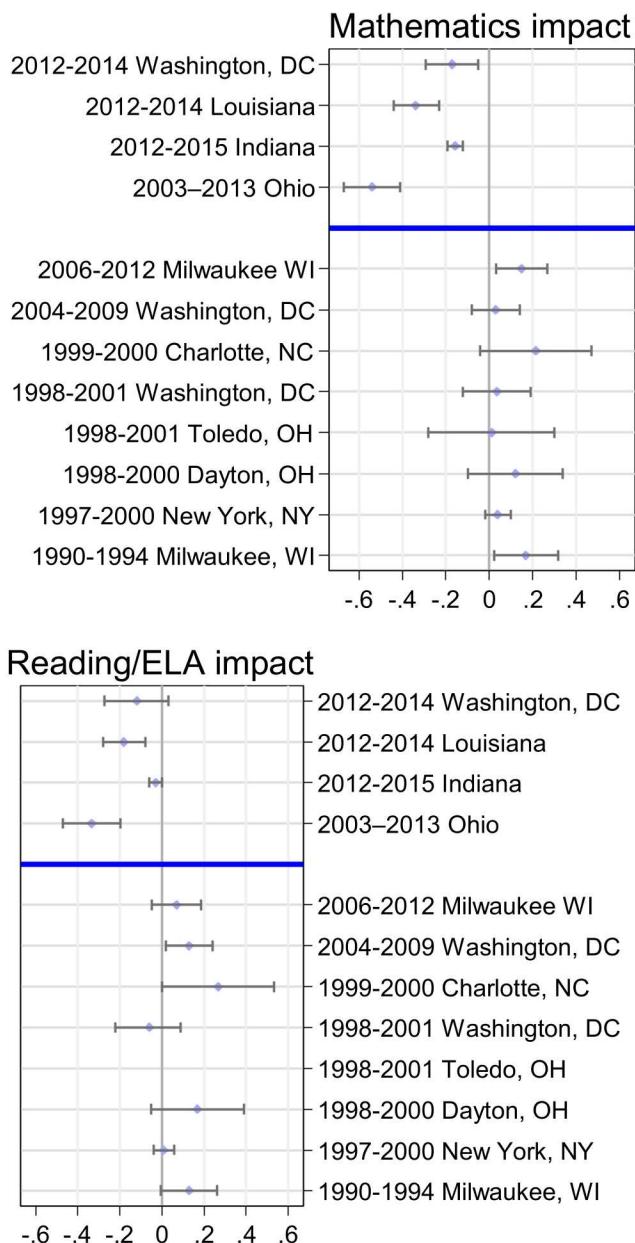
The four different studies use four different designs but arrive at the same result: on average, students that use vouchers to attend private schools do less well on tests than similar students that do not attend private schools. With voucher programs expanding rapidly and with each of the four studies measuring effects of vouchers differently, it's worth unpacking each study a bit to see what they say and do not say about effects of vouchers.

Table 1, in the appendix, compares features of the four studies, including the populations served by the programs, sample sizes of the studies, the test that studies used as their outcome, and how the studies measured impacts on those tests.

Figure 1, below, shows measures of test-score impacts, starting with the four studies at the top, and then effects on test scores from previous studies, roughly in reverse historical order. The point in the middle of the bars for each study is the estimate of the score effect, which is negative in both subjects in all four studies. The bars are confidence intervals for the estimated effects—when confidence intervals include zero, the effect is not statistically significant. Below the blue divider, we show effects for eight prior studies of vouchers. These show some positive effects for both subjects, though most are not statistically significant.

The question is why the pattern of recent studies differs from previous studies. As Dynarski had written previously, public schools and private schools have been under different accountability pressures for the last 15 years or so, which might explain some of the findings. Recognizing that researchers often call for more research, we think that call is merited here. It is rare for policy initiatives to be expanding in the face of evidence that those initiatives may have negative effects on key outcomes.

Figure 1. Findings from four current studies of vouchers and eight previous studies



Source: For discussion and references see <http://www.abtassociates.com/Perspectives/March/Do-Vouchers-Lead-to-Greater-Learning-Depends-Where.aspx>.

The District of Columbia Opportunity Scholarship Program

This study is a classic ‘field experiment’ consistent with the authorizing legislation that called for the program to be studied using the ‘strongest appropriate design.’

¹ Students selected to receive a voucher could attend private schools that agreed to accept the voucher as payment, which was more than half of all private schools in the District. Students and families had no obligation to use the voucher and, after a year, the study reported that about 30 percent of students in fact had not used their vouchers. This is a useful reminder that being offered a voucher expands options for parents but does not by itself require parents to do anything.

The study administered the *Terra Nova* test at the time students applied for vouchers (generally spring or early summer), and again about a year later. It also collected other data about students and families such as demographic characteristics, parent education, length of time at current residence, and parent ratings of the child’s current school. These characteristics were used in statistical models to adjust for whatever differences remained between students who were offered and not offered vouchers.

The findings showed that after one year, students who had been offered a voucher scored lower on the math part of the test, and the amount by which they were lower was statistically significant (the difference could not be explained by random variation). Students also scored lower on the reading test but the amount by which they were lower was not statistically significant.²

The study considered three possible explanations for the negative results. One was that students not offered vouchers went on to attend high-performing public schools (either traditional or charter schools). This did not explain much, however—students not offered vouchers attended public schools that had achievement levels that were average for the District. A second possible explanation was that students did less well on tests because they were adjusting to new schools. This explanation also did not hold up, in part because more than half of students *not* offered a voucher also switched schools, either because they had to (such as students who were moving from an elementary school to a middle school) or because they wanted to.

The third explanation was that private schools provided less instruction in reading and math. Data from a survey of principals that the study administered found that instructional time was lower by about an hour a week in both subjects, about twelve minutes a day. The District was not unusual in this regard—the difference in instructional time between private and public schools was about the same as the National Center for Education Statistics reported from a national survey of principals.³ But it’s at least plausible that students in private schools may have scored lower because they received less instruction in reading and math.

The Louisiana Scholarship Program

The Louisiana Scholarship Program (LSP) began in 2012.⁴ It is a statewide program and almost 10,000 students applied in its first year, making it considerably larger than the DC program, which averaged about 600 eligible applicants a year during the three years when students were enrolled in the study sample. Private schools that elected to participate by accepting vouchers as payment also had to administer the Louisiana state assessment to voucher-receiving students and were graded by the state using the same A-F scheme the state used for its public schools. Private schools whose voucher-receiving students scored poorly and received low grades from the state could be removed from the program.

The study of the LSP is an experiment but it is more complex than the one in DC. The lottery at the heart of the LSP experiment was conducted only when schools did not have enough available spaces at a grade level for the number of students that wanted to attend that school and grade level. A school may have had enough spaces for the number of applying fourth-graders, for example, but not enough spaces for the number of applying third-graders. That would have triggered a third-grade lottery at the school. The combination of applicant priorities, preferences parents expressed for schools, and available spaces resulted in a complex structure with 150 different lotteries, which required a complex analytic approach to measure voucher effects that is described in study reports.

The study estimated that students using vouchers had lower math scores on the Louisiana state assessment—in fact, scores were quite a lot lower. The study presented results for two samples, one that was restricted to students who had baseline scores

because they had previously participated in the state tests in public school before they applied for a voucher, and another that included the full sample of students that had a test score three years later regardless of whether they had a baseline score. Scores were negative and statistically significant for the full sample, but less negative and not statistically significant for the sample that was restricted to students with baseline scores. Experiments do not have to use baseline data to estimate effects because simple differences of outcomes at follow-up are effects of the program. And using larger samples can yield more precise estimates. It depends on whether the sample is sufficiently larger to offset not having baseline test scores. In this case, our preference is for the results from the full sample, but the results from both samples point in the same direction.

Media reporting of the findings pointed to the larger negative effects in the first year and smaller negative effects in the third year as good news.⁵ This is an odd conclusion. There are different arguments for vouchers, such as that they would give parents more choice, reduce the role of government in education, enable parents to transmit values and religion to their children, and deliver cost-effective education. But certainly one of the arguments for vouchers is to enable students to thrive academically in private schools. If this is the case, there should have been no catching up to do in the first place, beyond whatever adjustments students need to make when they change schools. And it's noteworthy that Louisiana students have not yet caught up after three years.

Some commenters have concluded that the negative effects in Louisiana were the consequence of the program being 'over-regulated.'⁶ But the conclusion that the Louisiana program was overregulated relies on unstated premises that private schools that agreed to participate were academically inferior to ones that did not agree but would have if the state did not impose requirements, or that regulation itself impairs academic achievement. Evidence of either is noticeably lacking in the argument. Also, the other three studies discussed here do not have the same regulatory structure.

The Indiana Choice Scholarship Program

Indiana currently operates the largest school voucher program in the country. More than 34,000 students received vouchers to attend more than 300 private

schools in the recently ended (2016-2017) school year. Unlike other voucher programs, Indiana gives vouchers to students living in relatively middle-income families, though students living in families closer to the poverty line are eligible for larger vouchers. And, unlike other states operating voucher programs, Indiana requires its private schools to administer the state assessment. Private schools are not new to the test.

The recently released study of the program examines its effects on test scores for students that have used vouchers for one, two, three, or four years.⁷ These are not the same students—a student that uses a voucher for, say, two years, and then returns to a public school, is not in the sample of students that used a voucher for three or four years. In the study's sample of students used to measure effects, the number of students that used a voucher for one year is ten times larger than the number that used a voucher for four years (Appendix Table 1).

Indiana's program did not use lotteries and the research team used quasi-experimental approaches to measure effects. It did this by matching students who switched schools and used vouchers with students who did not, and compared outcomes at later points. The matching approach is the equivalent of looking at a large crowd and picking out a person who most looks like you. A student who is using a voucher and is attending fifth grade, has family income near the poverty line, a particular race or ethnicity, and has low math and reading test scores, for example, would be matched to one or more students who are also attending fifth grade, have incomes near the poverty line, are of that race or ethnicity, and have low reading and math scores, but do not use vouchers.⁸

This approach sounds a lot like an experiment, but it differs on a crucial dimension—the characteristics of students or families that explain why some did and some did not use vouchers may not be the same. For example, voucher-using students might have more motivation to succeed academically, or parents of those students might be so inclined, or parents may have attended private schools themselves and want their children to attend them, too. There can also be 'negative' selection, such as if students struggling in public schools are more likely to use vouchers. In either case, these 'unobserved' variables get in the way because students using vouchers may have had different academic outcomes even if there were no voucher program. Not being able to control for these unobserved variables is what separates quasi-experiments from experiments. Lotteries, which are

true experiments, are blind to unobserved variables and they end up equally distributed among those who win them and lose them.

The study takes pains to look at alternative matching approaches and different ways to estimate effects on test scores. But the main finding is the same as the other two studies discussed above. Students who used their vouchers to switch from public to private schools were more likely to score less well in math, and were about the same in reading.⁹

The study notes that students using the voucher for more years appear to have smaller negative effects, but, as noted above, these are not the same students being followed for more years, which is the case in Louisiana (and will be in future reports for the DC study). They are different students that have used vouchers for longer periods. That some students used vouchers for longer periods puts more strain on the matching method because the case that unobserved variables are affecting their outcomes gets stronger. A useful opportunity exists here to explore differences between 'long stayers' and 'short stayers,' which may improve our understanding of which kinds of students benefit from voucher programs.

The Ohio EdChoice Program

The Ohio 'EdChoice' program provided vouchers for more than 18,000 students in the 2013-2014 school year, and recent legislated changes are likely to expand this number. Ohio did not conduct lotteries, but state assessments were administered to all students receiving vouchers, and the study matched students using vouchers to similar students that did not use them.¹⁰

One of the eligibility criteria for the Ohio program was that the public school that students currently were attending had to score below a threshold on the Ohio 'Performance Indicator' measure. The study used that threshold to identify schools near the threshold, and it matched students in schools that were on one side of the threshold with students that were in schools on the other side. Doing so is likely to reduce issues about unobserved variables, though the study acknowledges that it pays a price in terms of the representativeness of the findings. Most schools are well above or well below the threshold and they are not represented in the sample.

Comparing scores on the Ohio state assessment

for matched students found large negative effects for mathematics and for reading. The other three studies found evidence of negative effects for math—the Ohio study is the only one that found negative and statistically significant effects for reading as well. The main findings were not affected when the study estimated different kinds of models and made the sample larger by including students that became eligible for a voucher in any year after the program initially started in 2007.

The Ohio study also looked at whether the program led to changes in academic achievement for students that were in schools that were close to being eligible for the program. It found that students in these schools had higher academic achievement, a 'competitive effect' that echoes a previous study of competitive effects in Florida.¹¹ Competitive effects are interesting because they potentially include many students not using vouchers but benefiting academically from the voucher program. However, they also create a tension. Students using vouchers experience academic losses that are larger than the academic gains experienced by students not using vouchers.

Where are we now?

Four recent studies, four different programs, different research approaches, but the same general finding—using vouchers to attend private schools leads to lower math scores and, in one study, lower reading scores too.

Some previous studies showed positive outcomes for older students such as higher graduation rates and higher college-going rates.¹² Citing these and other studies, Greene has argued that test scores should be downplayed because they are weakly correlated with adult outcomes such as college-going and earnings.¹³

This argument begs the question about how large correlations should be to be considered as indicators of adult outcomes, and it also discounts recent research showing that test scores improvements related to effective teachers were correlated with gains in adult labor-market outcomes.¹⁴ This research suggests being very cautious when presented with evidence about public programs producing negative effects on test scores. Researchers need to consider ways to measure other outcomes that are meaningful in the debate, such as by designing studies with long follow-up periods to enable future research on high school graduation, college-going, and labor-market outcomes.

It means waiting longer for answers, but the value of knowing the answers is clear.

None of the four studies unpacked the education that is happening inside the public and private schools that study participants attend. The Indiana study mentions using qualitative approaches to interview private school administrators about their experiences adjusting to incoming voucher students, and that seems like a fruitful vein. There are a range of tools that researchers could use here—value-added measures that distinguish between the level of a school's test

scores and gains of students on test scores (gains probably are what parents care about, and levels are a noisy signal of gains), school climate surveys, teacher observation instruments, descriptions of curricula. Without these measurements, we really have no idea how private and public schools compare in how they go about educating students.

If the four studies suggest anything, it's that private schools have no secret key that unlocks educational potential.

1 <https://ies.ed.gov/ncee/pubs/20174022/pdf/20174022.pdf>

2 The previous study of the OSP found no statistically significant differences in reading and math after one year. See <https://ies.ed.gov/ncee/pubs/20074009/>.

3 <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2017076>.

4 Results for the second and third years of the Louisiana study are at <http://www.uaedreform.org/louisiana-sseep-evaluation/>.

5 See for example, the July 9 *Wall Street Journal* editorial "New Evidence on School Vouchers: Some Optimistic Findings From Indiana and Louisiana."

6 See, for example, <http://educationnext.org/the-folly-of-overregulating-school-choice/>

7 http://creo.nd.edu/images/people/Waddington_Berends_Indiana_Voucher_Impacts_06.24.17.pdf

8 What is described here is 'exact matching.' Propensity score matching is quite a bit more sophisticated than exact matching because it allows for tradeoffs between characteristics—for example, the algorithm might find that two students of the same race but different test scores are the closest match. But the general idea is the same.

9 The study report also notes that it cannot study effects of vouchers for students who do not switch from public to private schools. This happens, for example, if students enter a private-school kindergarten with

a voucher and stay in private schools thereafter.

By 2016, 52 percent of voucher students had not attended a public school (https://www.washingtonpost.com/local/education/how-indianas-school-voucher-program-soared-and-what-it-says-about-education-in-the-trump-era/2016/12/26/13d1d3ec-bc97-11e6-91ee-1addfe36cbe_story.html?utm_term=.36aa6c504945).

10 https://edex.s3-us-west-2.amazonaws.com/publication/pdfs/FORDHAM%20Ed%20Choice%20Evaluation%20Report_online%20edition.pdf.

11 See David Figlio & Cassandra M. D. Hart, 2014. "Competitive Effects of Means-Tested School Vouchers," *American Economic Journal: Applied Economics*, American Economic Association, vol. 6(1), pages 133-56.

12 The results from the first DC study showing higher graduation rates are at <https://ies.ed.gov/ncee/pubs/20104018/pdf/20104018.pdf>. Results showing higher college-going rates from the New York City voucher program are at http://paulepeterson.org/sites/default/files/Impacts_of_School_Vouchers_FINAL.pdf.

13 See <http://educationnext.org/mostly-care-test-scores-private-school-choice-not/> and also <https://jaypgreene.com/2016/11/05/evidence-for-the-disconnect-between-changing-test-scores-and-changing-later-life-outcomes/>.

14 <https://www.aeaweb.org/articles?id=10.1257/aer.104.9.2633>.

Appendix

Table 1. Features of the four studies

	District of Columbia Opportunity Scholarship Program	Louisiana Scholarship Program	Indiana Choice Scholarship Program	Ohio Educational Choice Scholarship Program
What does the program do?	Students are eligible if they live in the District of Columbia, and their families have incomes below 185 percent of the federal poverty line. Vouchers are \$8,500 a year for students attending private elementary and middle schools, and \$12,500 a year for students attending private high schools. Amounts are adjusted annually for inflation.	Louisiana students are eligible if their family incomes are less than 250% of the federal poverty line, and students either are entering kindergarten or are enrolled in a public school with a C, D, or F letter grade under the state accountability system. The voucher is the lesser of 90% of the amount the state and local government provides in funding to the local school system, or tuition charged by the student's chosen private school. Median tuition at participating private schools currently is \$4,925. Average per pupil spending is \$8,500.	Students are eligible for a voucher for up to 90 percent of tuition at a participating private school if their family annual income is less than 100 percent of the amount to qualify for free or reduced-price lunch. They are eligible for a voucher for up to 50 percent of tuition if their family annual income is less than 150 percent of the amount to qualify for free lunch.	Students are eligible if they are currently attending a 'designated public school' (a low-performing school according to the state); will be assigned to one of the designated public schools for the upcoming school year; are currently attending a charter or community school whose assigned home school is a designated public school; are entering kindergarten and would be assigned to a designated public schools; or are enrolling in an Ohio school for the first time and would be assigned to a designated public school. The voucher is \$4,650 for grades K-8 and \$6,000 for grades 9-12, or private school tuition if it is lower.

	District of Columbia Opportunity Scholarship Program	Louisiana Scholarship Program	Indiana Choice Scholarship Program	Ohio Educational Choice Scholarship Program
How did the study select students for the sample?	Eligible applicants in spring of 2012, 2013, and 2014 were selected by lottery. The study sample includes those 1,771 students and their parents.	Eligible applicants in 2012 that had applied to schools that were oversubscribed in the student's grade level were selected by lottery. Two samples that were analyzed were 2,746 students that had test-score data in the third year, and 1,206 students that had test score data in the third year and a test score at the time of application.	The study matched 3,276 low-income eligible applicants that moved from public schools to private schools to 12,406 students in public schools. Students were matched on test scores, gender, race and ethnicity, special education status, English Language Learner status, whether the student was retained in the baseline year, whether the student attended a charter or magnet school, and public school urban locale. The technical approach was propensity-score matching.	The main analysis for the study matched 445 students that moved from public to private schools for the 2007-2008 year to 7,491 students that attended schools that were close to being 'designated' but were not. Students are matched on prior-year reading and mathematics scores, gender, race and ethnicity, grade in 2006–07, and student history of economic disadvantage. The technical approach was propensity-score matching. Other analyses used the same methods to create matched samples for students applying in later years.
What test did the study use?	The study administered the Mc-Graw Hill <i>Terra Nova</i> , reading and math, in all grade levels.	The study used scores from the Louisiana Educational Assessment Program, which is given in grades three through eight and which private schools were required to administer to voucher students. In the third year, the study used the PARCC test.	The study used the Indiana state assessment, iSTEP. Most private schools in Indiana already administer the iSTEP as a requirement for accreditation.	The study used the Ohio state assessment, which voucher students are required to take.

	District of Columbia Opportunity Scholarship Program	Louisiana Scholarship Program	Indiana Choice Scholarship Program	Ohio Educational Choice Scholarship Program
How did the study measure impacts of vouchers on test scores?	The study measured impacts by comparing average test scores for students that had been offered a voucher to test scores for students that had not been offered a voucher. It used regression models to adjust for test scores at the time of application, grade levels, year, and student and family characteristics. It measured impacts for students who used their voucher by dividing the overall impact of the voucher offer by the percentage of students that used their voucher.	The study measured impacts using a two-stage regression model. The first stage predicted whether students used their voucher, with the offer of the voucher for the student's first-choice private school as an explanatory variable. The second stage included the predicted value from the first stage and other student and family characteristics.	The study measured impacts using regression models with a variable indicating that a student was a voucher user, and other student and family characteristics.	The study measured impacts using regression models with a variable indicating that a student was a voucher user, and other student and family characteristics.