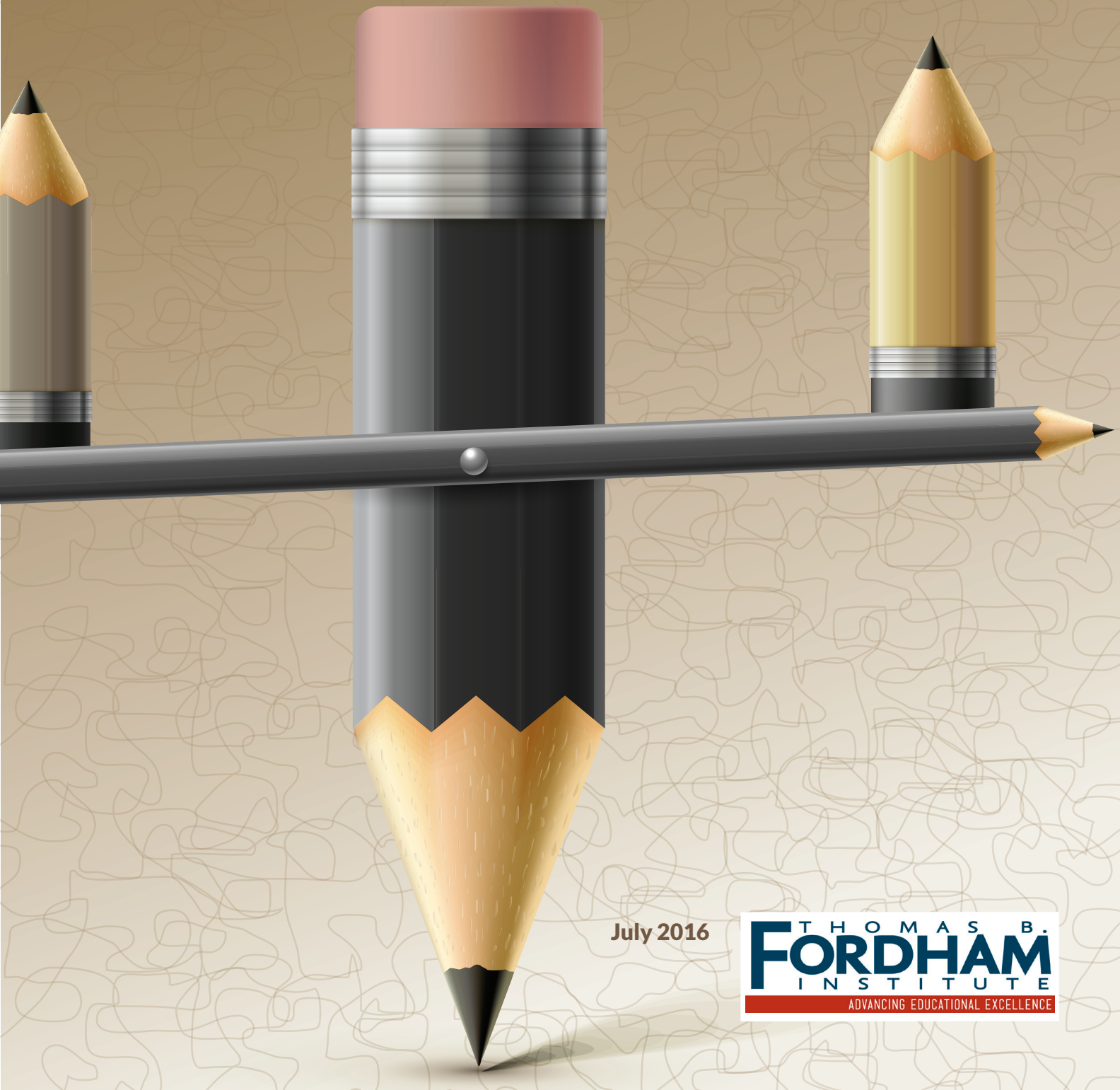
Evaluation of Ohio's EdChoice Scholarship Program:

Selection, Competition, and Performance Effects

By David Figlio and Krzysztof Karbownik



July 2016

THOMAS B. FORDHAM INSTITUTE
ADVANCING EDUCATIONAL EXCELLENCE

Evaluation of Ohio's EdChoice Scholarship Program:

Selection, Competition, and Performance Effects

By David Figlio and Krzysztof Karbownik

Foreword by Aaron Churchill and Chad L. Aldis

July 2016

The Thomas B. Fordham Institute is the nation's leader in advancing educational excellence for every child through quality research, analysis, and commentary, as well as on-the-ground action and advocacy in Ohio. It is affiliated with the Thomas B. Fordham Foundation, and this publication is a joint project of the Foundation and the Institute. For further information, please visit our website at www.edexcellence.net or write to the Institute at 100 E. Broad St., Suite 2430, Columbus, OH 43215. The Institute is neither connected with nor sponsored by Fordham University.

About the Authors

David Figlio is Director of the Institute for Policy Research and Orrington Lunt professor of Education and Social Policy and of Economics at Northwestern University. Figlio conducts research on a wide range of education and health policy issues from school accountability and standards to welfare policy and policy design. Figlio is the editor of the *Journal of Human Resources*, was the inaugural editor of *Education Finance and Policy*, and has served as associate editor of several other leading journals including the *American Economic Journal: Economic Policy* and *Journal of Urban Economics*. He is immediate past-president of the Association for Education Finance and Policy and a research associate at the National Bureau of Economic Research. Figlio joined Northwestern in 2008 from the University of Florida, where he was the Knight-Ridder Professor of Economics. He earned his Ph.D. in Economics from the University of Wisconsin-Madison in 1995.

Krzysztof Karbownik is a postdoctoral research fellow in the Institute for Policy Research at Northwestern University. His research focuses on family, health, labor, and education economics and the use of population-level administrative data sources to inform the public policy discussion. Karbownik is associated with CESifo in Germany, the Centre for Economic Analysis in Poland, and the Uppsala Center for Labor Studies in Sweden. He earned his Ph.D. from Uppsala University in 2014.

Contents

Foreword

By Aaron Churchill and Chad L. Aldis

Shortly after Ohio lawmakers enacted a new voucher program in 2005, the state budget office wrote in its fiscal analysis, "The Educational Choice Scholarships are not only intended to offer another route for student success, but also to impel the administration and teaching staff of a failing school building to improve upon their students' academic performance." As economist Milton Friedman had theorized decades earlier, Ohio legislators believed that increased choice and competition would boost education outcomes across the board. "Competition" in the words of Stanford's Caroline Hoxby, "would be the proverbial rising tide that lifts all boats."

Today, the EdChoice program provides publicly funded vouchers (or "scholarships") to more than eighteen thousand Buckeye students, youngsters previously assigned to some of the state's lowest-performing schools, located primarily in low-income urban communities.[1] That much is known. Yet remarkably little else is known about the program. Which children are using EdChoice when given the opportunity? Is the initiative faithfully working as its founders intended? Are participating students blossoming academically in their private schools of choice? Does the increased competition associated with EdChoice lead to improvements in the public schools that these kids left?

The present study utilizes longitudinal student data from 2003–04 to 2012–13 to answer these important questions. Specifically, the analysis utilizes the results from state tests—which all EdChoice students are required to take—to examine the vouchers' effects on two groups of pupils. First, the study inspects the scores of public school students who were eligible for vouchers—but did not take one—in order to gauge the competitive effects of EdChoice (i.e., its impact on traditional public school students and their schools). Second, it examines the academic impact of EdChoice on those students who actually use the vouchers to attend private schools.

This is the first study of EdChoice that uses individual student-level data, allowing for a rigorous evaluation of the program's effectiveness. (Earlier analyses by Matthew Carr and Greg Forster used school-level data to explore its competitive impact.) To lead the research, we tapped Dr. David Figlio of Northwestern University, a distinguished economist who has carried out examinations of Florida's tax credit scholarship program. He has also written extensively on school accountability, teacher quality, and competition. Given his experience, Dr. Figlio is exceptionally qualified to lead a careful, independent evaluation of Ohio's EdChoice program.

In this report, he sets forth three main findings:

- While the students who participate in EdChoice—the pupils who actually use a voucher to attend private schools—are primarily low-income and minority children, they are relatively less disadvantaged than other voucher-eligible students. Figlio reports that more than three in four participants are economically disadvantaged, and three in five are black or Hispanic. Viewed in relation to Ohio's public school population as a whole, students in EdChoice are highly disadvantaged—not surprising, given eligibility rules that require participants to have attended a low-achieving public school. But relative to students who are eligible for vouchers but choose not to use them, the participants in EdChoice are somewhat higher-achieving and somewhat less economically disadvantaged. This finding may be, in part, an artifact of the program's basic design: It allows private schools to retain control over admissions, and a child must gain admission into a private school before he or she can apply for a voucher. This multi-step process might be more easily navigated by relatively more advantaged families; their children might also be more likely to meet the private schools' admissions requirements.

[1] In June 2013, Ohio lawmakers created a new voucher program, referred to as the EdChoice Expansion program, for which eligibility is based on family income. This program is starting by phasing in kindergarteners and expanding by one grade level per year. The present research does not cover the income-based EdChoice Expansion. It is limited to the original EdChoice program for which eligibility depends on having attended a low-performing district school.

- EdChoice improved the achievement of the public school students who were eligible for the voucher but did not use it. When examining the test results of pupils attending public schools just above and below the eligibility threshold, the analysis finds that achievement in math and reading rose modestly as a result of voucher competition. (The analysis leverages the state's voucher eligibility rules to isolate voucher competition from other potential competitive effects, such as charter schools.) In other words, the voucher program has worked as intended when it comes to competitive effects. Importantly, this finding helps to address the concern that such programs may hurt students who remain in their public schools, either as a result of funds lost by those schools or the exodus of higher-performing peers. Quite the opposite has occurred in the case of EdChoice: Achievement improved when the voucher program was introduced and public schools faced stiffer competition (and the risk of losing their own students).

- The students who use vouchers to attend private schools have fared worse academically compared to their closely matched peers attending public schools. The study finds negative effects that are greater in math than in English language arts. Such impacts also appear to persist over time, suggesting that the results are not driven simply by the setbacks that typically accompany any change of school.

Let us acknowledge that we did not expect—or, frankly, wish—to see these negative effects for voucher participants; but it's important to report honestly on what the analysis showed and at least speculate on what may be causing these results. One factor might be related to the limits of credible evaluation: while the rigor of the methodology ensured "apples-to-apples" comparisons of student achievement, Dr. Figlio was limited to studying students who attended (or had left) public schools that were just above or below the state's cutoff for "low-performing." By definition, this group did not include the very lowest-performing schools in the state. It's possible that students who used a voucher to leave one of the latter schools might have improved their achievement; we simply cannot know from this study. The negative effects could also be related to different testing environments—higher stakes for public than private schools—or to curricular differences between what is taught in private schools and the content that's assessed on state tests. Finally, although this analysis does not enable us to identify individual schools as high- or low-performing, it may be the case that some of the private schools accepting EdChoice students are themselves not performing as well as they should.

***

Taken as a whole, the results reported here for Ohio's EdChoice program—one of the nation's largest voucher programs—are a mixed bag. The program benefitted, albeit modestly, thousands of public-school students; yet among the somewhat small number of participants studied here, the results are negative. The study mirrors important trends that can be seen in other voucher research. The modest, positive competitive effect on public school achievement replicates findings from jurisdictions like Florida, Louisiana, and Milwaukee, findings that also offered evidence that voucher competition improved public school outcomes. These are, of course, encouraging for advocates of competition and choice. Yet this study also extends a recent (and, to us, unwelcome) trend that finds negative effects for voucher participants in large statewide programs. While earlier evaluations of privately and publicly funded scholarship programs—usually administered at the city level—found neutral-to-positive impacts on participants, newer studies of Louisiana's and Indiana's statewide programs have uncovered negative results, particularly in math.

There's been much discussion about what might be behind these participant results. Is too much regulation discouraging high-quality private schools from joining the program? Are state exams failing to capture important private school contributions to student success? Do large, statewide programs lack the tools and resources to ensure quality at scale? Or are private schools simply struggling to raise achievement—especially in math—in relation to their public school counterparts? Some or all of these (or other) factors may be at work, but no one really knows for certain. More research on the effects of statewide voucher programs is obviously warranted.

Even though we don't have all the answers, we believe that thoughtful policy makers can draw from the extant

research as well as on-the-ground experience to give these programs the best chance of succeeding for more students, whether attending public or private schools. The pertinent lessons seem to us applicable both in states considering new private school choice programs and in states (like Ohio) that are seeking to improve an existing program.

First, we need to foster a healthy, competitive environment in K–12 education. A competitive jolt can awaken sleepy, lazy, or slipshod schools to clean up their act and attend more closely to the academic needs of their students. On the policy side, this means that lawmakers should continue to encourage a rich supply of school options, including not just private schools (in their many flavors, including religious and non-sectarian) but also public charter, STEM, and career and technical schools. At the same time, families can do their part by demanding more quality school choices. Competition and choice—two sides of the same coin—can incentivize all schools to work harder at meeting the needs of their pupils.

Second, policy makers should resist calls to pile more input-based regulations upon voucher-accepting private schools. Ohio's private schools already face heavier regulation than those in many states. For example, they must adhere to state operating standards and hire state-licensed or certified teachers. Most of this was true before EdChoice came along (which makes less likely the "overregulation" explanation for disappointing participant results, at least in Ohio). Policy makers should tread lightly when adding to schools' regulatory burdens: After all, freedom from regulation is precisely what makes private schools different and—for many—worth attending in the first place.

Third, as this study suggests, private schools likely vary when it comes to quality, and the public needs maximum transparency about this. Accordingly, state leaders should help families better understand the quality of their options by providing easy-to-compare information on the performance of voucher-accepting private schools. While Ohio already reports voucher students' proficiency rates at the school level (subject to FERPA limitations), we know that those results are likely to be conflated with non-schooling factors like family income. They are also hard to track down. To be fair to private schools that educate disadvantaged voucher pupils, we suggest the adoption of a value-added measure—a school quality indicator that is more poverty-neutral than conventional academic proficiency rates. States (including Ohio) should make sure that these academic outcomes for voucher-accepting private schools are easily accessible to parents, perhaps in a report-card-like format akin to those adopted for public schools. In Ohio, this would not add any additional testing or regulatory requirements on private schools.

Fourth, policy makers should craft simple, parent-friendly program rules. From the perspective of families, EdChoice is fairly complex, which may have influenced who participates in it. Eligibility hinges on public schools' annual ratings from the state—which can change from year to year—and the state has no obligation to notify parents of their children's eligibility. This means that families must bestir themselves to visit the state's website or seek eligibility information through other channels. To ensure awareness, states should require direct notification of eligibility from the state department of education or a competent nonprofit agency. (This should also happen when eligibility is based on income.) Making matters more complicated, current EdChoice application rules require eligible students first to gain admission to a private school; then *the school applies* to the state for a voucher. It would be far simpler for parents if they could apply directly to the state for a voucher and then shop for the right private school. This process would not only empower parents but also give policy makers a much clearer picture of the demand for vouchers.

The present report breaks important new ground, but it is by no means the final word on EdChoice. We still have much to learn, including whether vouchers impact non-testing outcomes such as post-secondary success. We also need a deeper understanding about the quality of individual private schools. But the information set forth in the pages that follow is critically important as thoughtful policy makers consider the design and implementation of voucher programs, both in Ohio and across the nation. Programs that aim to better the lives of children must face scrutiny from independent, credible evaluators. Even when its findings are unexpected and painful, rigorous, disinterested evaluation remains the best way to prod improvements and make progress toward the program's goals. In the case of EdChoice, the program appears to have met one of the two

objectives conceived by its founders: Competition has spurred some public school improvement. The challenge ahead is to forge a stronger EdChoice program, one that can lead to widespread academic improvements for children who take their scholarships to the state's private schools.

Acknowledgments

The time and talents of many individuals helped to create this report. In particular, we are deeply grateful to the report's authors, Dr. David Figlio and Dr. Krzysztof Karbownik, for leading this careful, rigorous evaluation of Ohio's largest voucher program. We are especially appreciative of Dr. Figlio's thoughtfulness throughout the drafting and review process. In addition, we extend special thanks to Dr. Patrick Wolf of the University of Arkansas for providing expert feedback on earlier drafts of this report. We express our deepest appreciation to the staff of the Ohio Department of Education for making this research possible. On the Fordham team, we thank Michael J. Petrilli and Chester E. Finn, Jr. for their helpful feedback on earlier drafts; both contributed significantly to the report's foreword. On the Fordham-Ohio team, we wish to thank our colleagues Jeff Murray, Jamie Davies O'Leary, and Jessica Poiner for their assistance in production and report dissemination. We are appreciative of the skillful work of copy editor Pamela Tatz and report designer Andy Kittles. This report was made possible through the generous support of the Walton Family Foundation and our sister organization, the Thomas B. Fordham Foundation.

Chad L. Aldis, Vice President of Ohio Policy and Advocacy

Aaron Churchill, Ohio Research Director

Thomas B. Fordham Institute

This report was funded under contract with the Thomas B. Fordham Institute. We are grateful to the Ohio Department of Education for providing the data used in this report, as well as for technical assistance in helping us to understand the data and programs; we are particularly thankful to Heather Boughton, Matt Cohen, Sue Cosmo, Eben Dowell, Colleen Grady, and Erin Whitt for their hard work and numerous conversations and explanations. We appreciate the helpful feedback from Chad Aldis, Aaron Churchill, Checker Finn, Mike Petrilli, and Patrick Wolf. The opinions and interpretations expressed in this report are those of the authors and not of the Ohio Department of Education or the Thomas B. Fordham Institute. All errors are our own.

- David Figlio and Krzysztof Karbownik

1. Introduction

In June 2005, the State of Ohio enacted the Educational Choice Scholarship Program (EdChoice, initially called the Educational Choice Scholarship Pilot Program), which offered scholarships to students assigned to public schools considered consistently poor-performing by the Ohio Department of Education, to take effect during the 2006–07 academic year.[1] At first, only students assigned to schools receiving the lowest rating, academic emergency, for three consecutive years were to be eligible for scholarships, which were worth up to $4,250 for elementary and middle school–aged students and up to $5,000 for high school students, but in Spring 2006 the rules for qualification were relaxed such that students assigned to schools rated either as under academic emergency or under academic watch for three consecutive years would be eligible for scholarships. In the 2006–07 academic year, 3,141 students from 99 traditional public schools attended private schools under the EdChoice program. In December 2006, the Ohio Legislature further relaxed the eligibility criteria such that students assigned to a school under either academic emergency or academic watch in two of the preceding three years would be eligible for EdChoice scholarships. As a consequence, the number of public schools with voucher-eligible students increased considerably, to 213 schools, and in the 2007–08 academic year 6,943 students attended private schools under the EdChoice program. By 2013–14, 18,080 students were attending private schools under the voucher program.[2] In the 2013–14 academic year, the program was further expanded to become available to economically disadvantaged students, regardless of school-quality measures; in the 2013–14 academic year, this expansion was for Kindergarteners only. However, in subsequent years, the program will phase in one grade at a time.

The purpose of this report is to provide an analysis of the effects of the EdChoice program on students and schools in Ohio using the most appropriate tools for causal inference possible, as determined by the authors, given how the program was implemented. Throughout the report, we present occasional references from the related scholarly literature; these references are intended to be representative but not comprehensive. Although the nature of program implementation precludes the use of experimental methods, it does still provide opportunities for quasi-experimental research designs. We investigate three interrelated questions:

(1) When students are offered the opportunity to attend private schools under the EdChoice program, which students ultimately attend private schools?

(2) What are the effects of the EdChoice program on the reading and mathematics performance of students who continue to attend traditional public schools?

(3) What are the effects of participation in the EdChoice program on the reading and mathematics performance of students who move to the private sector as a consequence of the program?

We make use of anonymized student-level data between the 2003–04 and 2012–13 academic years, provided by the Ohio Department Education, to address each of these questions in turn in the sections below. Ohio requires that EdChoice program participants in private schools take the Ohio Achievement Assessments, and the strong majority of these students complied with this requirement (and were successfully matched to Ohio Department of Education administrative data systems), starting in the 2007–08 academic year and especially beginning in the 2008–09 academic year.[3] In each case, we chose the years of the program that allowed for the most credible causal inference, regardless of whether the years were the most current. Our general summary of the evidence is as follows:

(1) There appears to be positive selection, as measured by prior academic performance and family advantage, among voucher-eligible students into private schools as part of the EdChoice program. Although a substantial majority of the students participating in the program, as well as their peers remaining in public schools, tend to be from low-income backgrounds, those students leaving for private schools under the program tend to be more advantaged and higher performing than their peers who were eligible to participate in the program but who remained in public schools.

(2) Although the estimates are sensitive to the specific assumptions made, and some assumptions lead to zero rather than positive findings, the evidence in general suggests that the EdChoice program improved the performance of students eligible to participate—most of whom remained in the public schools. The estimated improvements are typically in the range of one-eighth of the magnitude of the black-white test-score gap. This is particularly true regarding our analysis of schools newly eligible in 2007–08, the first year for which we feel relatively confident that we can make causal claims about the performance effects of the EdChoice program. Our research design estimates the competitive effects for public schools that are relatively high-performing compared to all eligible schools (that is, schools that have the highest performance index values but are still low enough to have students eligible for EdChoice scholarships), so we have less confidence in extrapolating these positive findings to public schools with considerably lower levels of performance.

(3) We can only credibly study the performance effects of moving to private schools under the EdChoice program for those students leaving comparatively high-achieving public schools. Those students, on average, who move to private schools under the EdChoice program tend to perform considerably worse than observationally similar students who remained in public schools. The magnitudes of this negative estimated effect are relatively large—around three times the positive estimated competitive effect (also estimated for the relatively high-performing public schools that were eligible to participate) of the EdChoice program. These differences cannot be explained by the disruptions associated with changing schools. It may be the case that there are less negative, or even positive, performance effects for students moving to private schools from lower-performing public schools, but we do not feel comfortable studying this group of students.

In summary, the evidence regarding the effects of EdChoice program suggests that while higher-performing students tend to leave public schools to attend private schools under the EdChoice program, the students who remain in the public schools—at least, those public schools that were comparatively high achieving—generally perform better on statewide tests as a consequence of EdChoice vouchers being available to students in a school. On the other hand, those students who leave these comparatively high-achieving public schools to go to private schools appear to perform worse than they would have had they remained in the public schools (which we estimate to have improved as a consequence of the introduction of EdChoice). Together, it appears that EdChoice has benefitted the majority of students, but the students who actually left the public schools—at least those on the margin of eligibility—perform worse on statewide tests. Although test performance is only one measure of educational success, these findings suggest that a detailed exploration of the possible causes of the negative test-score results (for instance, which private schools participate in the program, policies on school-grade retention, test-curriculum alignment, and the like) may be warranted.

2. History and background on the Ohio EdChoice Scholarship Program

In June 2005, Ohio lawmakers enacted the EdChoice program, with 2006–07 as the first year of implementation. A student's eligibility for an EdChoice voucher is premised on the academic performance of the traditional district school he or she is slated to attend. The rationale behind this model is captured in the following statement by the Ohio Office of Budget and Management: "The Educational Choice Scholarships are not only intended to offer another route for student success, but also to impel the administration and teaching staff of a failing school building to improve upon their students' academic performance."[4]

The initial EdChoice legislation (House Bill [HB] 66) defined a low-performing public school as one receiving three consecutive years of academic emergency ratings—the state's lowest classification. Shortly thereafter, legislators expanded the definition to also include schools receiving ratings of academic watch for three straight years (HB 530, enacted in March 2006). In December 2006, the legislature again modified the criteria by identifying low-performing schools based on whether they received ratings of either academic watch or academic emergency in two of the past three years (HB 79). After this flurry of early legislation, the eligibility criteria remained consistent until Ohio began its transition to A-to-F school ratings in 2012–13. To align with the new rating system, state law now designates EdChoice-eligible schools based on a D or F rating for two of the past three years on the state's key accountability measures (HB 555, enacted December 2012).

The state annually updates the list of designated EdChoice public schools. When the school ceases to meet the criteria, it is removed from the list. Students who received vouchers in previous years, however, remain eligible until they complete grade 12, provided they meet the following conditions: (1) they do not move to another district (though they remain eligible if assigned to another EdChoice-designated school in the new district); (2) they complete all required state achievement tests; and (3) they do not have more than twenty unexcused absences during a school year. Students across the state are eligible to participate in EdChoice if their assigned public schools meet the performance criteria defined in state law (except for Cleveland students, who are eligible to participate in a different voucher program).

Early EdChoice legislation (HB 79) set a cap on the number of available vouchers at 14,000. If the number of applications exceeded the cap, priority was, and still is, given to returning voucher recipients and new applicants who are low income. As the number of voucher applicants began to exceed the cap, a newly elected Ohio governor John Kasich and state lawmakers lifted the cap to 30,000 for 2011–12 and to 60,000 starting with the 2012–13 school year (HB 153, enacted in June 2011). To receive a voucher, parents and students first apply for admission to a participating private school. Once the student has been accepted, the private school submits a voucher application on behalf of the student.

As enacted in HB 66 in 2005, the initial amount of the EdChoice voucher was $4,250 for students in grades K–8 and $5,000 for students in grades 9–12 (or a smaller amount if the private school tuition is less than these amounts). Legislation passed in June 2015 (HB 64) that raised the maximum voucher amount to $4,650 for K–8 pupils and $5,900 for high school students (the high school amount increases to $6,000 starting in 2016–17). The state deducts the voucher amount from the state aid received by the student's district of residence. When a student's family income is less than or equal to 200 percent of the federal poverty level, private schools cannot charge tuition greater than the voucher amount. For families with income above this level, private schools may charge tuition that is equal to the difference between the voucher amount and the regular tuition rate.

Ohio defines two categories of private (or nonpublic) schools: chartered nonpublic schools and nonchartered, non-tax-supported schools. In order to accept EdChoice vouchers, a private school must be a chartered nonpublic school. (In this context, "charter" does not refer to a public charter school.) A chartered nonpublic school must be approved by the state and adhere to state operating standards. These schools retain the right to have admissions standards; can offer pupils education based on certain beliefs, values, or religions; and may

charge tuition. Chartered nonpublic schools receive a modest amount of state aid to offset administrative costs related to state regulation, and their students are entitled to district-provided transportation.

Since the program's inception, EdChoice students have been required to take all state exams, with results reported to the state. Although not part of the original EdChoice legislation, a provision in HB 1 (enacted July 2009) added a requirement for the state to publicly report proficiency rates disaggregated by voucher students' district of residence, by their private school of attendance, and by certain subgroups. These testing and transparency requirements remain in current state law.

According to the Fordham Institute's School Choice Regulations: *Red Tape or Red Herring?* (David Stuit and Sy Doan, January 2013), 39 percent of Ohio's private schools participated in the EdChoice program in 2009–10. Nonparticipation can be partly explained by the fact that many private schools are not located in the geographic vicinity of EdChoice-eligible students, as low-performing public schools are located primarily in the state's impoverished urban areas. Other nonparticipating private schools may be at capacity, while others may have elected not to accept voucher-bearing pupils.

3. Who attends private schools under the EdChoice program?

All students assigned to public schools that meet the performance criteria for EdChoice scholarships are categorically eligible to attend a private school under the program, but it is not obvious which students will be most likely to make use of a voucher. On the one hand, there exist examples from other states that suggest that relatively low-achieving students are more likely to use vouchers that are targeted at disadvantaged students, but there are also examples in which the nature of the selection is modestly positive.[5] On the other hand, there exists little evidence about the nature of selection in voucher systems that are based on the performance of schools rather than the family background of the student. It could be the case that the same factors that lead disadvantaged families to choose private schools under means-tested vouchers might be at play when voucher eligibility is linked to measured school performance, and given the correlation between average family advantage and measured school quality in Ohio, one might expect the voucher users to be disproportionately disadvantaged or low performing in the EdChoice context, as well. At the same time, it could be the case that relatively motivated families eligible for the voucher may be the group most likely to capitalize on the opportunity. Furthermore, unlike in the Florida context, where families can obtain a voucher before obtaining admission to a private school, in the EdChoice context families must obtain admission to a private school before they are able to seek a voucher.[6] These factors may make it more likely that higher-performing eligible students will be the voucher users in the Ohio context.[7]

In order to investigate this question, we measure the attributes of students moving to private schools under the EdChoice scholarship program versus those who were eligible to move but remained in the public schools. In order to obtain measures of student attributes, we are limited to students who have already spent at least a year in the public schools (and, in the case where we are looking at prior test scores as a measure of student selection, the students had to have been in public schools through at least third grade to be considered); therefore, we cannot describe the nature of selection for students who enter private schools under the voucher program in Kindergarten. As a consequence, we are limited principally to students who are changing schools at a nontraditional time (for example, not as Kindergarteners entering elementary school or as sixth graders entering middle school). This comparison is likely to favor those families motivated to change schools even when they do not have to do so.

With that proviso, we compare the attributes of students who chose a private school under the EdChoice program in each year to those who were eligible to do so but remained in the public schools. We make these comparisons along a number of lines: their most recent reading and mathematics test scores, whether the student has ever been observed as economically disadvantaged in school records, the student's gender, and the student's race.

We begin by comparing the test scores of students who transitioned to private schools under EdChoice to eligible students who remained in the public schools. In the figures below, reading and mathematics scores are standardized statewide in every year and every grade, and it is apparent that test scores for voucher-eligible students are much lower than the state average—an unsurprising fact, given that school-quality measures are based in large part on the test performance of students who would become eligible for a voucher. As can be seen in the figures, those students who move to private schools have, on average, considerably higher performance levels than those eligible students who remain in public schools. The observed gap between those who transitioned to private schools and those who remained is about half of the black-white test-score gap[8]—quite a large difference in prior performance. It does appear that in the most recent years of observation, the prior test-score gap between those who moved to private schools and those who remained in public schools shrank modestly, because the students who entered the EdChoice program in more recent years were not as far ahead of those who remained in the public schools as they were in the earlier years of the program. Nonetheless, the gap has remained large in recent years, as well.

Figure 1: Prior year standardized mathematics test score: EdChoice participants versus eligible nonparticipants, first year of program participation

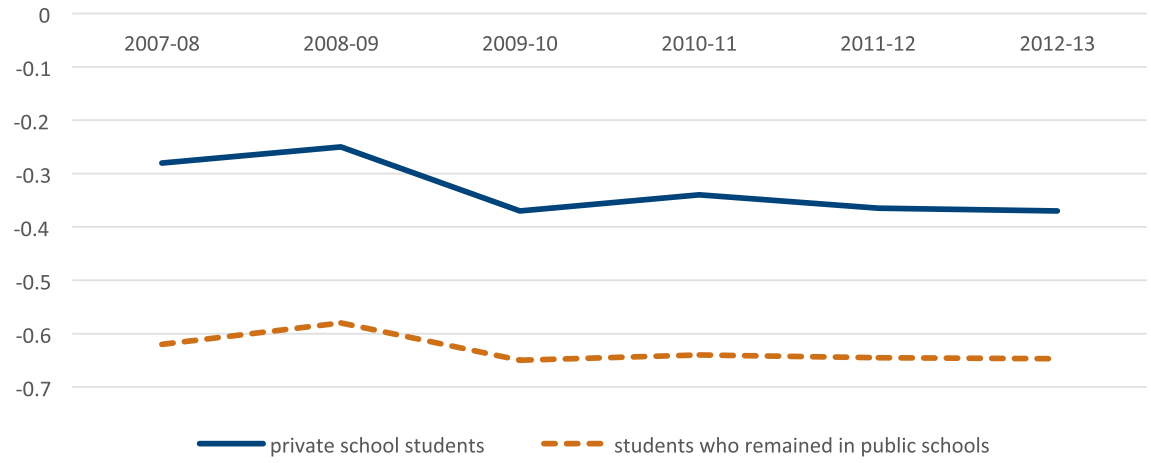
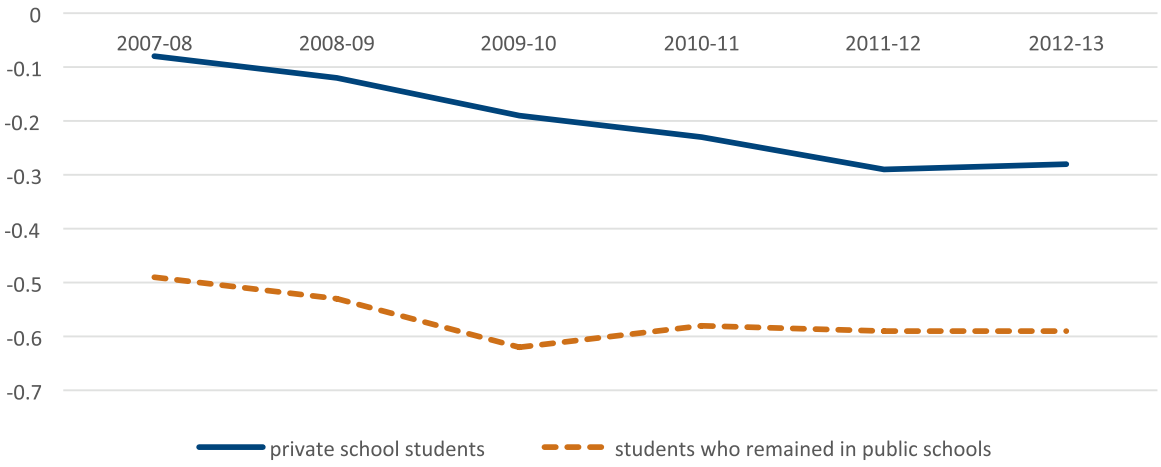


Figure 2: Prior year standardized reading test score: EdChoice participants versus eligible nonparticipants, first year of program participation



The preceding figures are based on those students observed moving to private schools in any given year compared to all eligible students who remained in the public schools regardless of how many years that they had the opportunity to change schools. Next, we perform the same analysis for students who made the choice to stay in public schools or left for private schools in their very first year of voucher eligibility. As can be seen in the figures below, the differences between movers and stayers in the public sector are similar regardless of whether we limit them to their first year of voucher eligibility or perform the comparison regardless of voucher eligibility.

Figure 3: Prior year standardized mathematics score: EdChoice participants versus eligible nonparticipants, first year of eligibility

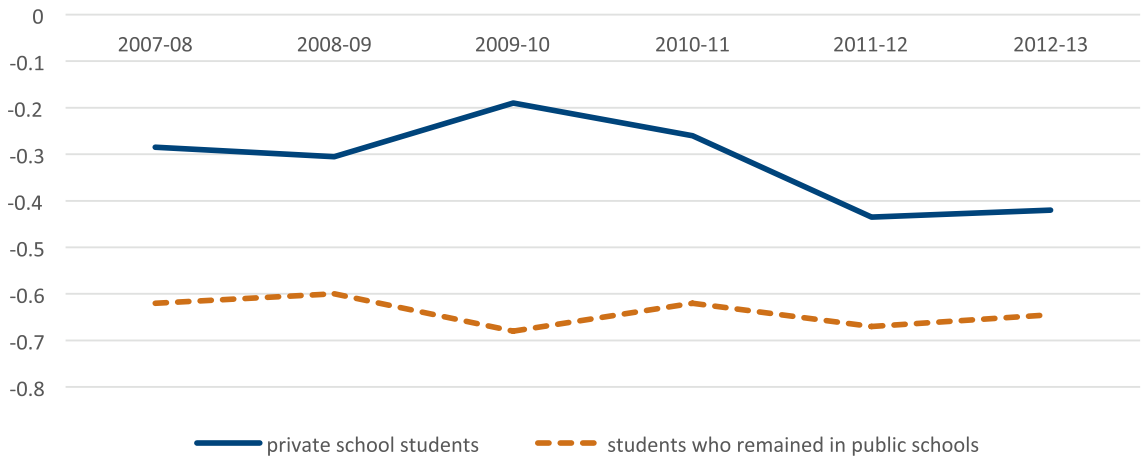
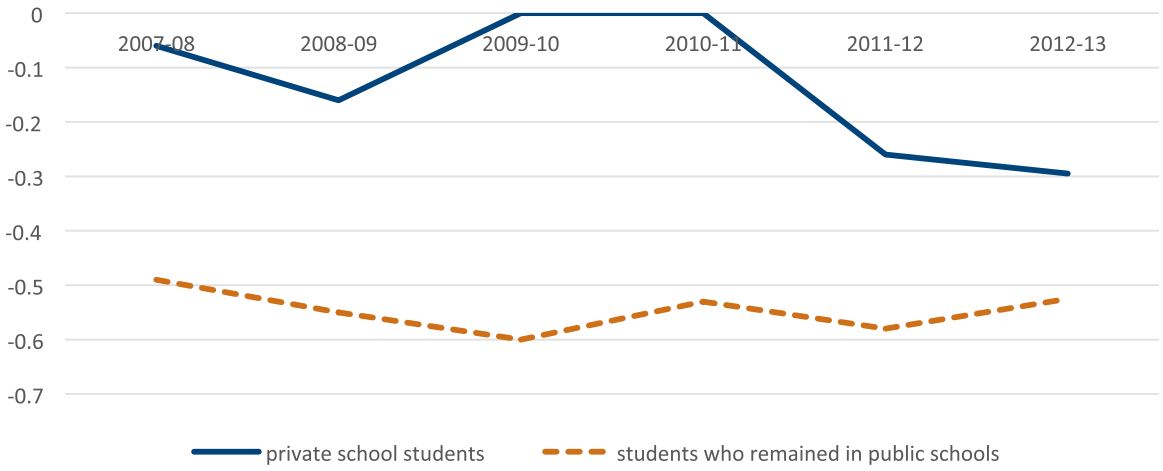


Figure 4: Prior year standardized reading score: EdChoice participants versus eligible nonparticipants, first year of eligibility



We observe similar patterns with regard to economic disadvantage. The overwhelming majority of students eligible for a voucher have been economically disadvantaged in the past. However, although around 95 percent of voucher-eligible students who remained in public schools have had a history of economic disadvantage,[9] the comparable figure tends to be around 85 percent for those who moved to a private school on an EdChoice scholarship. This gap, nevertheless, has been closing over time, and the narrowing of the gap in prior economic disadvantage is particularly pronounced in the case of first-time voucher-eligible students. At the same time, it is important to note that the overwhelming majority of students eligible for EdChoice scholarships, whether or not they make use of the vouchers, are economically disadvantaged, so we are comparing one very disadvantaged group to another very disadvantaged group.

Figure 5: Share of students ever economically disadvantaged: EdChoice participants versus eligible nonparticipants, first year of program participation

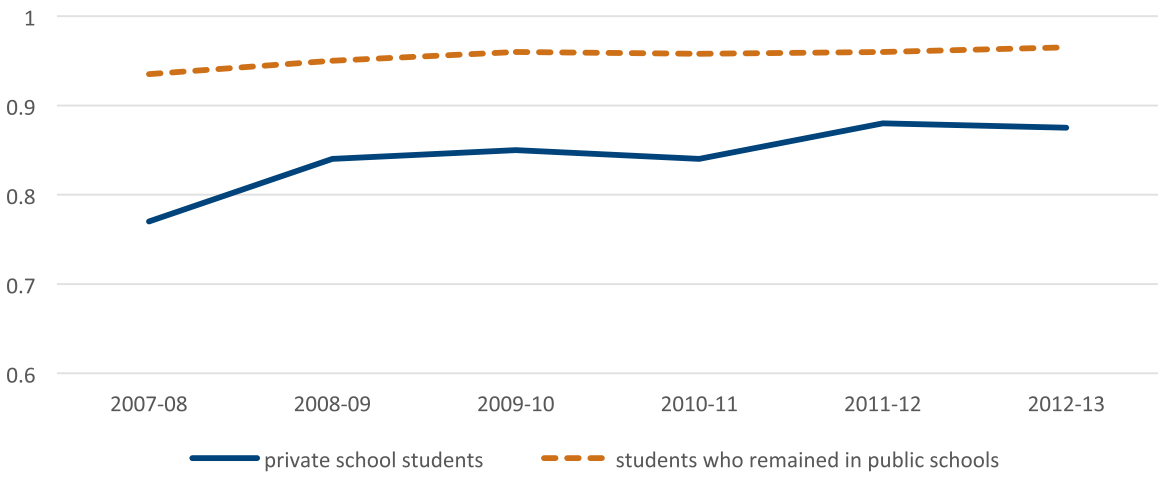
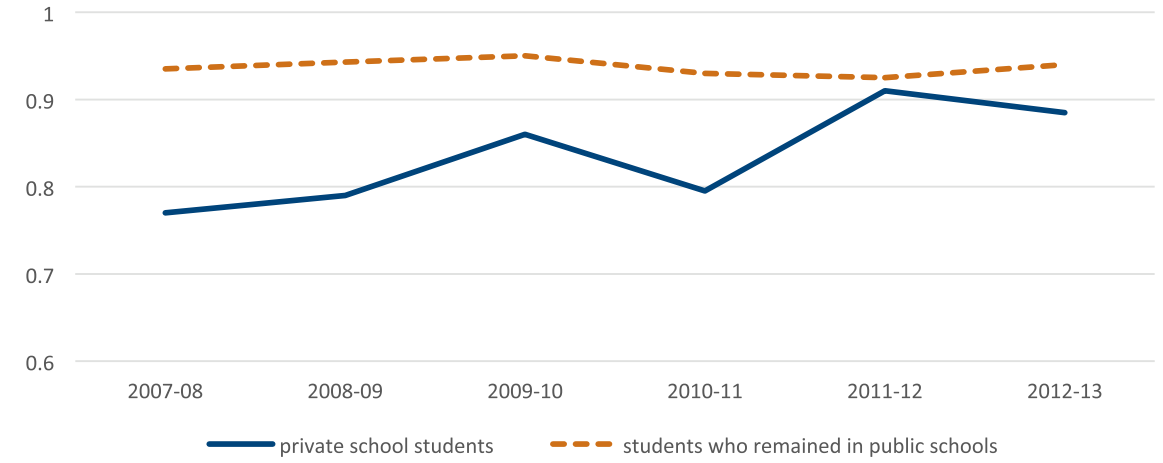


Figure 6: Share of students ever economically disadvantaged: EdChoice participants versus eligible nonparticipants, first year of program eligibility



We also compare private school movers to those remaining in the public school based on student gender. As can be seen in the figures below, female voucher-eligible students are more likely to move to private schools when given the opportunity under the EdChoice program than are male voucher-eligible students, perhaps because female students tend to be more educationally successful than male students, especially those from low-income and minority families, in Ohio and throughout the United States and industrialized nations.[10]

Figure 7: Share female students: EdChoice participants versus eligible nonparticipants, first year of program participation

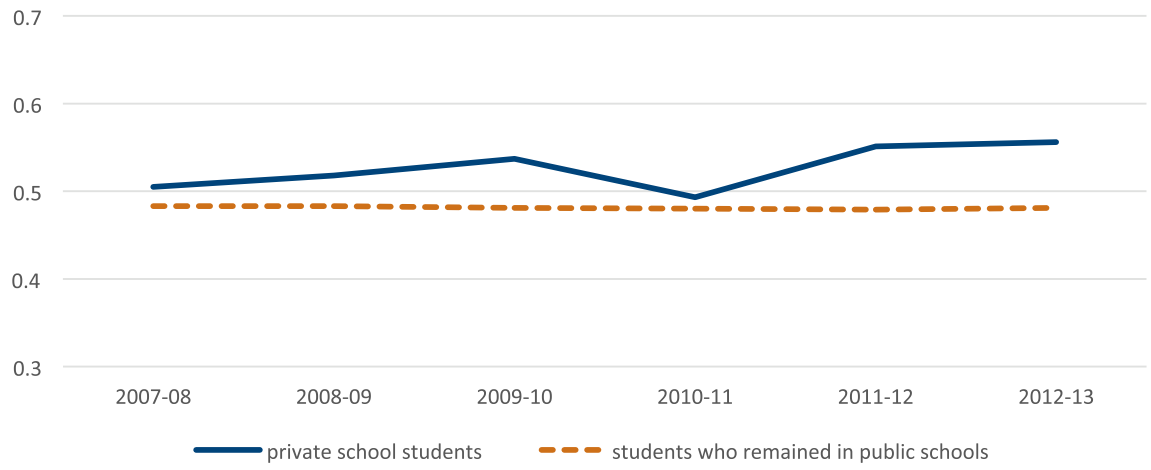
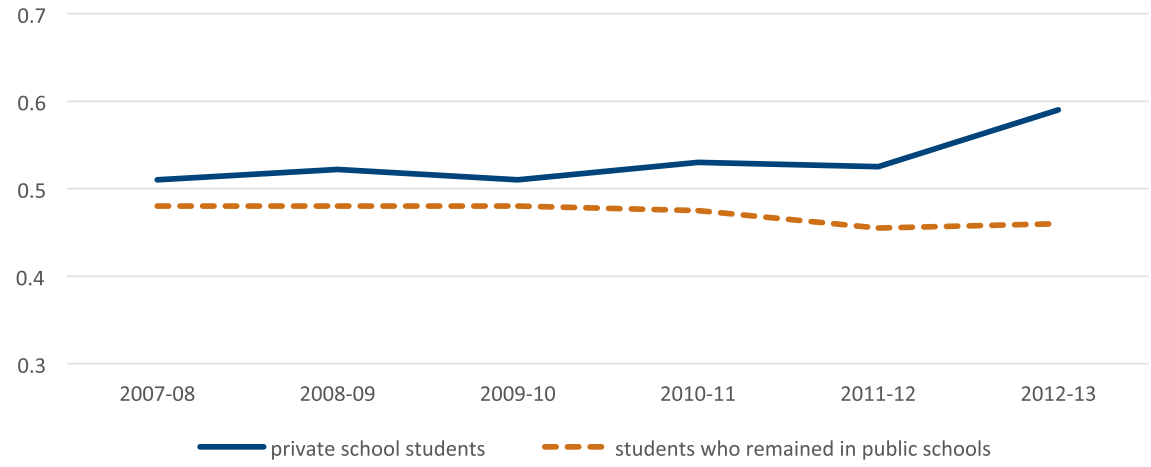


Figure 8: Share female students: EdChoice participants versus eligible nonparticipants, first year of program eligibility



Finally, we compare the racial and ethnic composition of voucher-eligible students who moved to private schools to those who remained in the public sector. The majority of EdChoice-eligible students are black, and we observe that the rate of private school attendance for black students is approximately proportionate to the black population. On the other hand, there is a difference between Hispanic and white student participation: we observe that private school movers are somewhat more likely to be Hispanic and somewhat less likely to be white than are eligible students who remain in the public schools, especially in the more recent years. Interestingly, this gap is driven primarily by students who have been eligible for multiple years rather than first-time eligible students, as there are no appreciable racial or ethnic differences in selection rates of first-time eligible students. These patterns suggest that Hispanic families may have required more time to act on their eligibility, perhaps because of language difficulties that hamper knowledge about eligibility or other aspects of the school-choice process or perhaps because of differences in school advising networks. These explanations, of course, are only speculative.

Figure 9: Racial/ethnic composition of EdChoice participants versus eligible nonparticipants, first year of program participation

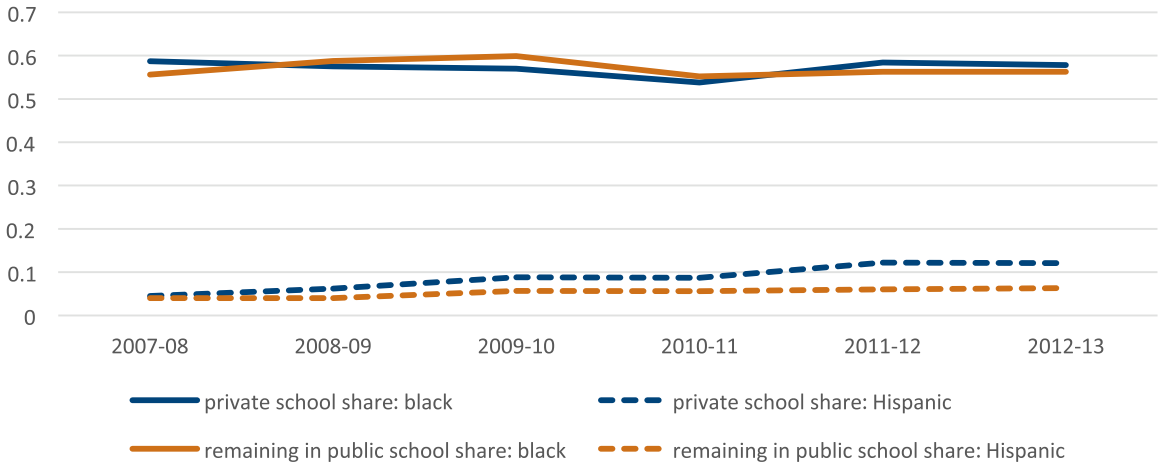
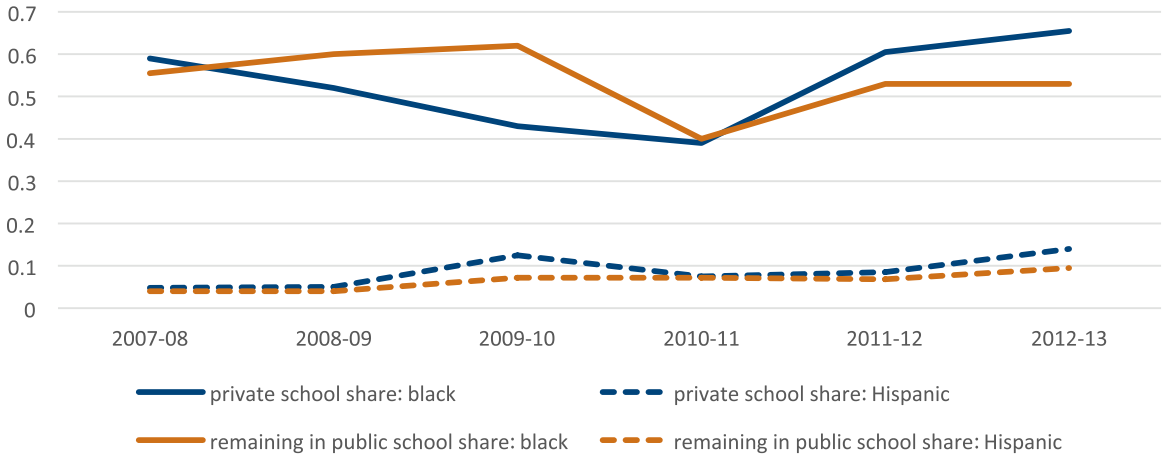


Figure 10: Racial/ethnic composition of EdChoice participants versus eligible nonparticipants, first year of program eligibility



In summary, it appears that comparatively high-achieving and comparatively well-off families (keeping in mind that the overwhelming majority of those participating in the program are still low income) are the groups more likely to use an EdChoice voucher to attend private school when offered the opportunity. There exist some racial and ethnic differences and some differences by gender, but these are relatively small in comparison to those seen regarding economic disadvantage and prior test performance. The fact that these gaps are considerably larger than those typically seen in other locations with voucher programs may suggest that there are specific features of the ways in which the EdChoice program is designed and implemented that make it more difficult for struggling students and comparatively disadvantaged families to make use of the voucher. Allowing schools to use their admissions standards almost certainly affects selection by achievement, both directly and indirectly through its attraction of comparatively motivated families.[11] Future work that investigates selection following the program's eligibility expansion in 2013–14 to include all economically disadvantaged students (beginning with Kindergarteners in 2013–14 and expanding one grade at a time in subsequent years) will help policymakers and analysts to understand the degree to which the means-tested channel for voucher receipt makes selection in Ohio more similar to that observed in Florida and other means-tested voucher locales.

4. Overall 'competitive effects' of EdChoice on public school students

Next, we turn to the question of whether the EdChoice program has affected student performance in reading and mathematics. We are interested both in the effects of program participation on the participants themselves (which we will study directly in section 5 of this report) as well as on the students who remain in the public schools.

There are numerous reasons to believe that the EdChoice program would affect performance of students in the public schools. One potential effect might come through changes in the **composition** of the student body in the affected public schools. We observed in section 3 of this report that the EdChoice program disproportionately attracted relatively high-performing students (albeit, as they are eligible for the vouchers, from a distinctly disadvantaged population) to the private schools, so the student body remaining in the public schools is somewhat lower achieving, on average, as a consequence of the program. If students benefit from higher-achieving peers, either directly or indirectly, this could lead to reduced performance in the public schools as a consequence of the EdChoice program. Of course, it's also possible that the direction of peer effects is different, at which point the composition effect of the EdChoice program would be different. Another potential effect of the EdChoice program could come through increased **competition** for students as a result of the voucher option. Other studies[12] have demonstrated that school vouchers can potentially lead to positive effects on public schools through this channel, so it is possible that Ohio public schools would also improve as a consequence of voucher competition. The introduction of school vouchers as an augmentation to school ratings could also increase the **salience** of the school ratings and induce performance improvements for schools that were previously rated poorly. Other studies[13] have found that introducing choice threats into school-accountability regimes has the potential to improve outcomes by more than the accountability systems alone, though these findings are not universal.[14] In summary, there are reasons to believe that the EdChoice program might either improve or reduce the performance of students attending affected public schools and that these effects depend on the nature and magnitudes of a variety of factors.

Two previous studies have investigated various aspects of the effects of the EdChoice program on traditional public schools. Greg Forster (*Promising Start: An Empirical Analysis of How EdChoice Vouchers Affect Ohio Public Schools*, August 2008) followed school-level cohorts of students from one grade in 2005–06 to the next in 2006–07 and found that for some grade transitions, schools where students were voucher eligible performed better than other schools.[15] Matthew Carr ("The Impact of Ohio's EdChoice on Traditional Public School Performance," *Cato Journal*, Spring/Summer 2011) carried out school-level analyses and found that schools where students became voucher eligible improved in terms of aggregate reading and mathematics scores following the introduction of the EdChoice program. In both the Forster and Carr studies, the authors provided evidence that their results were not due to "regression to the mean," a phenomenon that one might expect given that the schools subject to competitive pressure through the EdChoice program had low performance levels in the time prior to the introduction of the voucher program.

Though the existing studies of the effects of the EdChoice program on public schools provide much valuable information, there are opportunities to improve upon the existing research along two important dimensions. First of all, with the use of **student-level individual data**, it is possible for the first time to fully take into account the fact that school composition changes over time and, in fact, might be directly affected by the introduction of the voucher program. When we make use of student-level data, it is possible to follow an individual student's progress over time in a manner that is not possible with school-level data when researchers are forced to compare groups of students with different compositions. Second, and even more importantly, we make use of a **regression-discontinuity research design** that allows for much more of an apples-to-apples comparison than has been utilized in the extant literature on the EdChoice program. In essence, we will be able to compare schools that just barely became voucher eligible to schools that just barely missed becoming voucher eligible. Although this means that we are necessarily focusing our attention on a certain set of schools—those on the margin of becoming voucher eligible, rather than schools that were very far from the eligibility threshold—

the benefit of this research design is that we are able to study the effects of the EdChoice program using a comparison set of schools that are extremely similar to those that were directly affected by the program.

The biggest challenge to determining the effects of the EdChoice program on performance is that schools whose students become voucher eligible are systematically different from those whose students do not. This is, of course, by design: the Ohio Department of Education assigns school ratings based on average performance, and only schools with relatively poor performance are affected directly by the EdChoice program. As a consequence, merely comparing schools in which students become eligible for vouchers to those that did not—even if making a before-versus-after comparison—is unlikely to produce an apples-to-apples comparison.

Our solution, as mentioned above, is to implement a regression-discontinuity design. The benefit of a regression-discontinuity approach is that we can compare schools whose students just barely became voucher eligible to those schools whose students just barely missed voucher eligibility. The idea here is that these two sets of schools are going to be extremely close in terms of attributes, both observed and unobserved, so a comparison is more likely to be truly apples to apples. Our preferred regression-discontinuity analysis is for the **second** year of the program—students who would first become voucher eligible (or not) in the 2007–08 academic year. The rationale here is that in this second year of the program, the rules were such that it became much more difficult to predict exactly which schools' students would become voucher eligible and which schools' students would just barely miss the eligibility category, because the rule change made it so that schools under academic emergency or academic watch in at least two out the three years preceding the academic year in which the determination was made would become voucher eligible. Put differently, eligibility for vouchers in 2007–08 was based on a school's second-best performance between the three academic years of 2003–04, 2004–05, and 2005–06. Furthermore, because the program rules changed dramatically in the Fall 2006, making many more schools eligible for future rounds of vouchers, many schools that were "untreated" in 2006–07 essentially became partially "treated" by the threat of vouchers in the 2006–07 academic year. Therefore, we strongly believe that the second round of voucher eligibility, during which students could attend private schools for the first time in 2007–08, is by far the cleanest for causal-inference purposes.[16] (For completeness, we report some results from a considerably more flawed study of the first round of EdChoice program implementation in an appendix, but we strongly prefer the second-round implementation for the purposes of causal inference and encourage the reader to give the results based on the second round of implementation considerably more credence than those based on the first round of implementation.) Importantly, to the degree to which schools that just missed the threshold (on the positive side)—especially after the EdChoice policy was announced—might have been motivated to improve their performance, our estimates of competitive effects are underestimates of the true competitive effect of the EdChoice program.

A regression-discontinuity design requires that the researcher order all of the schools along some continuous "running variable"—the variable that determines whether or not an individual receives one treatment or another. Because the primary determinant of whether a school is designated as being under academic emergency or watch is that school's state-assigned performance index (PI)[17], our running variable is the second-best performance index received in the 2003–04, 2004–05, or 2005–06 academic year. This empirical approach only works for the set of schools not already eligible for vouchers in the initial 2006–07 academic year. Therefore, we exclude the initial ninety-nine schools eligible for vouchers in 2006–07 when carrying out this analysis.[18]

Our outcome of interest is either the student's test score in 2007–08, the first year that the students in question are eligible for a voucher, or in 2008–09, the second year of voucher eligibility. We are interested in both of these outcome years to see whether there is any change over time in the effects of the program between the first and second years of eligibility. Because the students attending different schools are fundamentally different, we measure our outcomes as the change in a student's standardized test scores (measured statewide with mean zero and standard deviation one) between 2005–06 and either 2007–08 (year one) or 2008–09 (year two). We identify 2005–06 as the baseline year of interest because this is the last year of outcomes prior to the introduction of the voucher program; 2006–07 test scores may plausibly be affected

by the introduction of the program. We assign students to schools for the purposes of voucher eligibility based on the schools they attended during the 2005–06 academic year, before the program was implemented. Therefore, for example, an estimated effect of 0.05 would mean that a student enrolled in 2005–06 in a school that would become voucher eligible in 2007–08[19] experienced 5 percent of one standard deviation better test scores, relative to their score in 2005–06, than would have been the case absent voucher eligibility. In our analysis, we consider 419,047 students with observed reading test-score growth between 2005–06 and 2007–08 and 418,749 students with observed mathematics test-score growth. When we add a second year of post-eligibility data (2008–09 scores), our analysis population necessarily declines considerably, to 300,270 (299,874) students in reading (mathematics). In most of the analyses that follow, we combine all voucher-eligible students together, regardless of whether they use the vouchers to attend private school or remain in the public school.[20]

We employ both graphical analyses and linear-regression analyses to present estimates of the EdChoice program effect on test scores. Because the results of regression-discontinuity analyses are often sensitive to the specific assumptions that a researcher employs, we present several variations on a theme for each type of analysis. This is especially important in cases like this EdChoice evaluation, because the running variable of the second-best PI is strong but does not perfectly predict voucher eligibility, as there are some idiosyncratic mechanisms through which schools with a second-best PI below eighty points can avoid voucher eligibility. In evaluation parlance, we refer to this type of situation as a "fuzzy" regression discontinuity. We deal with this fuzziness in a couple of different ways. In some analyses, we include all schools (except for the ninety-nine schools already eligible for vouchers in 2006–07) and assume that if the second-best performance index is below eighty points, the cutoff for academic-watch versus continuous-improvement status, the school becomes voucher eligible. In other analyses, we introduce a donut-hole approach in which we exclude entirely from the analysis all schools that have a second-best performance index either between 77 and 79.999 or between 75 and 79.999; we do this to reduce the number of schools that we call voucher eligible based on the second-best performance index when they are truly not voucher eligible due to exceptions. There exists a tradeoff in making this exclusion: the larger the donut hole, the more likely that schools just below the threshold are actually voucher eligible, but the larger the donut hole, the less likely the schools that are just below the donut hole are very similar to the schools that are just above the donut hole. We also lose 3.2 percent of all observations when we drop schools with second-best performance indices between 77 and 79.999, and we lose 5.1 percent of all observations when we drop performance indices between 75 and 79.999. Therefore, we present three different donut-hole variations to gauge the sensitivity of the results to this assumption.

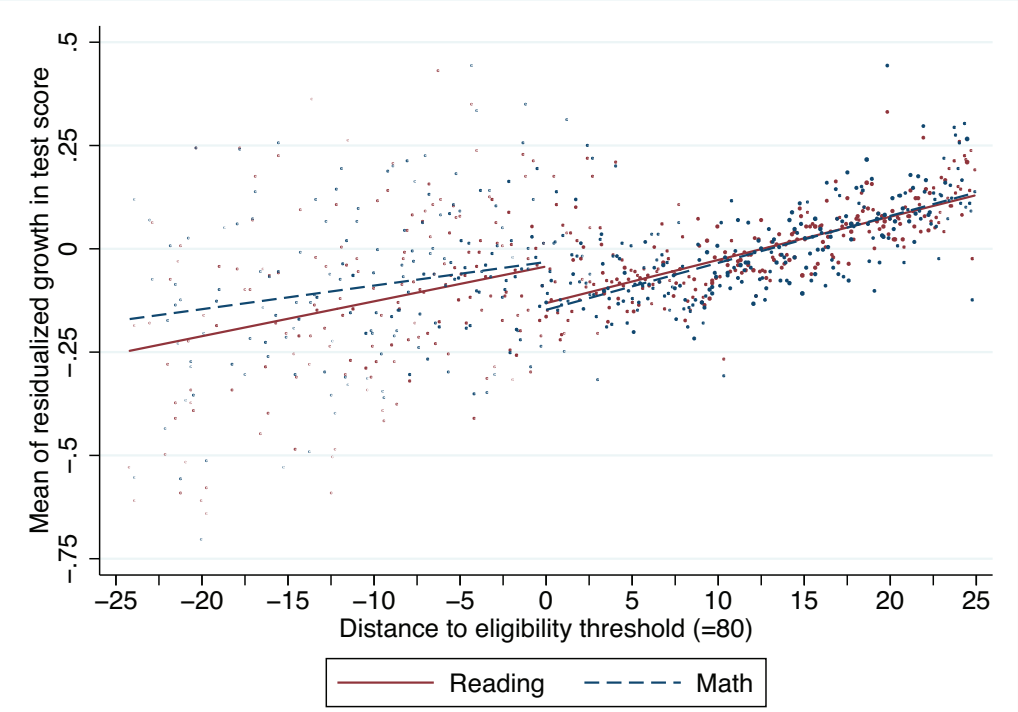
Second, for our regression analyses, we investigate the degree to which the estimated effects of EdChoice eligibility are affected by (1) controlling for a variety of student background characteristics (namely, sex, race and ethnicity, and economic-disadvantage status); (2) allowing the relationship between test-score outcomes and the second-best performance index to differ depending on whether the school is above or below the relevant threshold; or (3) both.

4.1. Graphical analysis

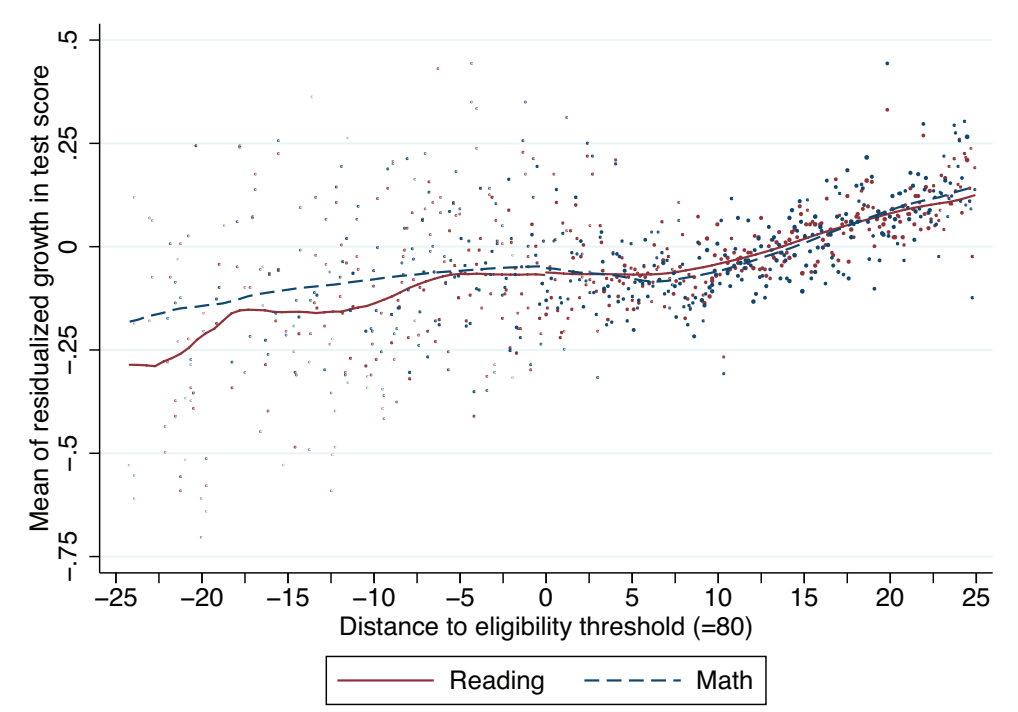
We begin with a set of graphical analyses of student-level test-score growth between 2005–06 and 2007–08. We present the graphical analyses in two ways: a linear analysis and a local-polynomial analysis. The linear analysis makes more use of the full range of schools, regardless of proximity to the threshold, while the local-polynomial analysis heavily weights the schools very close to the threshold. Each point in the graphs below is a separate value of a second-best performance index; if more than one school has exactly the same second-best performance index, we average those schools together in the graphs for ease of presentation.[21] The red lines represent reading scores, while the dashed blue lines represent mathematics scores. The estimated effect of EdChoice eligibility is the difference between the line to the left of the zero threshold and the line to the right of the zero threshold. As can be seen in the figures below, whether there appears to be a positive benefit

of EdChoice or a zero benefit depends on the assumption about whether the linear-fit analysis or the local-polynomial analysis is more appropriate.

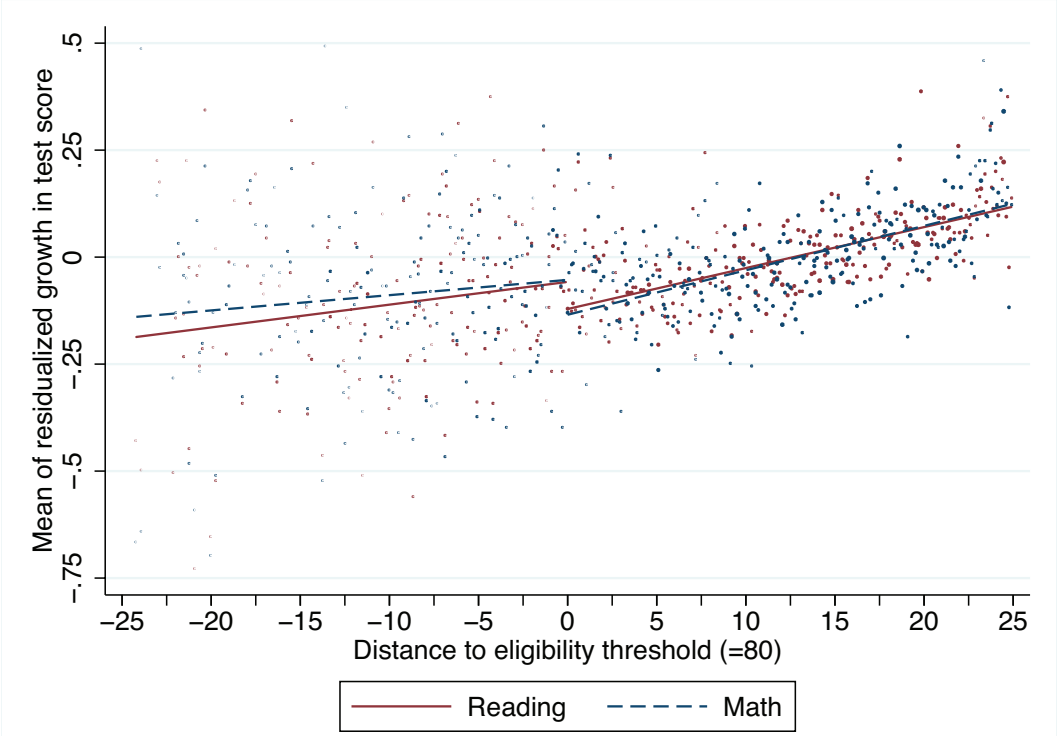
In the case of the linear fit, there is an apparent improvement in test-score growth for voucher-eligible students:

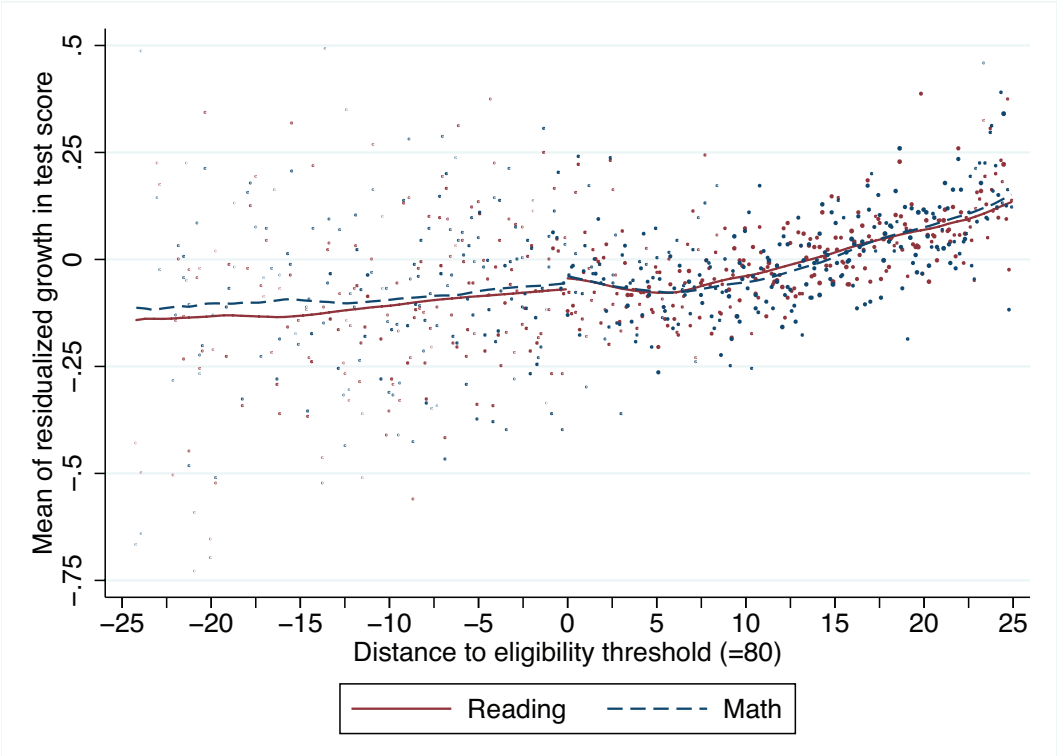


On the other hand, in the case of the local-polynomial analysis, there is no apparent jump at the zero threshold point:

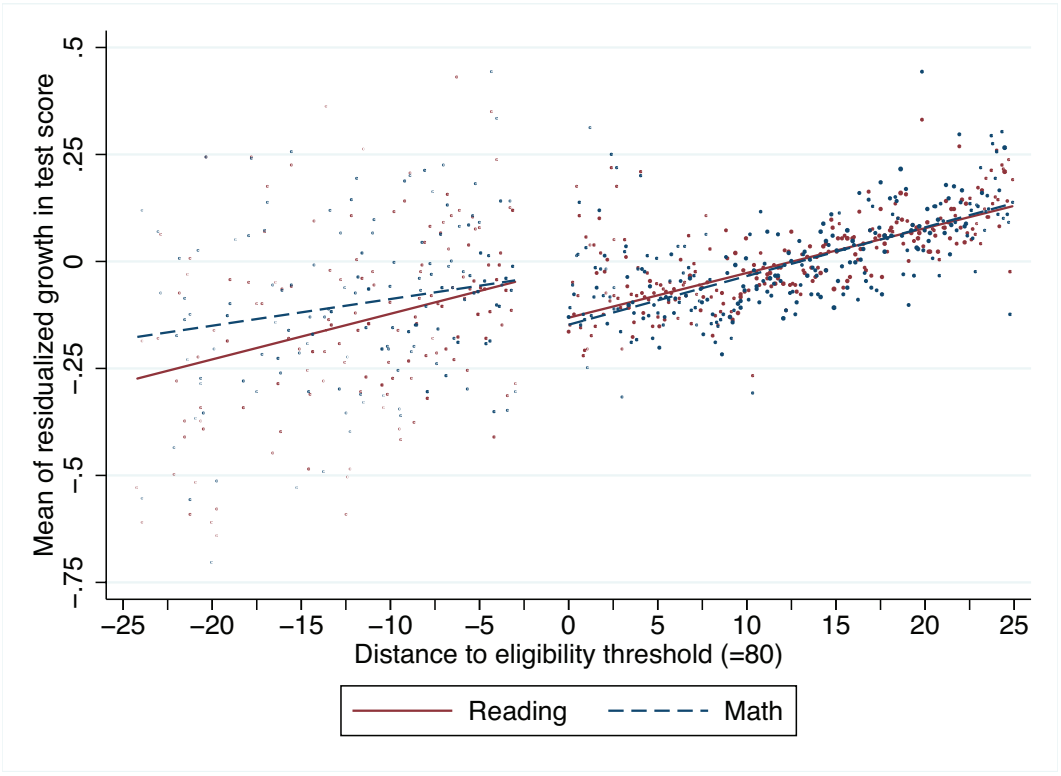


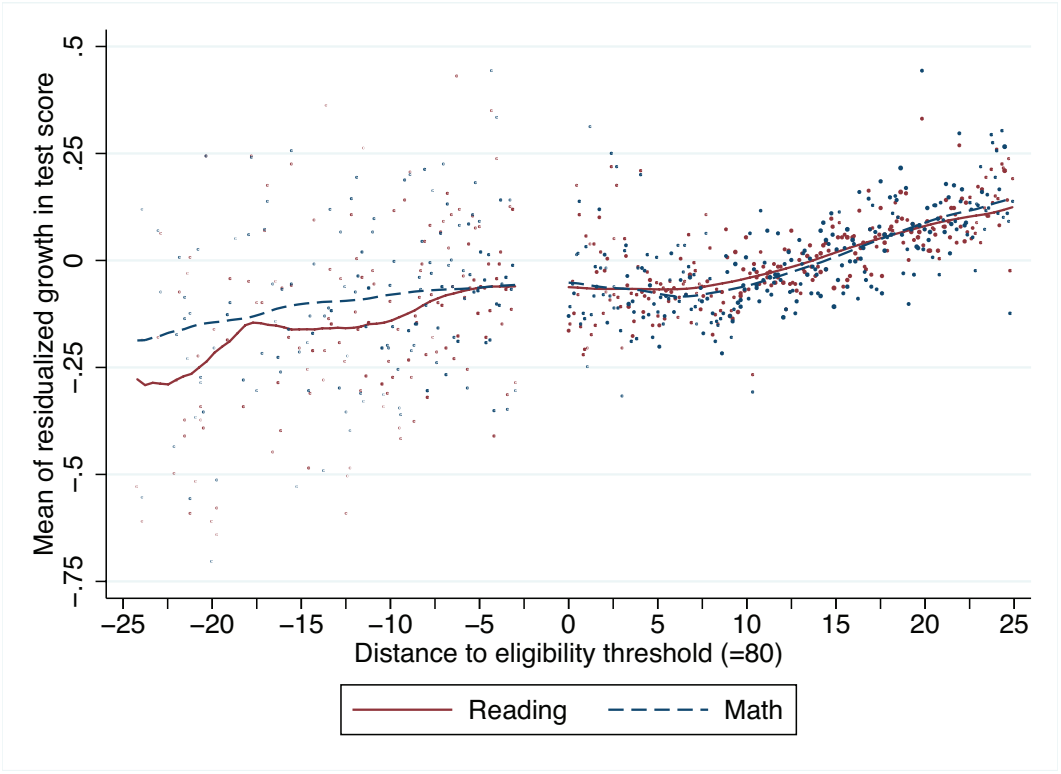
The same pattern of findings is apparent if we investigate test-score growth between 2005–06 and 2008–09. There is a clear positive estimate of EdChoice eligibility in the linear-fit case but a zero estimate in the local-polynomial case:



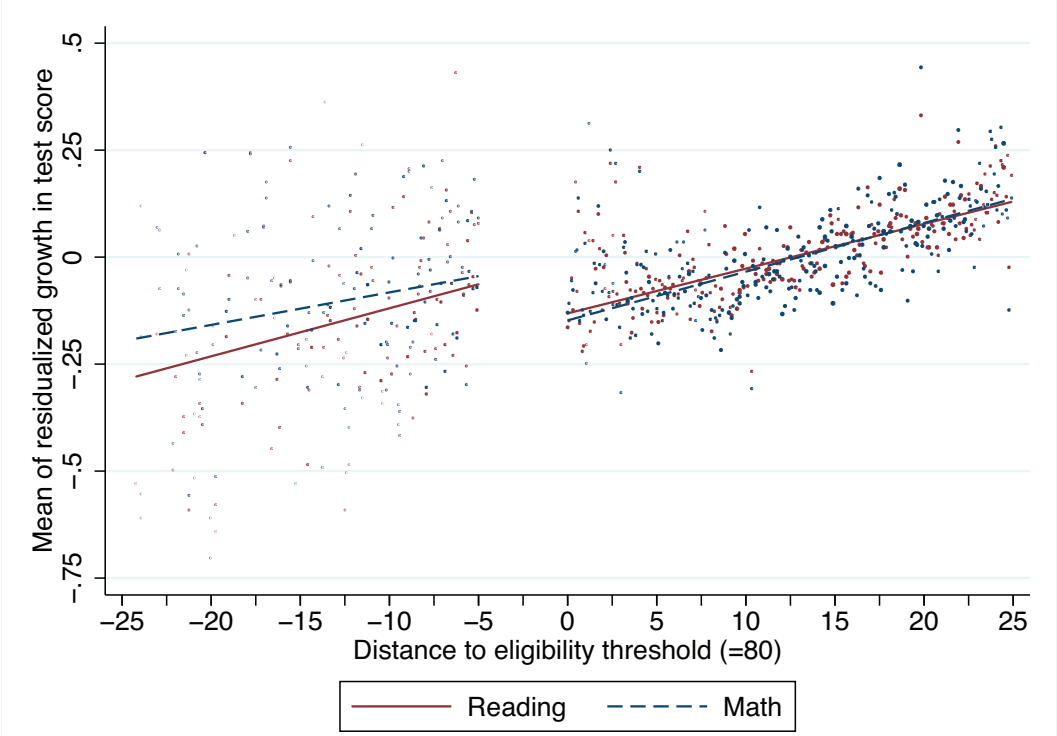


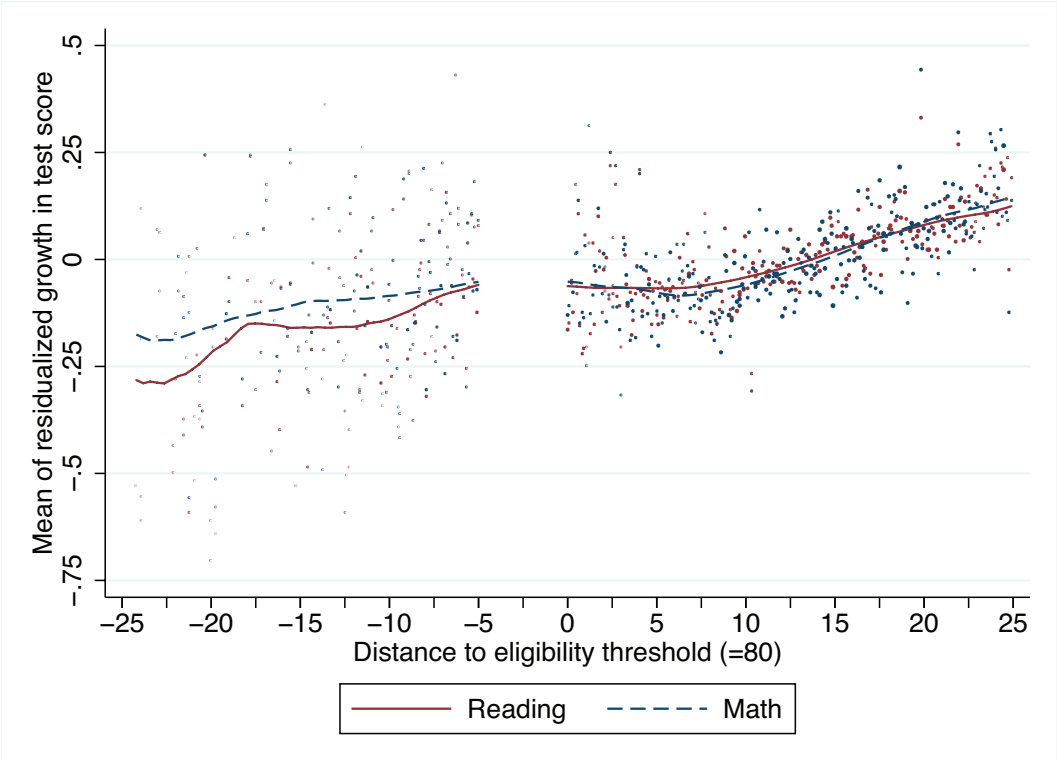
The same patterns—positive estimates for the linear-fit analysis and zero estimates for the local-polynomial analysis—are apparent regardless of the donut hole chosen or regardless of whether we look at growth between 2005–06 and 2007–08 or between 2005–06 and 2008–09. For instance, the following are the figures seen for the donut hole dropping performance indices 77 to 80, for the case of growth between 2005–06 and 2007–08:



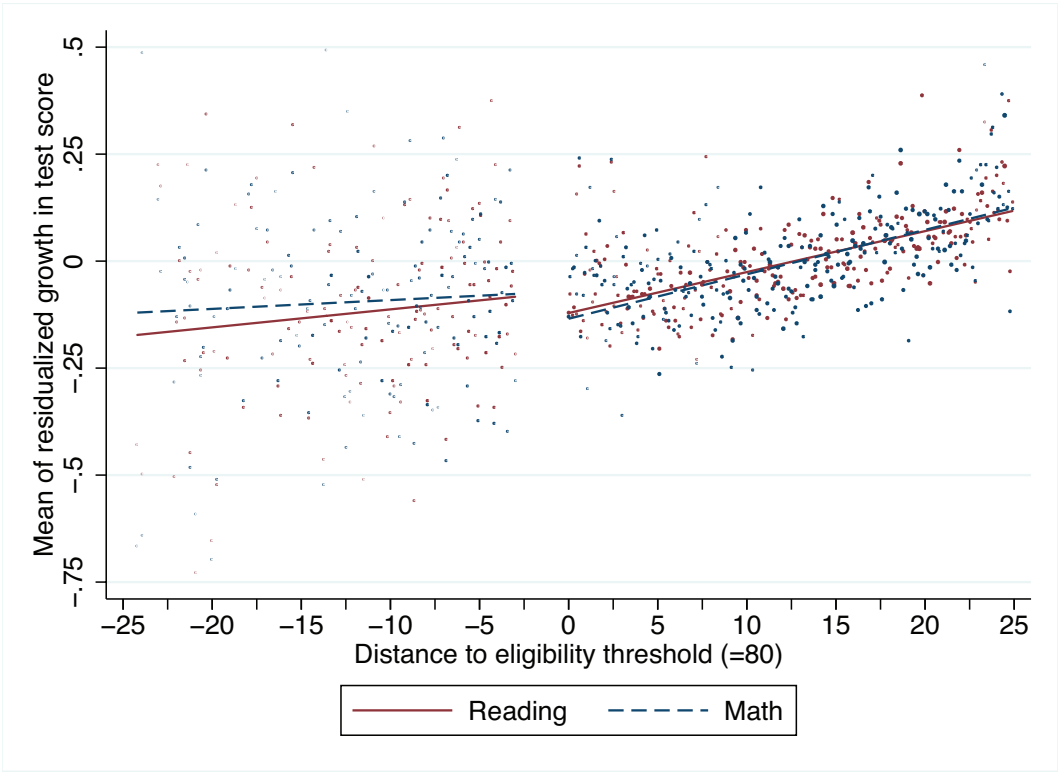


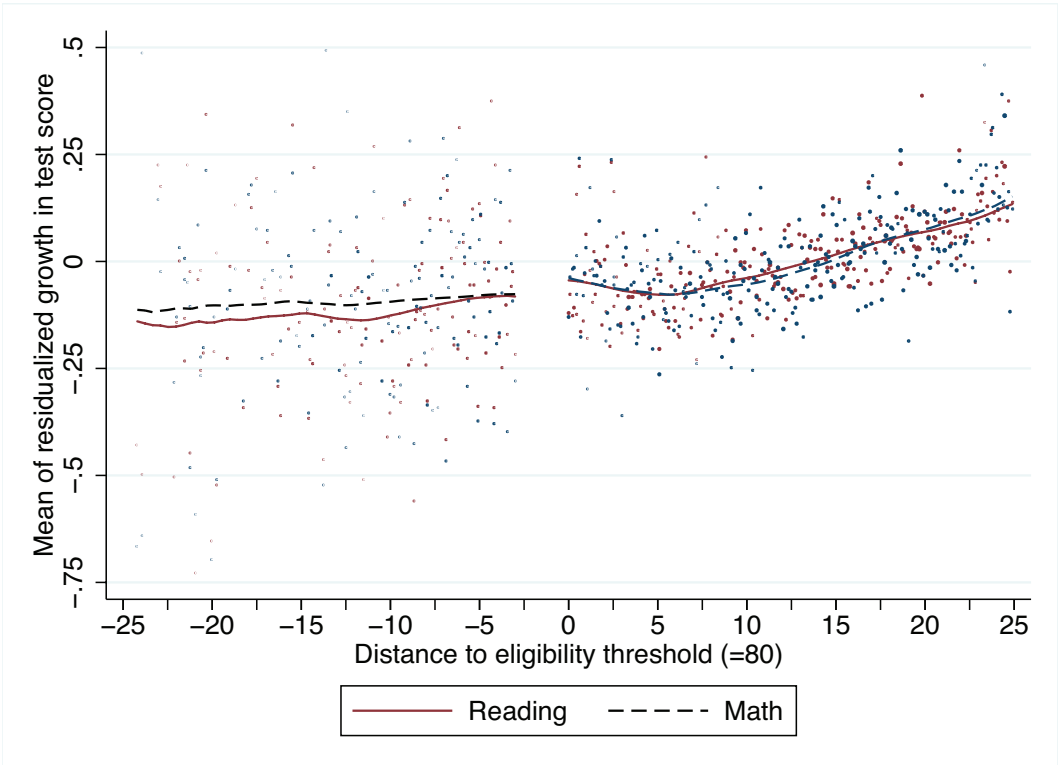
The following are the figures seen for the donut hole dropping performance indices 75 to 80, for the case of growth between 2005–06 and 2007–08:



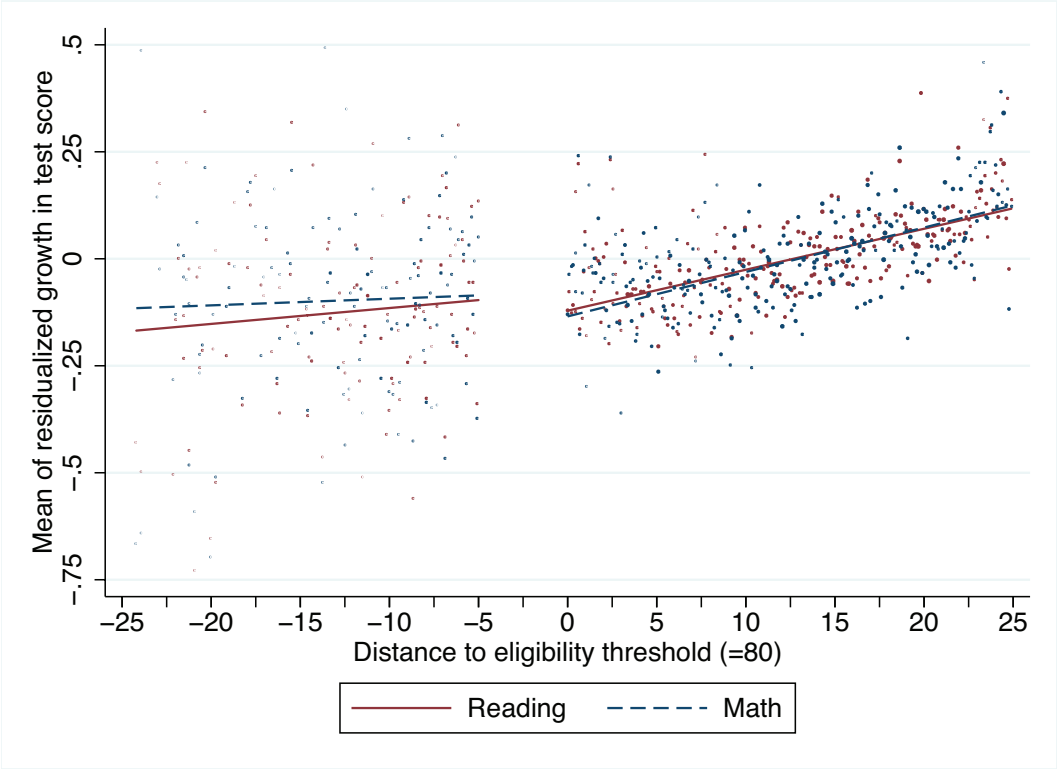


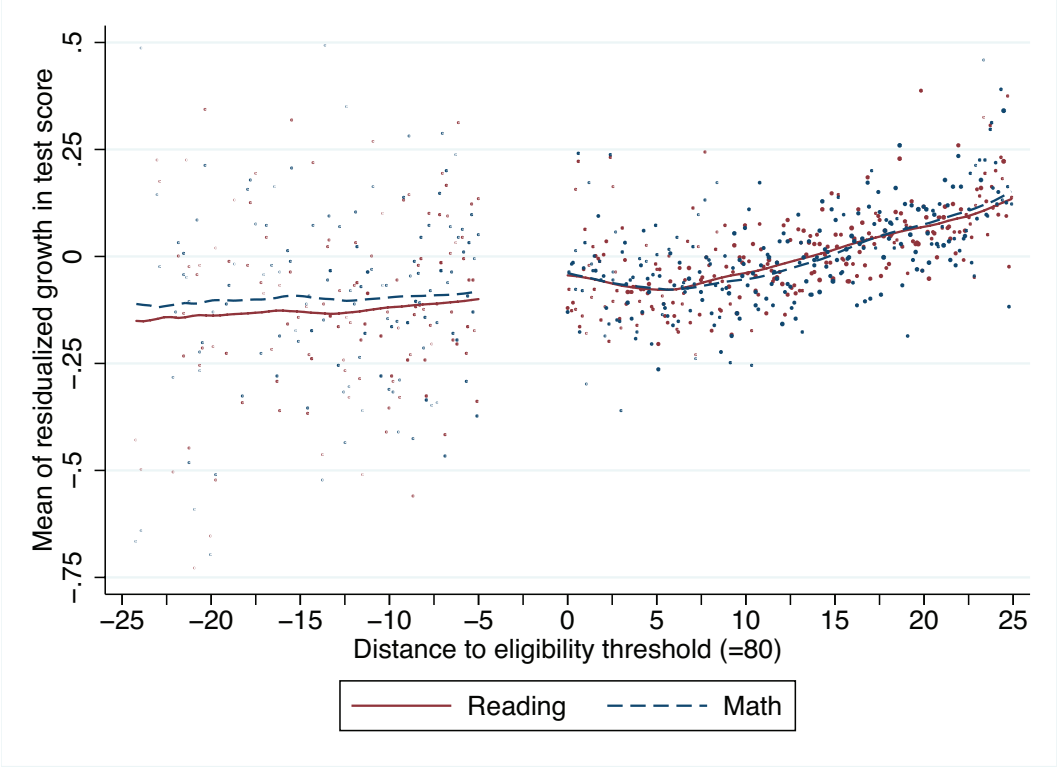
Likewise, the following are the figures seen for the donut hole dropping performance indices 77 to 80, for the case of growth between 2005–06 and 2008–09:





The following are the figures seen for the donut hole dropping performance indices 75 to 80, for the case of growth between 2005–06 and 2008–09:





In summary, the graphical analyses indicate that the effects of EdChoice on student performance for voucher-eligible students are not negative but are either zero or positive depending on the specific assumptions made regarding the fit of the empirical model. The empirical models that make use of a wider range of data to estimate the underlying relationship between the performance index and student performance tend to show positive results, while those that emphasize the data points extremely close to the threshold for voucher eligibility are more likely to suggest a result closer to zero (but not negative). The following subsection provides additional evidence that might help to serve as a "tiebreaker" between the solidly positive estimated effects of the EdChoice program and those that are closer to zero in magnitude.

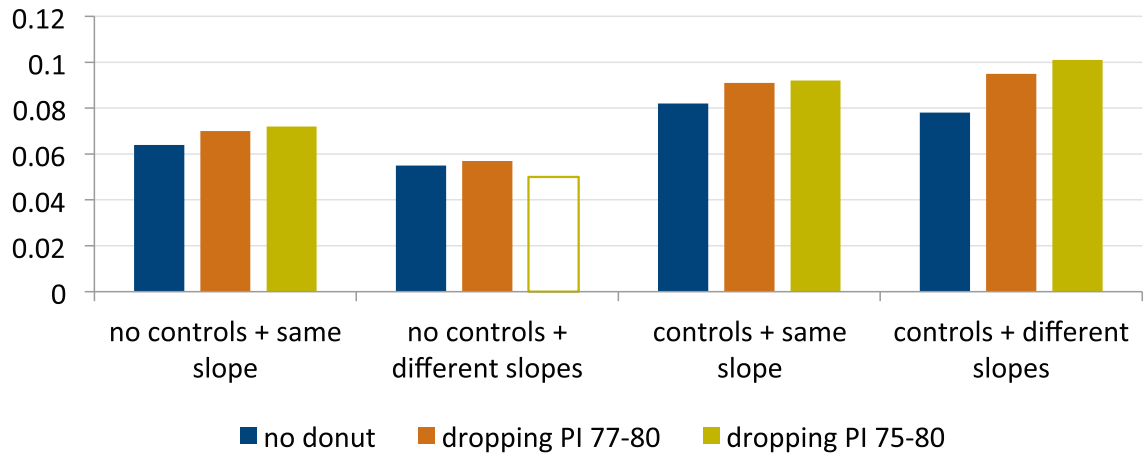
4.2. Linear-regression analysis

We next turn to linear-regression analyses that are analogous to the graphical analyses presented above. As before, we concentrate on the schools that became eligible in the second round of EdChoice eligibility—that is, with students able to attend private schools for the first time in 2007–08 and excluding the ninety-nine schools whose students became eligible in the 2006–07 round that is less well suited to this analysis.

In the discussion that follows, we consider a number of variations to discern the degree to which the results are sensitive to particular assumptions. We begin by presenting the estimated effects of EdChoice eligibility on students' reading and mathematics growth between 2005–06 and 2007–08 for twelve different specifications.[22] Specifically, for each outcome, we conduct four analyses apiece with no donut hole around the eligibility threshold, dropping schools with the second-best PI between 77 and 80 and dropping schools with the second-best PI between 75 and 80. For each of these cases, we present analyses where (1) we include no control variables and force the relationship between PI and test-score growth to be the same on both sides of the threshold (same slope); (2) we include no controls but allow the relationship to be different above versus below the threshold (different slopes); (3) we include controls for grade in 2005–06, sex, race and ethnicity, and economic disadvantage but impose the same-slope assumption; and (4) we include controls and also allow the different-slopes assumption. All regression models include a control for the running variable—the difference between the second-best PI and the eligibility threshold of 80. Also, because the treatment is a school-level treatment, we adjust the standard errors for clustering at the level of the school the student attended in 2005–06; doing so leads to larger standard errors and a steeper but more appropriate test for discerning the degree of statistical significance than would occur without adjusting the standard errors. In the text of this report, we present the estimated effects of EdChoice eligibility on test-score growth in graphical form; appendix table 1 presents the same results in a statistical tabular form.

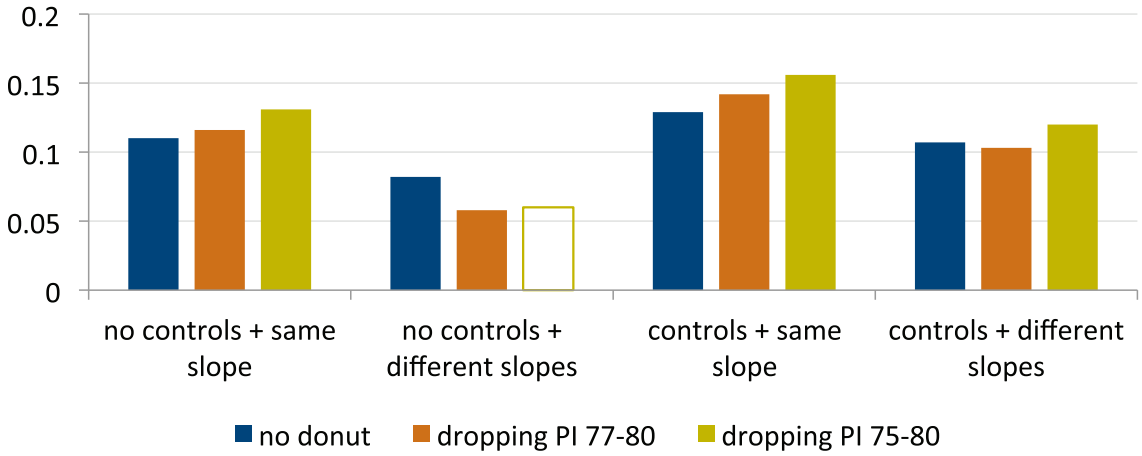
The following figure shows that the estimated effect of EdChoice eligibility on reading growth is invariably positive and statistically distinct from zero at conventional levels in eleven of twelve specifications:[23]

Figure 11: Estimated effects of EdChoice eligibility on reading growth from 2005–06 to 2007–08



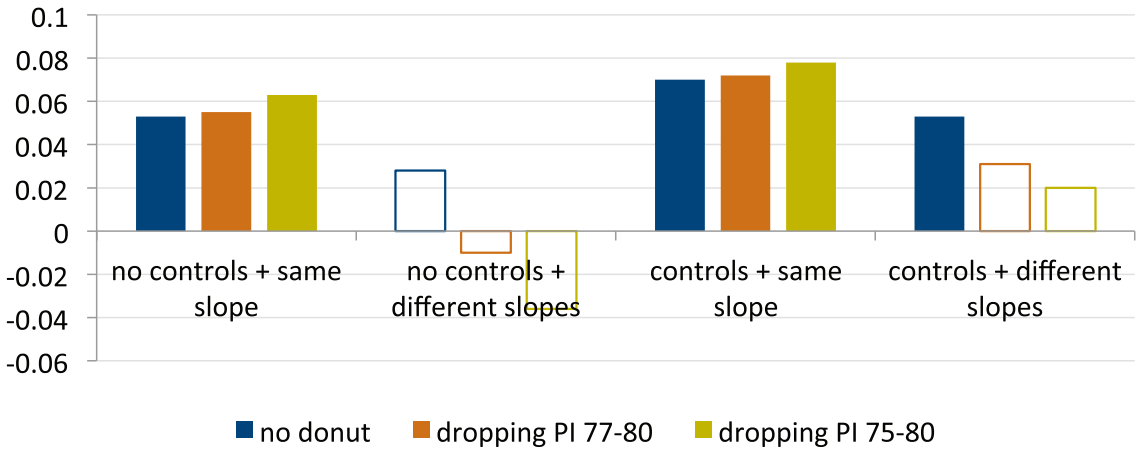
A similar pattern is apparent with regard to mathematics:

Figure 12: Estimated effects of EdChoice eligibility on mathematics growth from 2005–06 to 2007–08



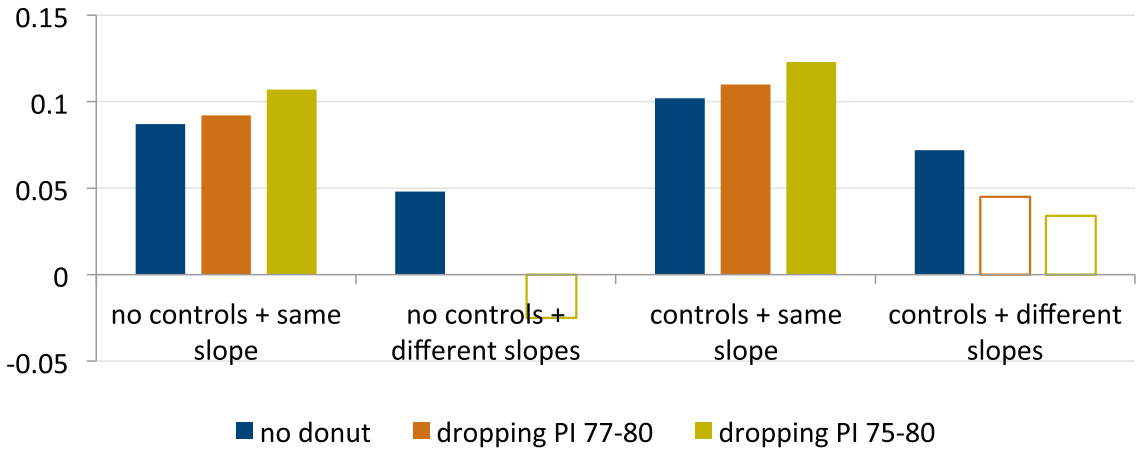
The results are not as universally strong when we consider reading growth between 2005–06 and 2008–09 but are still disproportionately positive and statistically distinct from zero, though more dependent on model specification. The assumption of same versus different slopes of the relationship between the PI and test-score growth is more consequential in this case than with growth through 2007–08:

Figure 13: Estimated effects of EdChoice eligibility on reading growth from 2005–06 to 2008–09



As before, a similar pattern emerges with mathematics, as well:

Figure 14: Estimated effects of EdChoice eligibility on mathematics growth from 2005–06 to 2008–09



The overwhelming majority of voucher-eligible students remain in the public schools after eligibility (this is especially the case because these analyses require that a student have a test score observed in 2005–06), but the analyses presented above include both eligible students remaining in public schools and those going to private schools on an EdChoice voucher. Therefore, we repeat the above analyses, restricting our attention only to those eligible students remaining in public schools:

Figure 15: Estimated effects of EdChoice eligibility on reading growth from 2005–06 to 2007–08 (public only)

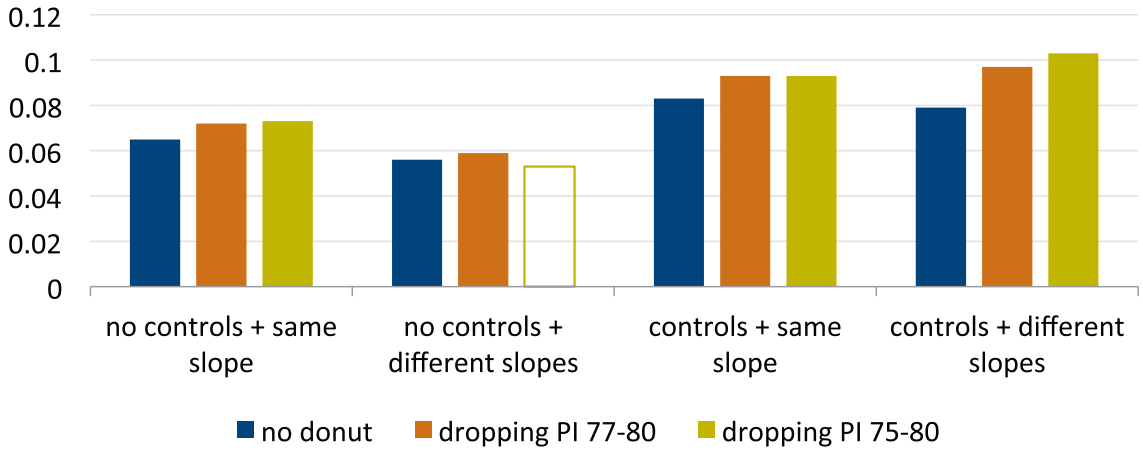


Figure 16: Estimated effects of EdChoice eligibility on mathematics growth from 2005–06 to 2007–08 (public only)

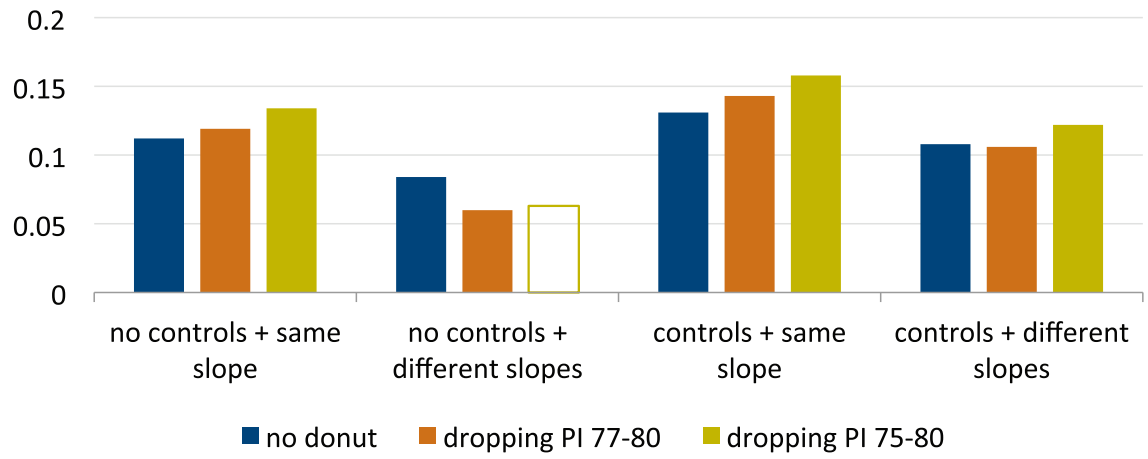


Figure 17: Estimated effects of EdChoice eligibility on reading growth from 2005–06 to 2008–09 (public only)

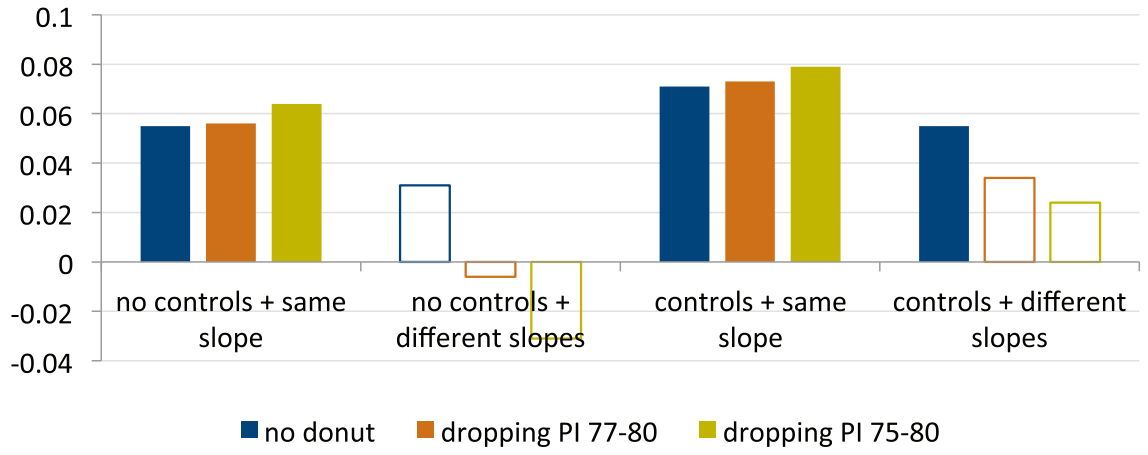
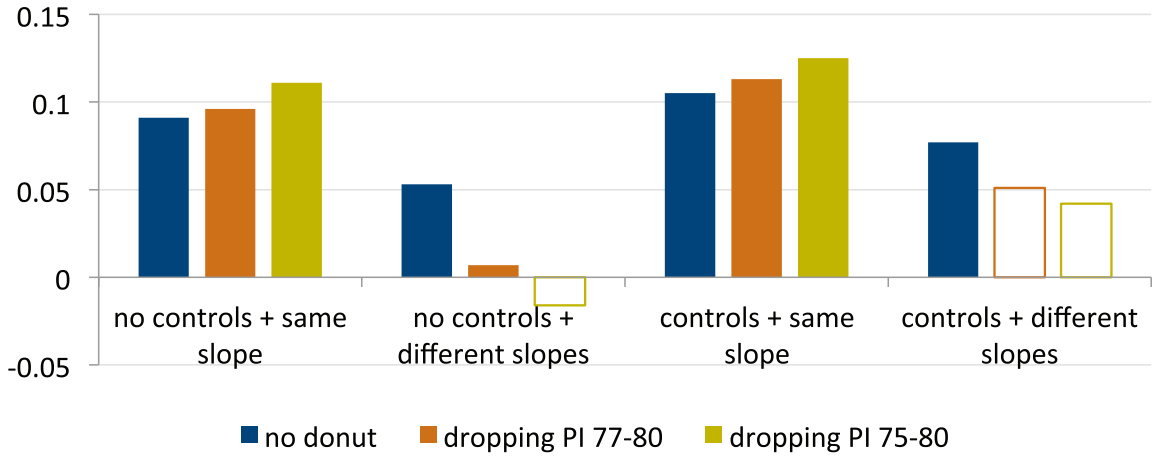


Figure 18: Estimated effects of EdChoice eligibility on mathematics growth from 2005–06 to 2008–09 (public only)

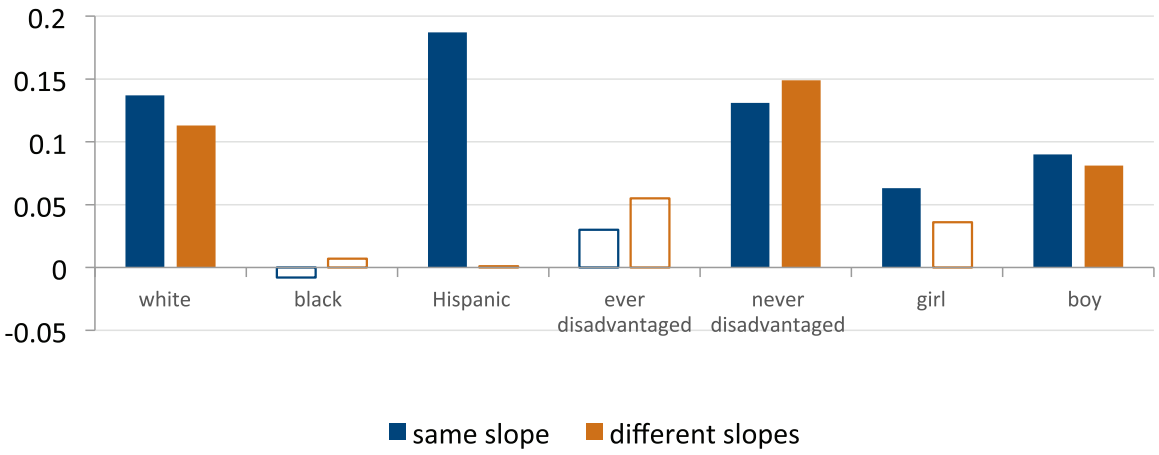


As is apparent from these graphs, the results are extraordinarily similar to those presented for all eligible students. These results are presented in tabular form as well in appendix table 2.

Next, we consider whether the performance increases associated with EdChoice eligibility are similar across a variety of groups. Because the above-mentioned analysis makes clear that the results are quite similar regardless of whether we include a donut hole (or regardless of the size of the donut hole considered), as well as whether or not we control for student background variables, for ease of explication from this point onward all specifications that we estimate include (1) no donut hole and (2) controls for grade in 2005–06, sex, race and ethnicity, and economic disadvantage. Because the results often appreciably vary depending on whether or not we allow the underlying relationship between the PI and test-score growth to vary depending on which side of the threshold a school is located, we continue to report two sets of findings: those where the slope of the relationship between the PI and test-score growth is constrained to be the same and those where the slopes are allowed to vary.[24] We consider seven different subgroups: three racial and ethnic groups (white, black, and Hispanic), ever disadvantaged versus never disadvantaged, and girls versus boys. In the body of the report, we limit the discussion to growth between 2005–06 and 2007–08, but in appendix table 3 we present the statistical tabular version of not only the graphs discussed below but also for the specifications based on growth between 2005–06 and 2008–09.

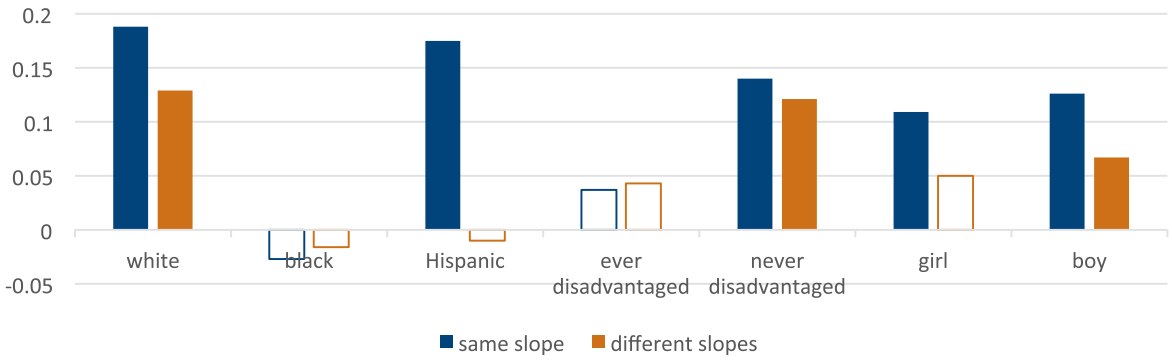
Although we never observe a statistically significant negative estimated effect of EdChoice eligibility for any subgroup, it appears that the positive benefits are disproportionately concentrated among white students and relatively advantaged students and are modestly stronger for boys than for girls. In addition, in some specifications, there is a large positive estimated effect for Hispanic students, but this effect is more sensitive to model specification. The results for reading growth are as follows:

Figure 19: Estimated effects of EdChoice eligibility on reading growth from 2005–06 to 2007–08, by subgroup



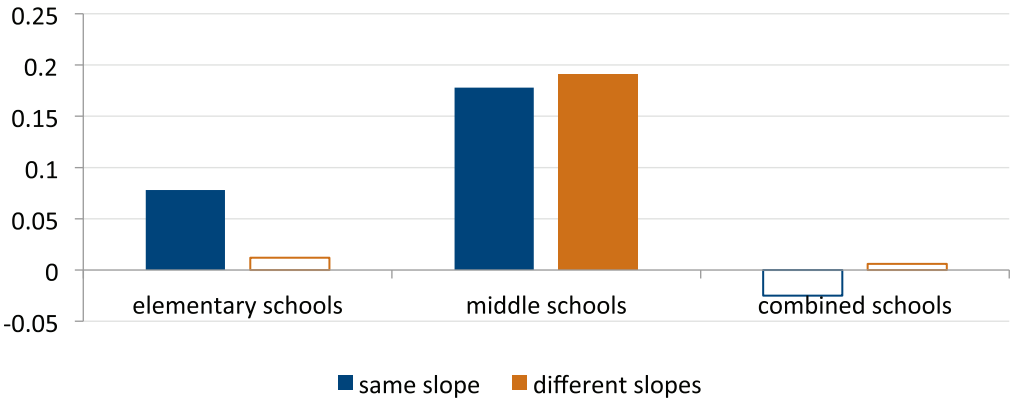
The estimated effects of EdChoice eligibility look very similar if instead we focus on mathematics growth:

Figure 20: Estimated effects of EdChoice eligibility on mathematics growth from 2005–06 to 2007–08, by subgroup



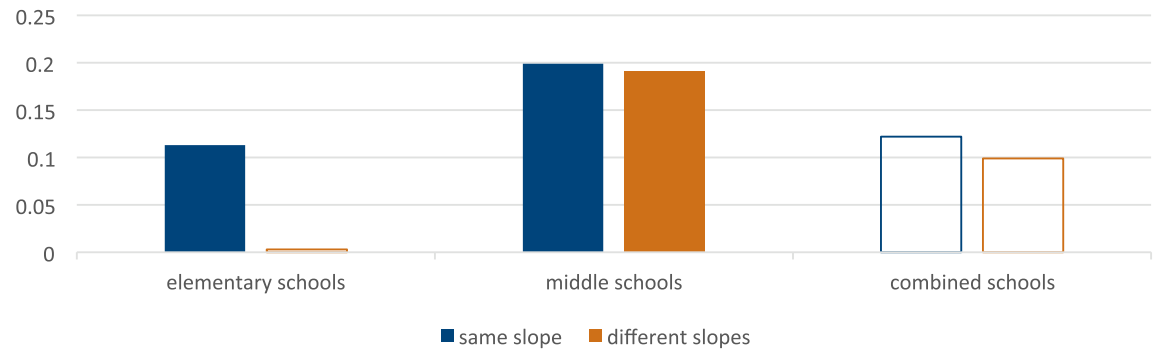
We also stratify the schools by elementary only, middle only, and combined elementary-middle schools.[25] The results, also reported in tabular form in appendix table 4, indicate that the positive effects of EdChoice eligibility are particularly concentrated among middle schools:

Figure 21: Estimated effects of EdChoice eligibility on reading growth from 2005–06 to 2007–08, by school type



The pattern of results appears similar for mathematics and for reading, though the mathematics results are much more pronounced (though still statistically indistinguishable from zero at conventional levels, given the small number of relevant schools) for combined schools:

Figure 22: Estimated effects of EdChoice eligibility on mathematics growth from 2005–06 to 2007–08, by school type



5. Effects of EdChoice participation on private school attendees

We have found that, on average, EdChoice eligibility appears to have benefited students (though, as mentioned in the previous section, whether we find a positive or zero estimated effect depends to some degree on the assumptions we make). However, although the empirical approach implemented in section 4 can help to identify the effects of EdChoice **eligibility**, it cannot help to identify the effects of EdChoice **participation**.

The reason that the regression-discontinuity design is inappropriate to study the effect of program participation is that the random element it exploits (some schools are just barely on one side or the other of the threshold for reasons that cannot be predicted or manipulated) is not relevant for those who actually participate in private schooling as a consequence of the program. Indeed, we've observed in section 3 that EdChoice-eligible students who participate in the program are different along a number of observable dimensions versus EdChoice-eligible students who do not participate in the program, and there's every reason to believe that there would be differences along unobservable dimensions as well, thanks to the factors associated with being motivated to change schools, obtaining admission to private school, and securing a voucher.

The ideal way to identify the effects of EdChoice program participation would be to make use of the random variation associated with a lottery, as a number of voucher evaluations in other locations have been able to do. However, EdChoice vouchers are not allocated via lottery, so this precludes this possibility in the Ohio context. The best that we can do in the EdChoice setting is to attempt to match EdChoice participants as closely as possible to nonparticipants in order to follow two observationally similar groups of students through their different settings for the sake of comparison.[26]

Specifically, we carry out a **propensity score matching** (PSM) approach in order to approximate as closely as possible the apples-to-apples comparison that would have been possible with random assignment.[27] The idea behind this approach is to find the nonparticipating students who are most similar along observable dimensions to the program participants. In many applications of PSM models, researchers are forced to identify these control students from the set of people who were eligible but for some reason did not participate. Such a practice can be troublesome because we have no way of knowing whether the people who chose to participate (and were admitted to a private school) and the selected people who did not choose to participate are really the same along unobservable lines. Fortunately, in the case of the EdChoice program, we are able to overcome the problems associated with identifying control students from the population of eligible individuals (as is common with PSM approaches) for the same basic reason that adds credibility to the analysis in section 4: because some schools just barely became voucher eligible and other schools barely missed becoming voucher eligible, we can identify potential comparison students who are nearly identical along observable dimensions to program participants but were **ineligible** to participate because of the school they attended before the EdChoice program was announced. Additionally, because we can identify the nearly identical (on observed performance) schools, we can argue that program participants and the comparison students attended schools that were reasonably similar. In sum, while the PSM approach that we employ is not without its problems,[28] we are convinced that it is as close as we can get to approximating random assignment given the EdChoice setting.

Our approach to PSM is to first isolate the schools whose second-best PI was just slightly above the threshold for voucher eligibility. We begin by limiting the comparison schools to those within three points of the eligibility threshold, and then we consider an even more restrictive case in which we limit comparison schools to those within one point of the eligibility threshold. We also consider cases in which we consider only the **treatment** students (that is, those using the EdChoice vouchers) who attended schools within three points of the eligibility threshold.[29] The idea here is to find treatment and comparison students who not only look very similar on observable factors but also attended public schools in 2006–07 that were extremely similar, except that the treatment students attended public schools that were voucher eligible and the comparison students attended public schools that were not. We match students by finding, for each EdChoice-voucher user, the comparison

student or students with the closest possible combination of prior-year reading and mathematics scores, student sex, student race and ethnicity, grade in 2006–07, and student history of economic disadvantage. There are several different ways to statistically identify the closest matches. In the tables and figures, we report the results when we match students using logit matching and when we match students using probit matching. Because we need to observe prior test scores in order to carry out a credible match, we can only study a maximum of 445 students who first moved to private schools in 2007–08.[30] Because we are focusing on such a narrow band of schools for the purposes of our comparison, the number of observations is necessarily much lower than the total set of EdChoice-participating students.

Although this implementation of the PSM estimator improves the scientific credibility of the estimates, it does so at significant cost to generalizability. Specifically, we can only identify with relative confidence the estimated effects of EdChoice-scholarship participation for those students who had been attending the highest-performing EdChoice-eligible public schools and not those who had been attending lower-performing public schools. It may be the case that those attending the poorest-performing public schools would have had very different performance effects than those attending the relatively well-performing public schools. Therefore, these comparisons may not be generalizable to the full set of public schools. It is certainly possible that children coming from worse-performing public schools to the private schools under the EdChoice program might have had considerably better outcomes; that said, we have no way of credibly investigating this possibility using existing data.

We begin by comparing reading and mathematics scores under a number of different PSM modeling specifications and study not just the first year of a student's attendance in private school (or comparison public school following eligibility), 2007–08, but also the next two years, 2008–09 and 2009–10, as well. The reason we follow students for multiple years is because we want to reduce the likelihood that any estimated effects of private school participation are affected by any short-term changes in test scores associated with school switching. We report both repeated cross-sectional results where the set of students in each year differs depending on how many students remain in tested grades as well as panel results in which we follow the same students (including 205 treated students attending private schools on EdChoice vouchers). We report statistical tabular results in appendix table 5.

Figure 23: PSM estimates of effects of EdChoice participation on mathematics test scores, 2007–08 cohort, zero-to-three-point comparison (relative to 2006–07)

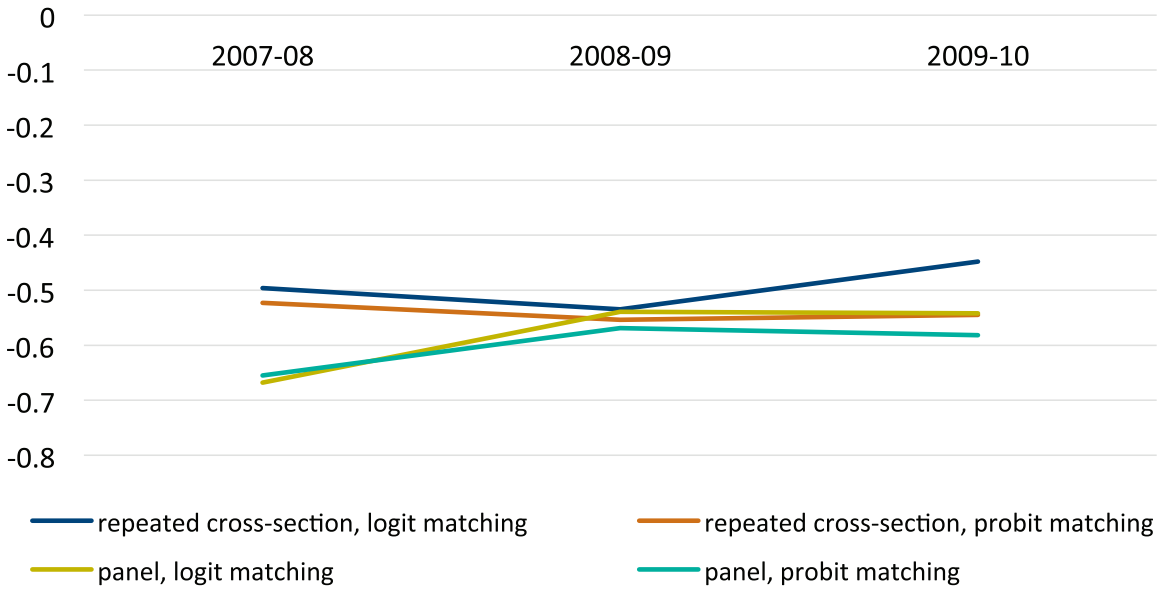
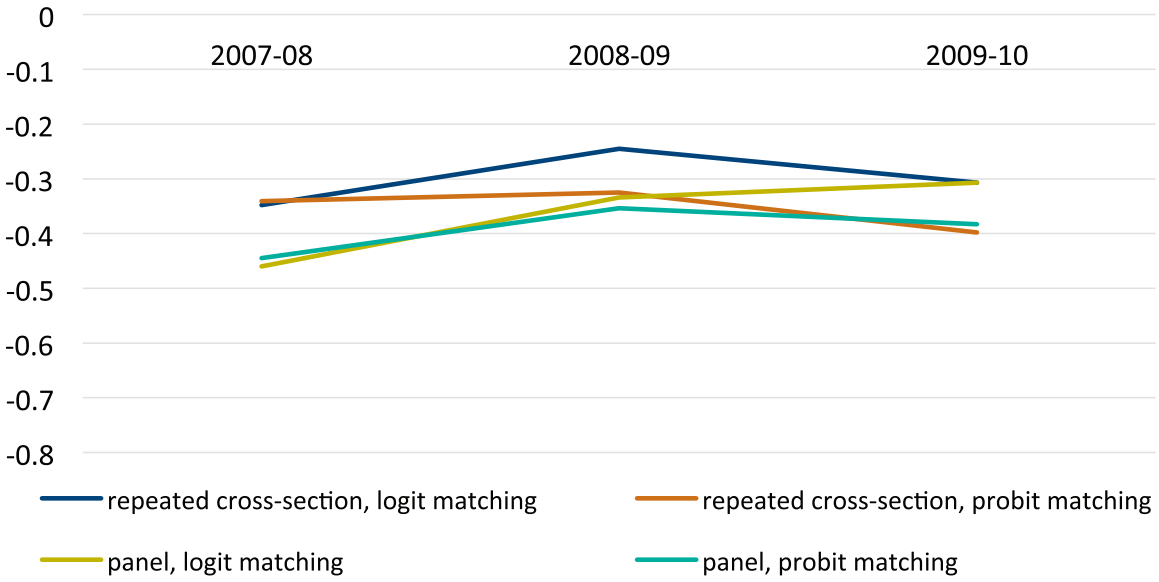


Figure 24: PSM estimates of effects of EdChoice participation on reading scores, 2007–08 cohort, zero-to-three-point comparison (relative to 2006–07)



As is apparent from the above graphs, the estimated effects of EdChoice participation on test scores are unambiguously negative across a variety of model specifications, for both reading and mathematics (though more negative for mathematics than for reading). The negative results are present regardless of whether we look at the same students in a panel setting or different sets of students, and they do not appear to change much over time, indicating that the initial negative results are not due to the fact that EdChoice participants all were newcomers in a new private school.

As seen in the following figures and in appendix table 6, the same patterns are clear if we restrict our comparison group to students attending schools within one point of the EdChoice eligibility threshold, as well as when limiting the EdChoice treatment group to schools within three points of the eligibility threshold (though this limits the analysis to just eighty-two treated students in year one and fifty-one treated students by year three).[31] In this last comparison in reading, the results become small and statistically insignificant in year three—but the overwhelming evidence indicates a substantial negative effect on test scores of attending private schools under an EdChoice voucher for those students who were attending the highest-performing schools amongst those that were eligible for the voucher. We cannot generalize these findings to students who had previously attended much lower-performing public schools because we cannot conceive of a credible way to make that type of comparison.

Figure 25: PSM estimates of effects of EdChoice participation on mathematics scores, 2007–08 cohort, different models (relative to 2006–07)

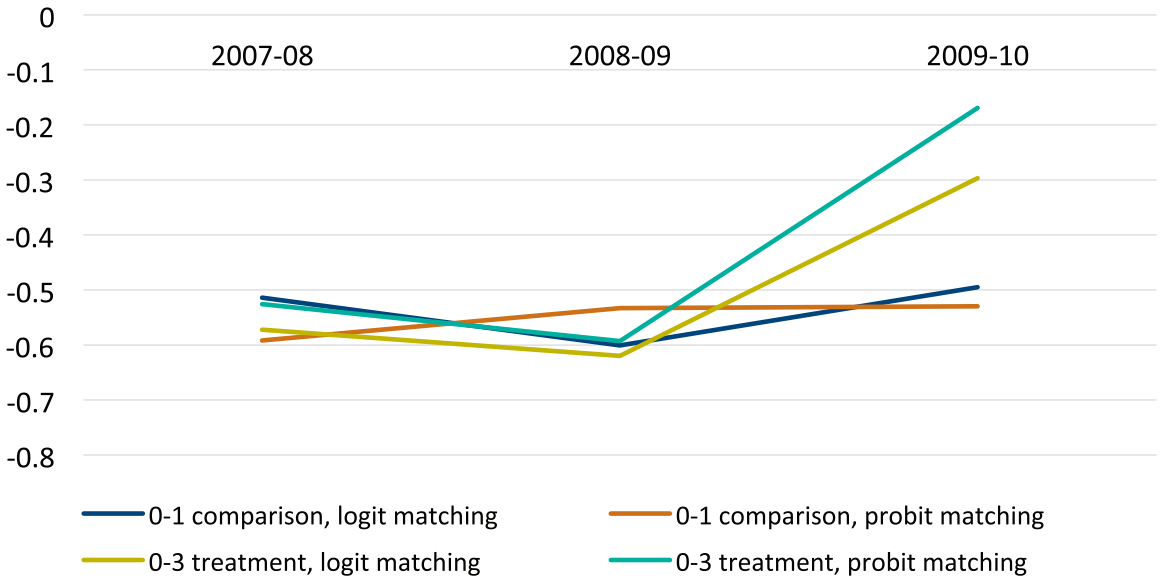
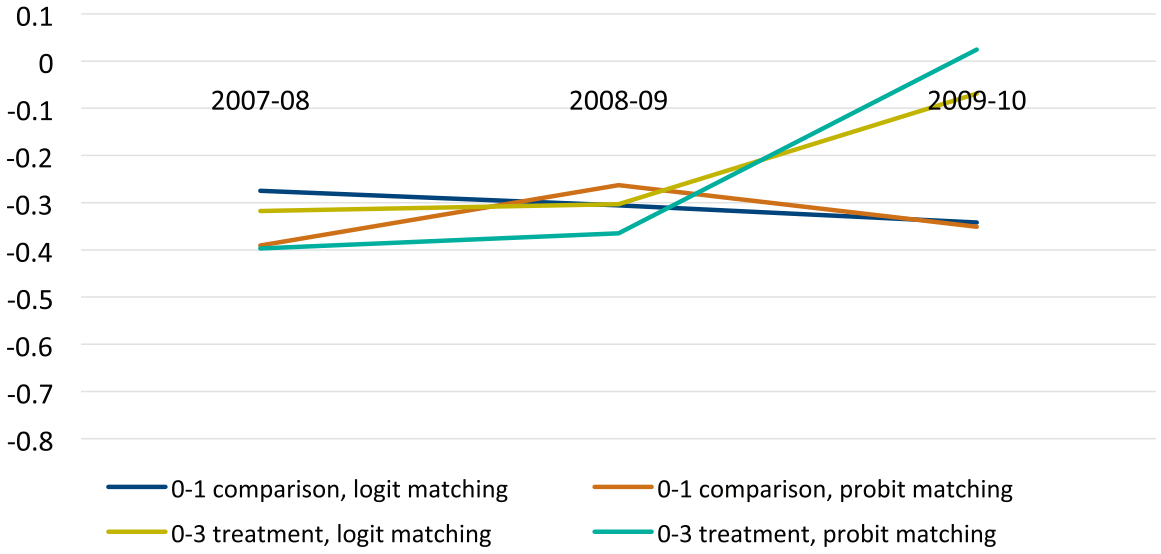


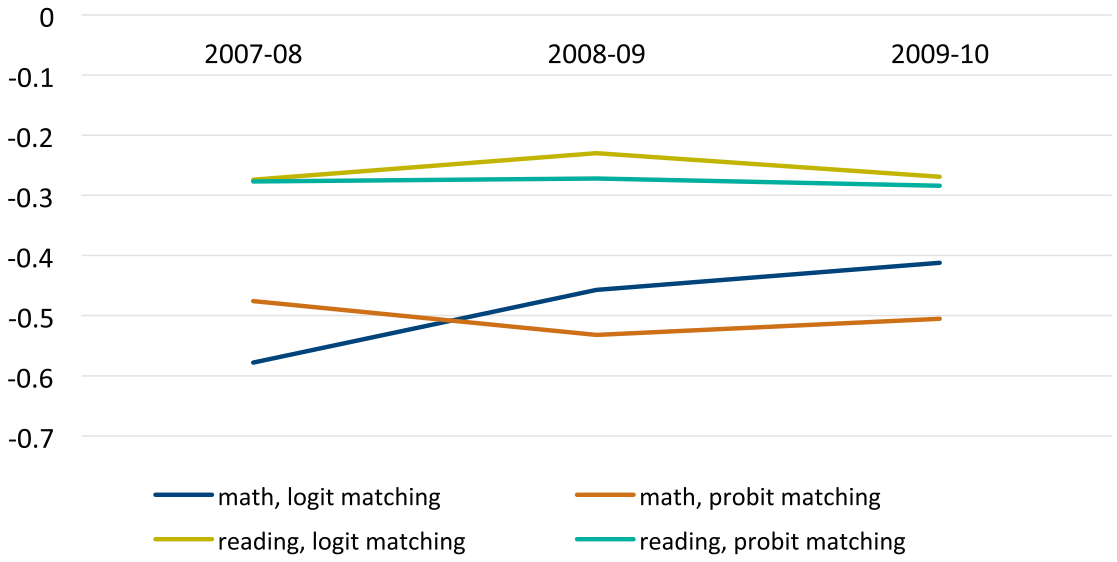
Figure 26: PSM estimates of effects of EdChoice participation on reading scores, 2007–08 cohort, different models (relative to 2006–07)



It may, for some reason, still be the case that the negative estimated effects of EdChoice participation are due to multiyear negative consequences of school moves for students who were old enough to have been tested in public schools before they moved. In order to directly address this question, we compare the students who moved to private schools under the EdChoice program to their closest matches among students from comparison schools (here, we choose the schools in the 0–1 comparison category, those schools who just barely missed eligibility for EdChoice vouchers) who also changed schools in 2007–08—but to other public schools rather than to private schools (as, recall, these comparison students were not eligible for a voucher under the EdChoice program). There were 2,576 closely matched public school movers in the comparison schools to compare with the 445 EdChoice participants in the analysis. These comparisons are seen in the

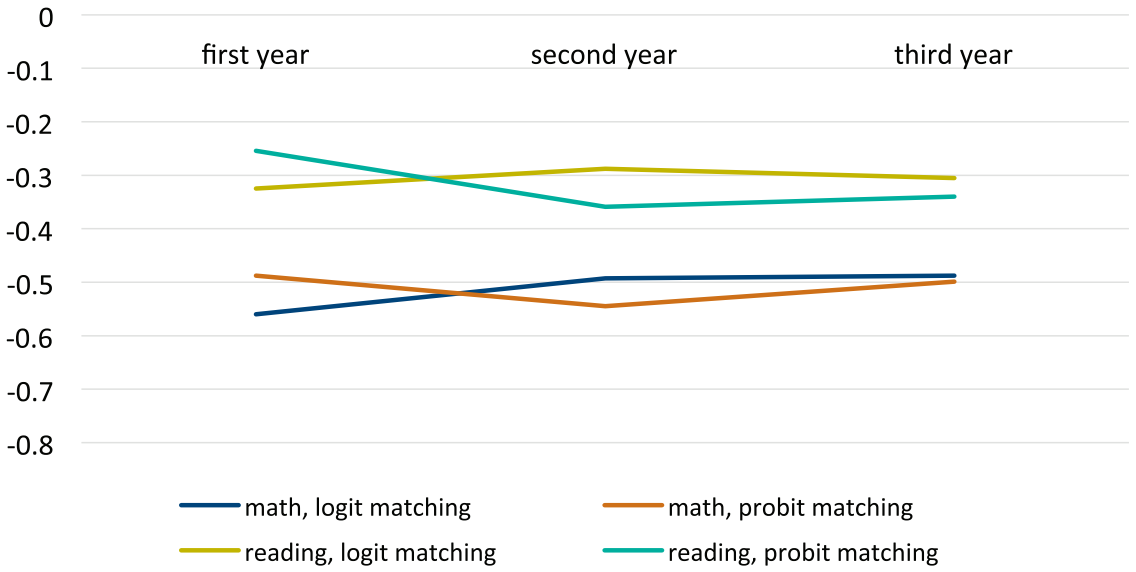
figure below and in appendix table 7. As is evident, the comparisons between EdChoice participants and closely matched public school changers in comparison schools look quite similar to those with all closely matched students in comparison schools. As a consequence, the evidence suggests that the negative findings are **not** due to the fact that EdChoice participants necessarily changed schools in 2007–08.

Figure 27: PSM estimates of effects of EdChoice participation on test scores, comparing participants to public school movers, zero-to-one-point comparison (relative to 2006–07)



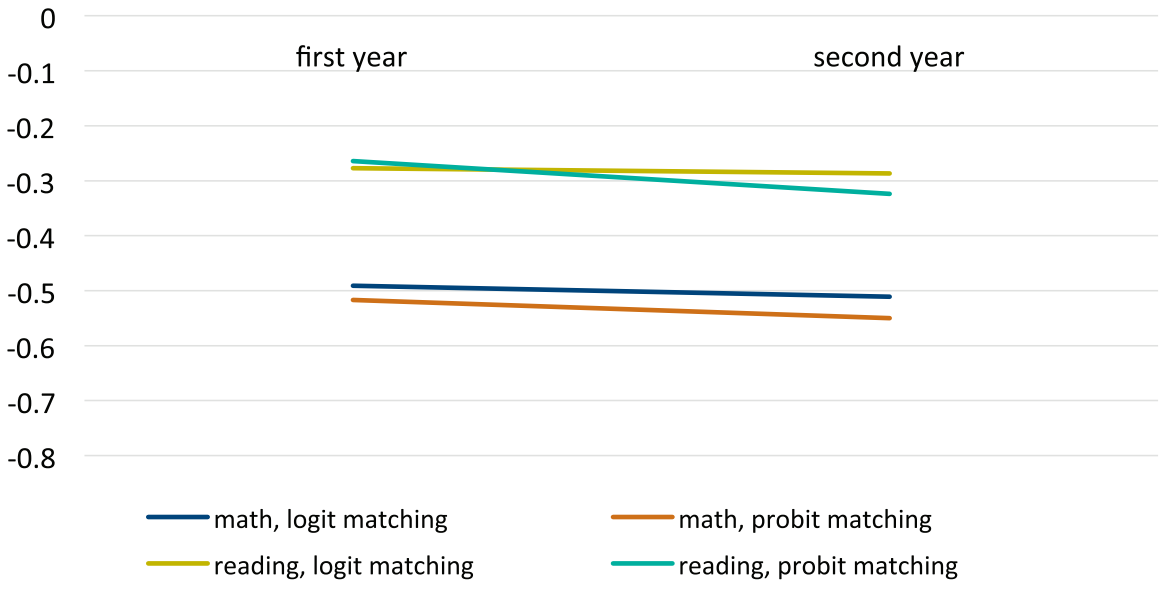
So far, we have concentrated on the students newly eligible in 2007–08. However, we can carry out this analysis to later years as well, given that we observe test-score data through the 2012–13 academic year. In the analysis that follows, we pool all of the EdChoice participants who became newly eligible at any time between 2007–08 and 2010–11 and followed them over the first three years of their participation. (As before, we are presenting the results of the pooled cross-section approach rather than the panel approach, so many of the students in year one are not observed in subsequent years.) In the following figure as well as appendix table 8, we compare these 876 newly eligible EdChoice participants to the 19,776 most closely matched students from schools that were never eligible for EdChoice vouchers but also were in the very close (zero-to-one-point) comparison range. As can be seen, the results that include the more recent EdChoice participants look very similar to those from the 2007–08 new eligibility cohort.

Figure 28: PSM estimates of effects of EdChoice participation on test scores, 2007–08 to 2010–11 first-eligible cohorts, zero-to-one-point comparison (relative to 2006–07)



We can go one step further still and limit ourselves to cohorts where we only observe two post-eligibility years, thereby pooling together all newly eligible EdChoice participants entering private schools between 2007–08 and 2011–12 (and, of course, still have test-score histories). This brings in another seventy-four private school students who were first eligible for EdChoice vouchers in 2011–12, along with 3,884 additional public school students in comparison schools. As can be seen in the following figure, as well as in appendix table 9, the patterns of findings continue to present the same story of EdChoice participants faring considerably worse in the private schools than very similar comparison students fared in the public schools.

Figure 29: PSM estimates of effects of EdChoice participation on test scores, 2007–08 to 2011–12 first-eligible cohorts, zero-to-one-point comparison (relative to 2006–07)



These results are certainly not without their caveats. As we mentioned at the beginning of this section, the ideal approach to identify the effects of private school participation under an EdChoice voucher would have been to make use of random assignment generated by a lottery, but that is not an option given the ways in which the EdChoice vouchers are allocated. PSM models are not ideal because there may still be unobserved differences between EdChoice participants and nonparticipants who were very closely matched on the basis of observable factors including prior test scores; however, this is somewhat less of an issue in the PSM approach that we implemented because we are not comparing people who made the choice to participate versus people who made the choice not to participate, as is very often the case with PSM approaches. In our implementation of PSM, we are comparing people who made the choice to participate versus people who had no choice regarding participation but who have similar observable attributes, which lends some additional credibility to the approach—and, of course, in doing so we are only able to focus on public schools that were among the highest performing of the voucher-eligible schools. Therefore, we cannot generalize these findings to those students coming from lower-performing public schools.

Weighing the remarkable consistency of the evidence against the limitations of the PSM approach, our conclusion is that participation in the EdChoice program likely reduced students' reading and mathematics scores relative to what would have occurred in the public sector—for those students who had previously attended the highest performing of the EdChoice-eligible schools. This may be because the students attended lower-quality private schools than the public schools that they left (especially because the public schools likely performed somewhat better as a consequence of the EdChoice program, though the improvement in the public schools is nowhere near as large as the estimated reduction in participants' scores after going to private schools). It may also be that the private schools attended are not necessarily lower quality but are focused on different sets of skills and competencies, or it may be that the private schools attended under the EdChoice program may not have emphasized the state assessments to the degree to which the public schools did. Although Ohio state law has required the public reporting of private school average test scores for students participating in the EdChoice program since 2009–10, it is still almost surely the case that participating private schools did not have curricula as well aligned to the state assessments as did the public schools, and private schools face different degrees of public accountability tied to the state assessments than do public schools. It is clear that there remains a need for a deeper understanding of the factors (such as public-private differences in curriculum alignment, attributes of schools participating in the program, consequences of differences in accountability between public and private schools, potentially different effects for students coming from especially low-performing public schools versus relatively high-performing public schools from which students were voucher-eligible, and so forth) that contribute to the observed differences in student outcomes between public and private schools. Such an evaluation is beyond the scope of this present project.

Though most studies of voucher participation in other settings find positive or zero estimates of participation on student test scores, it is not unheard of to find negative test-score estimates of voucher participation, even estimates that are of similar magnitude to those found in Ohio. Most recently, a lottery study regarding school vouchers in Louisiana by Atila Abdulkadiroglu, Parag Pathak, and Christopher Walters ("School Vouchers and Student Achievement: First-Year Evidence from the Louisiana Scholarship Program," National Bureau of Economic Research Working Paper 21839, December 2015) found consistent evidence of very large negative consequences of voucher participation on student test scores, particularly in mathematics, though a subsequent study by Jonathan Mills and Patrick Wolf (*The Effects of the Louisiana Scholarship Program on Student Achievement after Two Years*, University of Arkansas School Choice Demonstration Project, 2016) indicates that year-two results are less negative than those observed in the first year. Abdulkadiroglu and his coauthors conducted a series of analyses to suggest that the schools participating in the program tended to be those that had been rapidly losing enrollment and might have been more likely to recruit voucher students; carrying out a similar analysis in Ohio is beyond the scope of this report but would be a very valuable area for future investigation. Likewise, it is possible that in the Ohio case, systematically different schools admitted Kindergarten entrants, whom we cannot study, versus the schools that admitted entrants in late-elementary or middle grades; this could also be a potential explanation for our findings.

6. Summary

Taken together, the results of this report present a mixed bag of findings regarding the EdChoice voucher program. Although the evidence is not completely unambiguous, the weight of the evidence indicates that EdChoice eligibility improved reading and mathematics outcomes for the students affected. We suspect that this is coming through increased competition for lower-ranked public schools as well as a desire for these schools to improve to avoid losing students to the voucher program; we suspect that the competition is a leading explanation rather than merely avoidance of grading stigma because the regression-discontinuity approaches focusing on the second-best PI are designed to concentrate particularly on the voucher-eligibility component of the system, rather than on the school ratings themselves. We find evidence that the program attracts relatively high-scoring and comparatively advantaged eligible students (though these students are still overwhelmingly low scoring and disadvantaged as a group, relative to the state as a whole) and that this may be due to programmatic rules that require private school admission before voucher application, rather than the reverse, which is seen in other locales such as Florida. And though EdChoice eligibility apparently improves student test scores in general, this is not the case for those who actually use their vouchers to attend private schools, having previously attended relatively high-performing public schools among the EdChoice-eligible schools. Those eligible students (coming from these relatively high-performing public schools) who attend private schools appear to fare considerably worse than we predict that they would have performed had they remained in the public schools. These are averages, of course, and there are some reasons to believe that the private school experiences of EdChoice participants may be better than what we estimate. For instance, private schools participating in the EdChoice program do not face the same high stakes associated with state testing that is aligned to public school curricula but not to any particular private school curriculum. Although since 2009–10, Ohio state law has required the public reporting of average performance of private school students participating in the EdChoice program, there are no formal sanctions or rewards for private schools associated with performance on the state tests. In addition, of course, the experiences of private school students coming from public schools farther away from the threshold of eligibility may have been considerably different from those observed using the methods employed in this report. Nonetheless, this analysis is the best that we were able to do with the information at hand, suggesting that deeper study into the causes of these performance differences—related to differences in school quality, test-curriculum alignment, or other factors—should be a priority.

Appendices

Appendix 1: Analysis of overall effects in the initial 2006–07 eligibility wave

In this report, we focus on the model specification from the second wave of eligibility, rather than the first wave of ninety-nine EdChoice-eligible schools in 2006–07. We strongly prefer the second-wave analysis for two principal reasons. First, the more complicated nature of eligibility—the fact that the second-best school rating determined eligibility, as opposed to the best school rating determining eligibility—makes causal inference more defensible. Second, because the program rules changed dramatically in Fall 2006, making many more schools eligible for future rounds of vouchers, many schools that were "untreated" in 2006–07 essentially became partially "treated" by the threat of vouchers in the 2006–07 academic year. If schools that just missed voucher eligibility in 2006–07 faced additional competitive pressure in 2006–07 as a consequence of the program-rule-change announcement and responded accordingly by improving performance, this would introduce bias against finding a positive estimated effect of EdChoice eligibility in the first round of eligibility.

Nonetheless, for the sake of completeness, and with these provisos, we repeat the regression-discontinuity analysis for the initial 2006–07 wave of eligibility and, again (for the three different donut-hole specifications), present analyses where (1) we include no control variables and force the relationship between PI and test-score growth to be the same on both sides of the threshold (same slope); (2) we include no controls but allow the relationship to be different above versus below the threshold (different slopes); (3) we include controls for grade in 2004–05, sex, race and ethnicity, and economic disadvantage but impose the same-slope assumption; and (4) we include controls and allow the different-slopes assumption. In this set of analyses, we back up the initial test scores to the previous year as well, so we investigate test-score growth between 2004–05 and 2006–07. As can be seen in the figure below (as well as in statistical tabular form in appendix table 10), the results still indicate positive effects of initial-round EdChoice eligibility on reading growth from 2004–05 to 2006–07, but the level of statistical significance tends to be lower, as are the magnitudes of the estimated findings even when statistically significant. In the case of mathematics, there is no evidence of a systematic relationship, either positive or negative, between initial EdChoice eligibility and test-score growth. Therefore, while there is reason to believe that these results are underestimates of a true effect that is more positive that those presented below (because many of the untreated schools in this analysis were actually treated relatively early in the 2006–07 academic year)—and while we strongly prefer the estimates from the second round of the EdChoice program because we believe those results to be the more scientifically credible—these first-round results suggest that caution is still warranted when concluding that the EdChoice program led to improvements in the schools that became voucher eligible.

Figure A1: Estimated effects of first-round EdChoice eligibility on reading growth from 2004–05 to 2006–07

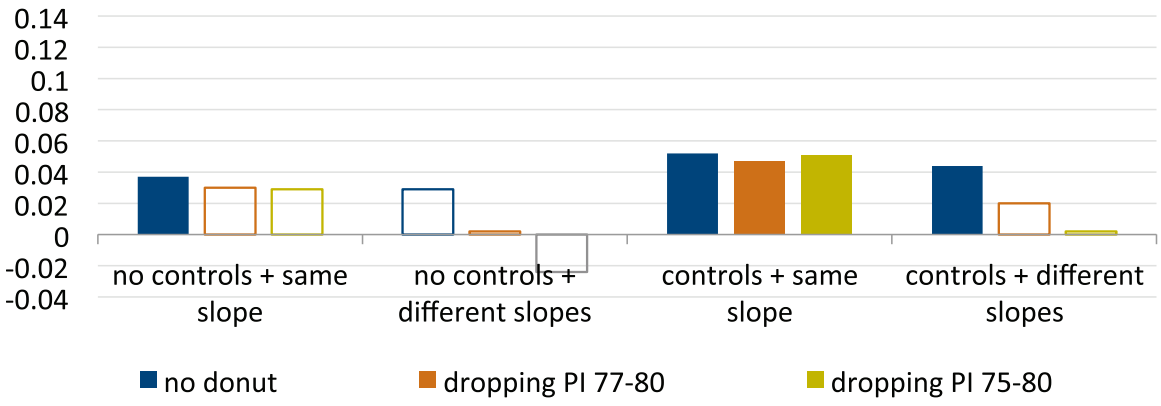
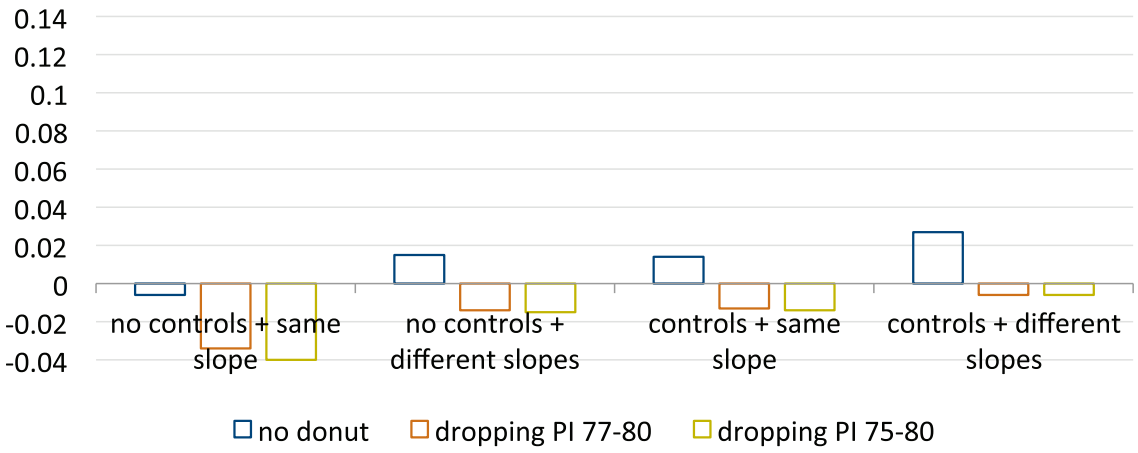


Figure A2: Estimated effects of first-round EdChoice eligibility on mathematics growth from 2004–05 to 2006–07



Taking these less rigorous first-round findings together with the more empirically valid 2007–08 EdChoice round results, our general conclusion is that the EdChoice program likely improved test scores for newly eligible students. Although the results are somewhat sensitive to model specification as well as timing, the preponderance of the evidence supports the notion of positive overall effects of the program.

Appendix 2: Tables

Table A1: Total effects of EdChoice eligibility, focusing on schools first eligible in 2007–08 (and dropping the ninety-nine schools first eligible in 2006–07)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
		No donut			Dropping schools with second-best PI between 77 and 79.999				Dropping schools with second-best PI between 75 and 79.999			
VARIABLES	A1	A2	A3	A4								
	Growth in standardized reading scores from 2005–06 to 2007–08											
Eligible	0.064***	0.055***	0.082***	0.078***	0.070***	0.057*	0.091***	0.095***	0.072***	0.050	0.092***	0.101**
	(0.018)	(0.021)	(0.018)	(0.021)	(0.022)	(0.033)	(0.021)	(0.031)	(0.025)	(0.047)	(0.024)	(0.044)
Observations	419,047	419,047	419,047	419,047	405,669	405,669	405,669	405,669	397,891	397,891	397,891	397,891
	Growth in standardized mathematics scores from 2005–06 to 2007–08											
Eligible	0.110***	0.082***	0.129***	0.107***	0.116***	0.058*	0.142***	0.104***	0.131***	0.060	0.156***	0.120**
	(0.021)	(0.025)	(0.021)	(0.025)	(0.024)	(0.035)	(0.024)	(0.033)	(0.028)	(0.053)	(0.027)	(0.049)
Observations	418,749	418,749	418,749	418,749	405,373	405,373	405,373	405,373	397,590	397,590	397,590	397,590
	Growth in standardized reading scores from 2005–06 to 2008–09											
Eligible	0.053***	0.028	0.070***	0.053**	0.055**	–0.010	0.072***	0.031	0.063**	–0.036	0.078***	0.020
	(0.020)	(0.022)	(0.019)	(0.022)	(0.024)	(0.033)	(0.022)	(0.032)	(0.027)	(0.047)	(0.026)	(0.046)
Observations	300,270	300,270	300,270	300,270	290,880	290,880	290,880	290,880	285,438	285,438	285,438	285,438
	Growth in standardized mathematics scores from 2005–06 to 2008–09											
Eligible	0.087***	0.048	0.102***	0.072**	0.092***	0.000	0.110***	0.045	0.107***	–0.025	0.123***	0.034
	(0.026)	(0.030)	(0.024)	(0.029)	(0.029)	(0.037)	(0.027)	(0.035)	(0.032)	(0.054)	(0.031)	(0.051)
Observations	299,874	299,874	299,874	299,874	290,491	290,491	290,491	290,491	285,060	285,060	285,060	285,060
Different slopes		X		X		X		X		X		X
Controls			X	X			X	X			X	X

Note: Standard errors are clustered at 2005–06 school level. Eligible is defined as below 0 threshold where threshold is PI–80. Controls include female, white-non-Hispanic, Black-non-Hispanic, Hispanic, economically disadvantaged, and grade dummies. All regressions also include running variable (PI relative to threshold) and interactions with being above or below threshold in the case of different slopes.

Table A2: Total effects of EdChoice eligibility, focusing on schools first eligible in 2007–08 (and dropping the ninety-nine schools first eligible in 2006–07): Public school attendees only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
		No donut			Dropping schools with second-best PI between 77 and 79.999				Dropping schools with second-best PI between 75 and 79.999			
VARIABLES	A1	A2	A3	A4								
	Growth in standardized reading scores from 2005–06 to 2007–08											
Eligible	0.065***	0.056***	0.083***	0.079***	0.072***	0.059*	0.093***	0.097***	0.073***	0.053	0.093***	0.103**
	(0.018)	(0.021)	(0.018)	(0.021)	(0.022)	(0.033)	(0.021)	(0.031)	(0.025)	(0.047)	(0.024)	(0.045)
Observations	418,477	418,477	418,477	418,477	405,179	405,179	405,179	405,179	397,449	397,449	397,449	397,449
	Growth in standardized mathematics scores from 2005–06 to 2007–08											
Eligible	0.112***	0.084***	0.131***	0.108***	0.119***	0.060*	0.143***	0.106***	0.134***	0.063	0.158***	0.122**
	(0.021)	(0.025)	(0.021)	(0.025)	(0.024)	(0.035)	(0.024)	(0.034)	(0.028)	(0.053)	(0.027)	(0.050)
Observations	418,179	418,179	418,179	418,179	404,883	404,883	404,883	404,883	397,148	397,148	397,148	397,148
	Growth in standardized reading scores from 2005–06 to 2008–09											
Eligible	0.055***	0.031	0.071***	0.055**	0.056**	–0.006	0.073***	0.034	0.064**	–0.031	0.079***	0.024
	(0.020)	(0.023)	(0.019)	(0.022)	(0.024)	(0.033)	(0.023)	(0.032)	(0.027)	(0.048)	(0.026)	(0.046)
Observations	299,385	299,385	299,385	299,385	290,124	290,124	290,124	290,124	284,748	284,748	284,748	284,748
	Growth in standardized mathematics scores from 2005–06 to 2008–09											
Eligible	0.091***	0.053*	0.105***	0.077***	0.096***	0.007	0.113***	0.051	0.111***	–0.016	0.125***	0.042
	(0.026)	(0.030)	(0.024)	(0.029)	(0.029)	(0.037)	(0.027)	(0.035)	(0.032)	(0.054)	(0.031)	(0.051)
Observations	298,990	298,990	298,990	298,990	289,736	289,736	289,736	289,736	284,371	284,371	284,371	284,371
Different slopes		X		X		X		X		X		X
Controls			X	X			X	X			X	X

Note: Standard errors are clustered at 2005–06 school level. Eligible is defined as below 0 threshold where threshold is PI–80. Controls include female, white-non-Hispanic, Black-non-Hispanic, Hispanic, economically disadvantaged, and grade dummies. All regressions also include running variable (PI relative to threshold) and interactions with being above or below threshold in the case of different slopes.

Table A3: Heterogeneity in total effects of EdChoice eligibility, by demographic characteristics, focusing on schools first eligible in 2007–08 (and dropping the ninety-nine schools first eligible in 2006–07)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
VARIABLES	White		Black		Hispanic		Ever disadvantaged		Never disadvantaged		Girl		Boy	
	Growth in standardized reading scores from 2005–06 to 2007–08													
Eligible	0.137***	0.113***	–0.008	0.007	0.187***	0.001	0.030	0.055	0.131***	0.149***	0.063***	0.036	0.090***	0.081**
	(0.023)	(0.035)	(0.046)	(0.051)	(0.053)	(0.075)	(0.023)	(0.034)	(0.031)	(0.053)	(0.022)	(0.032)	(0.024)	(0.036)
Observations	335,672	335,672	45,840	45,840	7,486	7,486	204,975	204,975	200,694	200,694	198,299	198,299	207,370	207,370
	Growth in standardized mathematics scores from 2005–06 to 2007–08													
Eligible	0.188***	0.129***	–0.027	–0.016	0.175***	–0.010	0.037	0.043	0.140***	0.121**	0.109***	0.050	0.126***	0.067*
	(0.026)	(0.038)	(0.047)	(0.052)	(0.062)	(0.091)	(0.026)	(0.036)	(0.036)	(0.056)	(0.026)	(0.037)	(0.025)	(0.036)
Observations	335,230	335,230	45,829	45,829	7,578	7,578	204,843	204,843	200,530	200,530	198,131	198,131	207,242	207,242
	Growth in standardized reading scores from 2005–06 to 2008–09													
Eligible	0.097***	0.040	–0.014	–0.029	0.142**	–0.011	0.009	–0.007	0.075**	0.062	0.033	–0.030	0.089***	0.019
	(0.026)	(0.036)	(0.043)	(0.049)	(0.062)	(0.096)	(0.024)	(0.034)	(0.031)	(0.047)	(0.024)	(0.034)	(0.026)	(0.036)
Observations	240,802	240,802	32,537	32,537	5,379	5,379	149,613	149,613	141,267	141,267	142,159	142,159	148,721	148,721
	Growth in standardized mathematics scores from 2005–06 to 2008–09													
Eligible	0.148***	0.059	–0.037	–0.049	0.137**	–0.053	0.008	–0.012	0.108***	0.025	0.075**	–0.012	0.109***	0.014
	(0.031)	(0.038)	(0.051)	(0.056)	(0.070)	(0.094)	(0.028)	(0.038)	(0.040)	(0.058)	(0.030)	(0.040)	(0.031)	(0.039)
Observations	240,355	240,355	32,488	32,488	5,441	5,441	149,407	149,407	141,084	141,084	141,978	141,978	148,513	148,513
Different slopes		X		X		X		X		X		X		X

Note: Standard errors are clustered at 2005–06 school level. Eligible is defined as below 0 threshold where threshold is PI–80. Controls include female, white-non-Hispanic, Black-non-Hispanic, Hispanic, economically disadvantaged, and grade dummies. All regressions also include running variable (PI relative to threshold) and interactions with being above or below threshold in the case of different slopes.

Table A4: Heterogeneity in total effects of EdChoice eligibility, focusing on schools first eligible in 2007–08 (and dropping the ninety-nine schools first eligible in 2006–07), by initial school type

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Elementary school		Middle school		Combined school	
			Growth in standardized reading scores from 2005–06 to 2007–08			
Eligible	0.078***	0.012	0.178***	0.191***	–0.025	0.006
	(0.024)	(0.037)	(0.052)	(0.074)	(0.102)	(0.100)
Observations	312,827	312,827	75,420	75,420	17,422	17,422
			Growth in standardized mathematics scores from 2005–06 to 2007–08			
Eligible	0.113***	0.003	0.199***	0.191*	0.122	0.099
	(0.028)	(0.040)	(0.066)	(0.099)	(0.089)	(0.090)
Observations	312,555	312,555	75,409	75,409	17,409	17,409
			Growth in standardized reading scores from 2005–06 to 2008–09			
Eligible	0.069***	–0.040	0.213**	0.241*	–0.033	–0.021
	(0.025)	(0.036)	(0.087)	(0.128)	(0.115)	(0.110)
Observations	264,047	264,047	14,954	14,954	11,879	11,879
			Growth in standardized mathematics scores from 2005–06 to 2008–09			
Eligible	0.095***	–0.036	0.280**	0.148	0.102	0.077
	(0.031)	(0.041)	(0.123)	(0.158)	(0.105)	(0.106)
Observations	263,689	263,689	14,954	14,954	11,848	11,848
Different slopes		X		X		X

Note: Standard errors are clustered at 2005–06 school level. Eligible is defined as below 0 threshold where threshold is PI–80. Controls include female, white-non-Hispanic, Black-non-Hispanic, Hispanic, economically disadvantaged, and grade dummies. All regressions also include running variable (PI relative to threshold) and interactions with being above or below threshold in the case of different slopes. The three most common schools at the elementary level are K–5, K–6, and K-4. The three most common schools at the middle level are 6–8, 5–8, and 4–8. The three most common combined schools are K–8, pre-K–8, and K–7.

Table A5: PSM estimates of the effects of EdChoice participation: Control students attended schools zero to three points above threshold

	(1)	(2)	(3)	(4)	(5)	(6)
	Repeated cross-section			Three-year panel		
	2007–08 scores	2008–09 scores	2009–10 scores	2007–08 scores	2008–09 scores	2009–10 scores
			Mathematics (logit matching)			
Private school 2007–08	–0.496***	–0.535***	–0.448***	–0.668***	–0.539***	–0.542***
	(0.044)	(0.049)	(0.057)	(0.067)	(0.066)	(0.071)
Number treated	445	405	301	205	205	205
Observations	19,667	16,403	11,712	11,197	11,197	11,197
			Mathematics (probit matching)			
Private school 2007–08	–0.523***	–0.554***	–0.545***	–0.655***	–0.569***	–0.582***
	(0.047)	(0.050)	(0.056)	(0.065)	(0.061)	(0.063)
Number treated	445	405	301	205	205	205
Observations	19,667	16,403	11,712	11,197	11,197	11,197
			Reading (logit matching)			
Private school 2007–08	–0.348***	–0.245***	–0.307***	–0.460***	–0.334***	–0.307***
	(0.049)	(0.054)	(0.061)	(0.067)	(0.069)	(0.074)
Number treated	445	405	301	205	205	205
Observations	19,667	16,403	11,712	11,197	11,197	11,197
			Reading (probit matching)			
Private school 2007–08	–0.341***	–0.325***	–0.398***	–0.445***	–0.354***	–0.383***
	(0.049)	(0.052)	(0.054)	(0.069)	(0.067)	(0.069)
Number treated	445	405	301	205	205	205
Observations	19,667	16,403	11,712	11,197	11,197	11,197

Table A6: PSM estimates of the effects of EdChoice participation: Investigating different criteria for inclusion into the analysis; repeated cross-section

	(1)	(2)	(3)	(4)	(5)	(6)
	Control students attended schools zero to one point above threshold			Control students attended schools zero to three points above threshold; treatment students attended schools zero to three points below threshold		
	2007–08 scores	2008–09 scores	2009–10 scores	2007–08 scores	2008–09 scores	2009–10 scores
			Mathematics (logit matching)			
Private school 2007–08	–0.514***	–0.601***	–0.495***	–0.572***	–0.620***	–0.297**
	(0.048)	(0.050)	(0.064)	(0.103)	(0.113)	(0.145)
Number treated	445	405	301	82	75	51
Observations	7,491	6,023	4,184	19,304	16,073	11,462
			Mathematics (probit matching)			
Private school 2007–08	–0.592***	–0.533***	–0.530***	–0.526***	–0.593***	–0.169
	(0.048)	(0.052)	(0.063)	(0.083)	(0.111)	(0.143)
Number treated	445	405	301	82	75	51
Observations	7,491	6,023	4,184	19,304	16,073	11,462
			Reading (logit matching)			
Private school 2007–08	–0.275***	–0.306***	–0.342***	–0.318***	–0.303***	–0.069
	(0.050)	(0.055)	(0.063)	(0.115)	(0.096)	(0.131)
Number treated	445	405	301	82	75	51
Observations	7,491	6,023	4,184	19,304	16,073	11,462
			Reading (probit matching)			
Private school 2007–08	–0.391***	–0.263***	–0.351***	–0.397***	–0.365***	0.024
	(0.052)	(0.054)	(0.066)	(0.074)	(0.114)	(0.142)
Number treated	445	405	301	82	75	51
Observations	7,491	6,023	4,184	19,304	16,073	11,462

Table A7: Comparing private school students newly eligible in 2007–08 to those ineligible but changing public schools in 2007–08: PSM estimates of the effects of EdChoice participation, where control students attended schools zero to one point above threshold; repeated cross-section

	(1)	(2)	(3)	(4)	(5)	(6)
		Mathematics			Reading	
	2009–10 scores	2010–11 scores	2011–12 scores	2009–10 scores	2010–11 scores	2011–12 scores
			Logit matching			
Private school 2007–08	–0.578***	–0.457***	–0.412***	–0.274***	–0.230***	–0.269***
	(0.057)	(0.064)	(0.077)	(0.063)	(0.072)	(0.084)
Number treated	445	405	301	445	405	301
Observations	3,021	2,634	1,833	3,021	2,634	1,833
			Probit matching			
Private school 2007–08	–0.476***	–0.532***	–0.505***	–0.277***	–0.272***	–0.284***
	(0.057)	(0.065)	(0.074)	(0.064)	(0.069)	(0.075)
Number treated	445	405	301	445	405	301
Observations	3,021	2,634	1,833	3,021	2,634	1,833

Table A8: Stacking first-time eligibility, PSM estimates of the effects of EdChoice participation: Students first-time eligible between 2007–08 and 2010–11 are pooled, and control students attended schools zero to one point above threshold; repeated cross-section

	(1)	(2)	(3)	(4)	(5)	(6)
	Mathematics			Reading		
	First-year scores	Second-year scores	Third-year scores	First-year scores	Second-year scores	Third-year scores
	Panel A1: Logit matching					
Private school	–0.560***	–0.493***	–0.488***	–0.325***	–0.288***	–0.305***
	(0.037)	(0.038)	(0.043)	(0.038)	(0.039)	(0.046)
Number treated	876	761	554	876	761	554
Observations	20,652	13,737	8,399	20,652	13,737	8,399
	Panel A2: Probit matching					
Private school	–0.488***	–0.545***	–0.499***	–0.254***	–0.359***	–0.340***
	(0.035)	(0.039)	(0.045)	(0.037)	(0.040)	(0.050)
Number treated	876	761	554	876	761	554
Observations	20,652	13,737	8,399	20,652	13,737	8,399

Table A9: Stacking first-time eligibility, PSM estimates of the effects of EdChoice participation: Students first-time eligible between 2007–08 and 2010–12 are pooled, and control students attended schools zero to one points above threshold; repeated cross-section

	(1)	(2)	(4)	(5)
	Mathematics		Reading	
	First-year scores	Second-year scores	First-year scores	Second-year scores
	Panel A1: Logit matching			
Private school	–0.491***	–0.511***	–0.277***	–0.287***
	(0.033)	(0.037)	(0.035)	(0.037)
Number treated	950	826	950	826
Observations	24,610	16,041	24,610	16,041
	Panel A2: Probit matching			
Private school	–0.517***	–0.550***	–0.264***	–0.324***
	(0.035)	(0.037)	(0.038)	(0.040)
Number treated	950	826	950	826
Observations	24,610	16,041	24,610	16,041

Table A10: Total effects of EdChoice eligibility, focusing on ninety-nine schools first eligible in 2006–07

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	No donut				Dropping schools with second-best PI between 77 and 79.999				Dropping schools with second-best PI between 75 and 79.999			
VARIABLES												
	Panel A1: Panel (2004–05 to 2006–07) in reading. Residualized growth between 2004–05 and 2006–07											
Eligible	0.037**	0.029	0.052***	0.044**	0.030	0.002	0.047**	0.020	0.029	–0.024	0.051**	0.002
	(0.018)	(0.021)	(0.017)	(0.020)	(0.021)	(0.030)	(0.020)	(0.030)	(0.023)	(0.038)	(0.022)	(0.037)
Observations	355,519	355,519	355,519	355,519	341,283	341,283	341,283	341,283	331,340	331,340	331,340	331,340
	Panel B1: Panel (2004–05 to 2006–07) in mathematics. Residualized growth between 2004–05 and 2006–07											
Eligible	–0.006	0.015	0.014	0.027	–0.034	–0.014	–0.013	–0.006	–0.040	–0.015	–0.014	–0.006
	(0.024)	(0.026)	(0.018)	(0.019)	(0.026)	(0.032)	(0.020)	(0.024)	(0.029)	(0.043)	(0.022)	(0.031)
Observations	248,881	248,881	248,881	248,881	238,446	238,446	238,446	238,446	231,370	231,370	231,370	231,370
	Panel C1: Panel (2004–05 to 2007–08) in reading. Residualized growth between 2004–05 and 2007–08											
Eligible	0.046***	0.023	0.054***	0.037*	0.047**	–0.005	0.056***	0.014	0.061***	–0.013	0.073***	0.016
	(0.018)	(0.020)	(0.017)	(0.020)	(0.021)	(0.029)	(0.020)	(0.029)	(0.023)	(0.039)	(0.022)	(0.038)
Observations	269,601	269,601	269,601	269,601	258,907	258,907	258,907	258,907	251,689	251,689	251,689	251,689
	Panel D1: Panel (2004–05 to 2007–08) in mathematics. Residualized growth between 2004–05 and 2007–08											
Eligible	0.008	–0.002	0.020	0.010	–0.006	–0.049	0.009	–0.028	–0.000	–0.070	0.016	–0.043
	(0.025)	(0.027)	(0.023)	(0.026)	(0.028)	(0.034)	(0.026)	(0.033)	(0.030)	(0.043)	(0.028)	(0.041)
Observations	165,701	165,701	165,701	165,701	158,679	158,679	158,679	158,679	154,230	154,230	154,230	154,230
Different slopes		X		X		X		X		X		X
Controls			X	X			X	X			X	X

Note: Standard errors are clustered at 2005–06 school level. Eligible is defined as below 0 threshold where threshold is PI–80. Controls include female, white-non-Hispanic, Black-non-Hispanic, Hispanic, economically disadvantaged, and grade dummies. All regressions also include running variable (PI relative to threshold) and interactions with being above or below threshold in the case of different slopes.

Endnotes

[1] Cleveland students and schools are not part of the EdChoice program, as Cleveland has a different longstanding school-voucher program in place.

[2] For details, see *Establishing a Baseline: Ohio's Education System As It Enters a New Era*, Public Impact and Thomas B. Fordham Institute, August 2015.

[3] In the 2006–07 academic year, only 58.8 percent of voucher students in testing grades were matchable to testing records, according to Ohio Department of Education data. This figure increased to 73.9 percent in 2007–08 and to 85.8 percent in 2008–09. Since that time, the figure has varied between 82.3 percent and 93.4 percent. It is not entirely clear why test reporting got much better after the initial year of private school testing, but the compliance patterns are consistent with what might be expected given a phase-in of a new program.

[4] Office of Budget and Management, *State of Ohio: Budget Highlights*, *Fiscal Years 2006 and 2007* (August 2005): http://obm.ohio.gov/budget/operating/doc/fy-06-07/budget-highlights.pdf

[5] Examples of cases in which relatively low-performing students are more likely to attend private schools with a means-tested voucher include David Figlio, Cassandra Hart, and Molly Metzger, "Who Uses a Means-Tested Scholarship, and What Do They Choose?" *Economics of Education Review*, April 2010, and Cassandra Hart, "Contexts Matter: Selection in Means-Tested School Voucher Programs," *Educational Evaluation and Policy Analysis*, June 2014. William Howell and Paul Peterson (with David Campbell and Patrick Wolf), *The Education Gap: Vouchers and Urban Schools*, Brookings Institution Press, 2006, find evidence of, at most, modest positive selection into private schools, a result generally consistent with Patrick Wolf, et al., *Evaluation of the DC Opportunity Scholarship Program: First Year Report*, U.S. Department of Education, 2005, and David Fleming, Joshua Cowen, John Witte, and Patrick Wolf, "Similar Students, Different Choices: Who Uses a School Voucher in an Otherwise Similar Population of Students?" *Education and Urban Society*, 2015.

[6] We know of no other voucher systems that operate exactly like the EdChoice case. In Washington, DC, for instance, families must first be deemed eligible for a voucher and students must be deemed admissible to their private school of choice before being entered into the lottery (or automatically awarded vouchers, in the case of no oversubscription). In Milwaukee, students are admitted to private schools before seeking financial aid through the voucher system, but private schools cannot apply admission standards to voucher students.

[7] We ascribe no value judgment regarding whether it is better or worse if comparatively high-achieving (or high-income, etc.) students are those who make use of the EdChoice vouchers. The purpose of this section is simply to make clear that the design of voucher eligibility can help to determine which students make use of a school voucher. Different families might avail themselves of an EdChoice voucher under different circumstances.

[8] In this report, we often describe the magnitude of our findings in terms of the black-white test-score gap in order to gauge how large the estimated effects or differences are in comparison to other differences that we observe in the data. It is beyond the scope of this report to speculate about the causes of the black-white test-score gap.

[9] In some cases, a history of economic disadvantage could come from attending a school where all students are categorically deemed economically disadvantaged for purposes of the National School Lunch Program. Blanket classification of students as eligible for free or reduced-price lunches would bias the comparison here only if participants or nonparticipants are disproportionately from such schools.

[10] See David Autor, David Figlio, Krzysztof Karbownik, Melanie Wasserman, and Jeffrey Roth, "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes," Northwestern University Institute for Policy Research Working Paper 15–16 (December 2015), for a detailed investigation of this pattern.

[11] Unfortunately, we do not have access to the data that would allow us to gauge directly the degree to which schools' application of admissions standards drives these differences.

[12] See, for example, David Figlio and Cassandra Hart, "Competitive Effects of Means-Tested School Vouchers," *American Economic Journal: Applied Economics*, January 2014, who show that the Florida means-tested school-voucher program induced positive competitive effects for public schools.

[13] See, for example, Martin West and Paul Peterson, "The Efficacy of Choice Threats Within School Accountability Systems: Results from Legislatively Induced Experiments," *Economic Journal*, March 2006, which contends that choice threats augment the effects of school accountability for low-rated schools.

[14] One example, David Figlio and Cecilia Rouse, "Do Accountability and Voucher Threats Improve Low-Performing Schools?" *Journal of Public Economics*, January 2006, argues that the positive effects on test scores of the initial roll-out of Florida's short-lived school-voucher system tied to school accountability were likely due more to grading stigma than to voucher threats per se.

[15] Though four of the ten transitions Forster studied had negative average transitions for voucher-eligible schools versus other schools, none of these four differences were statistically distinct from zero at conventional levels. On the other hand, three of the six positive average transition differences were statistically distinct from zero at conventional levels of significance.

[16] It would be desirable to study later rounds of the EdChoice voucher eligibility as well, but this becomes much more difficult to study because an increasingly large number of schools in the neighborhood of potential eligibility would have already become voucher eligible, and causal inference would become progressively more problematic. It is therefore best to limit the analysis to the first instance of a major policy change—that is, the change in the eligibility rules that first affected private school enrollment and eligibility in 2007–08.

[17] It is also possible for a school to receive a designation above academic watch if, for example, AYP status is met or there was a sufficient growth in the school's PI score. However, because these other factors are not continuous measures, we emphasize the PI as the principal driver of a school's designation.

[18] We very strongly prefer this approach to one analyzing the initial 2006–07 voucher-eligible schools, and we caution the reader that there are several important reasons to discount the first-round results. However, for completeness, we present results from the first-round EdChoice implementation in an appendix.

[19] This would occur if a school was rated under academic watch or below in two of the three years of 2003–04, 2004–05, or 2005–06. Eligibility determination occurred in 2006–07 for eligibility in 2007–08, but to be as certain as possible that we were not identifying the effects of EdChoice eligibility based on people who endogenously sorted into schools in anticipation of a voucher, we use the more conservative 2005–06 school attendance to determine eligibility.

[20] In some analyses, we show the results both including and excluding the voucher users. Invariably, the results are virtually identical regardless of how we treat the voucher users in the analysis.

[21] In order to estimate the lines reported in the figures, we weigh each point by the number of observations.

[22] As before, this is measured as the growth in standardized reading (mathematics) scores between 2005–06 and 2007–08.

[23] We present estimates that are not statistically distinct from zero at conventional levels in outline only.

[24] There are theoretical reasons to believe that the slopes of this relationship would vary on either side of the threshold. If some public schools affected by voucher competition change their behaviors as a consequence of the competition, this might influence the subsequently observed relationship between the PI and test-score growth. As a consequence, we report the results both ways to assess the degree to which the estimated effects of vouchers on public school performance varies depending on statistical-modeling assumptions.

[25] Here, we stratify based on the school type the student was attending during the 2005–06 academic year.

[26] One might wonder why we do not simply compare those students who use a voucher to those remaining in the public schools. This comparison would be highly problematic because there are obvious differences between those students who select into private schools and those who do not. In addition to the selection differences described in section 3 of this report, it is also certainly the case that those students who move to private schools were those who were the most motivated to change schools and, of course, students had to have obtained admission to a private school in order to make use of the EdChoice scholarship. Any credible estimate of performance effects must take this type of selection into account, which is what motivates our estimation approach.

[27] An example of the application of PSM in a participant-effects evaluation of a school-voucher program is John Witte, Patrick Wolf, Joshua Cowen, Deven Carlson, and David Fleming, "High Stakes Choice: Achievement and Accountability in the Nation's Oldest Urban Voucher Program," *Education Evaluation and Policy Analysis*, December 2014.

[28] One important continuing limitation is that everyone who is a participant has at least one type of unobserved factor that is correlated with participation in the program, while only some of the noneligible comparison group students will have the same unobservable variables correlated with participation. Therefore, it is still almost surely the case that the comparison groups are not identical on unobserved factors.

[29] In the analysis that follows, we consider comparison schools that never became voucher eligible in the future. We have also relaxed this restriction to look only at comparison schools that only remained voucher ineligible for the first three years of the program, and the results are extremely similar and available on request from the authors.

[30] We begin with the students newly eligible in 2007–08, because three-quarters or more of private school participants' testing records were only matchable to Ohio Department of Education student databases beginning in 2007–08. A considerably smaller fraction (58.8 percent) were matched in 2006–07, the first year of voucher eligibility for some students.

[31] We report the repeated cross-section analysis only in order to save space.

This page is left intentionally blank

This page is left intentionally blank

This page is left intentionally blank

Thomas B. Fordham Institute

100 E. Broad St., Suite 2430

Columbus, OH 43215

614-223-1580

www.edexcellence.net/ohio-policy

The Institute is neither connected with nor sponsored by Fordham University.