Effects of Scaling Up Private School Choice Programs on Public School Students

David N. Figlio, Northwestern University and NBER
Cassandra M.D. Hart, University of California, Davis
Krzysztof Karbownik, Emory University and NBER

February 2020

Using a rich dataset that merges student-level school records with birth records, and a student fixed effect design, we explore how the massive scale-up of a Florida private school choice program affected public school students' outcomes. Expansion of the program produced modestly larger benefits for students attending public schools that had a larger initial degree of private school options, measured prior to the introduction of the voucher program. These benefits include higher standardized test scores and lower absenteeism and suspension rates. Effects are particularly pronounced for lower-income students, but results are positive for more affluent students as well.

Acknowledgements: We are grateful to the Florida Departments of Education and Health for providing de-identified, matched data used in this analysis. Figlio acknowledges support from National Science Foundation, National Institute of Child Health and Human Development, the Bill and Melinda Gates Foundation, and the Institute for Education Sciences (CALDER grant), and the Shelter Hill Foundation for assistance in building the dataset and/or conducting this research. We thank Michael Dinerstein, Stefanie Fischer and seminar participants at the CESifo Economics of Education Meeting and University of New South Wales for helpful comments and suggestions. The conclusions expressed in this paper are those of the authors and do not represent the positions of the Florida Departments of Education and Health or those of our funders. All errors are our own.

I. Introduction

Programs using public funds for children to attend private schools of their choice are on the rise in the United States: As of 2019, 25 states, plus the District of Columbia and Puerto Rico, had voucher or scholarship programs in place, many of them targeted to specific populations like students with disabilities or low-income students (EdChoice, 2019). Among the most controversial issues associated with private school choice programs involves what happens to the students remaining in public schools. On the one hand, private school choice programs could encourage public schools that might otherwise have been complacent to vigorously improve the education they offer in order to avoid losing "clients" to a more-affordable private schooling sector (Epple, Romano, & Urquiola, 2017; Urquiola, 2016). On the other hand, public school students could be harmed by private school choice programs if the programs drain resources from the public schools or if choice-induced sorting of students disadvantages those remaining in public schools (Epple, Romano & Urquiola, 2017). The theoretical predictions assume an established program, so it is important to know what happens to traditional public schools as school choice programs expand and mature. We might expect to see the most pronounced effects when a new school choice program is initially announced, as the programs may be most salient when they are new and receiving publicity. But we might also expect the effects of school choice programs to become more pronounced as the programs grow; as schools see a growing share of students opting into choice programs, they may feel more compelled to respond.

The weight of the U.S. evidence shows small but positive effects of the introduction of private school voucher programs on public school students' test scores (Chakrabarti, 2008; Egalite, 2016; Egalite & Wolf, 2016; Figlio & Hart, 2014; Figlio & Karbownik, 2016; Greene &

Winters, 2007; Hoxby, 2003; Rouse et al., 2013; see Urquiola, 2016; Epple, Romano & Urquiola, 2017; and Egalite & Wolf, 2016 for recent overviews of this literature). But these studies are generally limited to the very immediate short-run effects, when both the pros and cons of the choice program may be constrained due to the small number of initial participants. What happens when the private school choice program scales up and the number of students using private school vouchers grows to encompass a sizeable fraction of the overall student body? To date, with the exception of an informative but single-market school-level analysis from Milwaukee (Chakrabarti, 2008), we do not know much about whether scaling up private school choice programs helps or harms public schools.

One paper provides some evidence on the impacts of scale-up of public school choice (charter) programs. Gilraine, Petronijevis, & Singleton (2019) show that North Carolina students who experience an increase in exposure to charter competition due to new school approvals resulting from the removal of a charter cap saw improvements in math (but null effects in reading). While they see some positive effects, the degree of scale-up that they study is considerably smaller than the degree that we consider here. In their setting, competition increases in a relatively narrow window of two years, resulting in roughly a 25% increase in the potential degree of charter school competition. Our longer time scale allows us to consider a program that has grown almost seven-fold from its original size and currently serves a participant population that is nearly 4% the size of the K-12 student population in Florida.[1] Moreover, our paper considers a different sector providing competition (private vs. public charter), explores

[1] Specifically, 108,098 students participated in the program as of 2017-18, compared to a K-12 public school student population of nearly 2.78 million students (Florida Department of Education, 2018; Florida Department of Education, n.d.a)

both cognitive and behavioral outcomes, provides extensive heterogeneity analyses, and is set in a different state.[2]

In this paper, we make use of extraordinary child-level data that matches birth records to school records, and employ student fixed effects to evaluate a statewide school voucher program, the Florida Tax Credit Scholarship Program, that grew over the course of about a decade from less than one percent to roughly four percent of the state's student body participating. We exploit differences in the initial competitive landscape faced by different schools – using five separate measures of voucher competition introduced by Figlio and Hart (2014) – as well as differences over time in the expansion and rollout of the voucher program, to determine whether students attending public schools that face increased exposure to private school choice as a result of a statewide voucher program experience educational (test scores) and behavioral (absenteeism and suspensions) benefits or losses.

We find evidence that as public schools are more exposed to private school choice, their students experience increasing benefits as the program scales up. In particular, higher levels of private school choice exposure are associated with lower rates of suspensions and absences, and with higher standardized test scores in reading and in math. These results are not uniform: We carry out an extensive heterogeneity analysis facilitated by matched birth and school records from Florida, and find that the public school students most positively affected by increased exposure to private school choice are comparatively low-socio-economic status students (those with lower family incomes and lower maternal education levels). Nonetheless, we also observe statistically significant but smaller gains for higher-SES students who are unlikely to be targeted

[2] Two other recent papers consider scale up of charter programs (Ridley and Terrier, 2018; Cohodes et al., 2019), however, these papers focus on even smaller expansions in Massachusetts and only the former is concerned with effects on students remaining in traditional public schools. They find small positive effects on test scores but do not explore behavioral outcomes.

by vouchers themselves. In terms of potential mechanisms, we are able to rule out alternative explanations related to changing composition of students remaining in the public schools and effects on the resources that public schools have. Thus, in our view, increase in competitive pressure resulting from increased voucher utilization is the more plausible channel for the estimated gains in test scores and behavior.

II. Florida Context

We focus on the competitive effects of the Florida Tax Credit (FTC) Scholarship Program as it scales up. Announced in spring 2001, the FTC program provides dollar-for-dollar tax credits to corporations that donate to non-profit Scholarship Funding Organizations (SFOs); the SFOs then use these contributions to offer scholarships to low-income students for use at private schools (Florida Department of Education, 2018). Until recently, and during the study period covered by this paper, in order to receive a scholarship the first time, students must either have spent the previous full year in a Florida public school or be entering kindergarten or first grade. In 2002-03, the first year of operation, the program spent $50 million to fund scholarships for 15,585 students, with a maximum value of $3,500 for each scholarship. Initially, eligibility was restricted to students with a family income no greater than 185 percent of the federal poverty line (see Florida Statute 220.187 (2001) and its subsequent amendments), or $47,637.50 for a family of four in 2019 dollars.

The program has expanded along several dimensions since 2002-03, its first year of operation. Table 1 charts the expansion of the program in terms of the designated funds for the program, realized spending, the number of participating private schools, the number of students enrolled, and the maximum scholarship level available. This growth may partially reflect a loosening of the income-based restrictions. In 2006 school-year, the program introduced a rule

allowing students using the scholarships to continue in the program if their family income did not exceed 200% of the federal poverty line. In 2010, this continuing-student eligibility threshold was raised to 230% of the federal poverty line. In 2016, the eligibility restrictions were changed to allow partial scholarships for entering students with incomes between 185%-260% of the federal poverty line (Table 1). By the 2017-18 school-year, the program cost roughly $640 million and awarded scholarships to 108,098 students (Florida Department of Education, 2018), or about 4 percent of all K-12 students in Florida.

II. Methods

A. Data and Sample

We draw upon data provided by the Florida Department of Education and Florida Department of Health. The Department of Education collects data on all students, including test scores, absences, and suspension data for students in grades PK-12. The Florida agencies merged these data to birth records for children born in Florida between 1992 and 2002, providing measures of families' socioeconomic status at birth, as well as neonatal outcomes such as birth weight. Because we also received the same data on the set of children born in Florida but never attending Florida public schools, we can characterize selection into our sample. We measure a public school's competitive landscape of nearby private schools based on files maintained by the Florida Department of Education, which provide locational data (latitude and longitude, as well as addresses) for public and private schools as well as the grades that each school serves (Figlio & Hart, 2014).

Our sample is limited in two key ways. First, we focus on outcomes for students in grades 3-8, because test scores serve as one of our main outcomes and they are most consistently

available for this set of grades. The second is that due to data availability and in order to have complete coverage of the rich set of measures provided by the birth records data, we restrict our sample only to those students with Florida birth certificates (i.e., those students born in Florida). Roughly 81 percent of children represented in Florida birth records are ultimately observed in the Florida public school data, a match rate that tracks closely with the share of Florida-born students who appear in Florida public schools according to the American Community Surveys (Figlio, Guryan, Karbownik, & Roth, 2014).[3] Records of children who started in a public Florida kindergarten but left the public school system prior to the start of testing in the third grade or had missing test score information in all years accounted for 14.8 percent and 0.8 percent of the remaining matched sample, respectively.[4] Additionally, 0.8 percent of the matched sample was excluded from testing because of severe disabilities. This suggests that our data provide good coverage of the overall universe of students affected by the competitive pressures from the school voucher program.

Overall, our main analytic sample includes student data for roughly 1.2 million unique students in the 2002-03 to 2016-17 academic years, although we use several additional prior years of data to characterize the initial schools for students in earlier cohorts as well. When we refer to academic years in data for the remainder of the paper, we will refer to spring of the academic year when the testing takes place.

[3] It is noteworthy that the voucher program's scale-up affected who shows up in public schools to begin with. As we discuss in Appendix A, in areas with greater competition, we see a diminishing share of students coming from lower-income families enrolling in public schools, consistent with the means-testing criteria for program eligibility.
[4] Leaving the public school system between kindergarten and the commencement of testing in grade 3 is not consistently correlated – in terms of sign and statistical significance – with competitive pressures faced at entry into the school system, suggesting that attrition from the sample is not endogenous to exposure to competitive pressure.

B. Models

We estimate the effects of expansions of school choice programs by estimating within-student models of the following form:

$$Y_{isglt} = \beta Expansion_t \times Competition_{sl} + \theta_{il} + \delta_{gt} + \varepsilon_{isglt}$$

where Y_{isglt} captures an outcome measure for student *i* who entered the FLDOE data in grade one school *s*, observed in grade *g* corresponding to school level (elementary or middle school) *l* in year *t*, θ_{il} is a student- by-school level fixed effect that allows separate within-student effects for elementary school (grades 3-5) and middle school (grades 6-8), but constrains school effects to be determined by schools that we anticipate students attending given their first grade placement.[5,6] The term δ_{gt} is a grade-by-year fixed effect. We include robust standard errors (ε_{isglt}) clustered by student's first-grade school.

The coefficient of interest is β, which estimates the interaction between a measure, $Expansion_t$, that captures the degree of program use statewide in year *t*, and a measure, $Competition_{sl}$, that captures whether each student's school is expected to face an above-median or below-median degree of competitive pressure, based on the pre-program competitive landscape. Importantly, for reasons we describe below in the Competition Measures section, the competitive pressure measures that we expect each student's school to face are projected based

[5] We focus on first grade rather than kindergarten because first grade is the first mandated grade of attendance in Florida. In practice, there is extremely high correspondence between kindergarten school attendance and first grade school attendance observed in the Florida data. As explained below we anchor each student to their grade one school and then rely on empirically observed flows of students between elementary and middle schools. Therefore, a school effect itself is not identified in this equation given the individual student fixed effects.

[6] Since our model includes individual fixed effects, it is problematic to further include lagged test scores in this estimation, which could presumably account for dynamic responses of cognitive skills to competitive pressures, because coefficient on β will be inconsistently estimated (Nickell, 1981). Nonetheless, since our sample sizes are very large, we have also estimated models with once-lagged test scores as control variable. This analysis produces, if anything, more positive estimates in the range of 0.2 to 1.0 percent of a standard deviation as compared to our baseline results of 0.3 to 0.7 percent of a standard deviation per 10 percent increase in the program size. Thus, we conclude that our results are robust to this specification check.

on the school that each child *initially attends* in first grade, rather than the actual school attended in any given grade, and we project these different measures based on whether the child is in elementary or middle school. This helps us avoid identifying off of changes in competitive pressure generated by endogenous moves by students. The competitive pressure measures are further based on pre-program competitive landscapes, as explored in Figlio & Hart (2014), rather than the actual competitive landscapes after the voucher program is introduced because the latter may be endogenous to public school quality. Given this design, the student-level fixed effect implicitly holds the initial level of pre-program competition constant within each student-by-school level cell. This means that the effect for the interaction term is identified off of program expansion rather than by any movement of the students between schools, or off the introduction of new private schools in response to the incentives introduced by the voucher program. Thus, the coefficient of interest describes whether expansion matters more for schools with higher initial degrees of competitive pressure than for schools with relatively little initial competitive pressure. We multiply our estimates by 100 to ease interpretation of effect sizes in our figures and tables.

C. Measures

Outcomes. Our main cognitive outcomes rely on standardized measures of math and reading scores for 3rd-8th grade students on Florida's state tests. We standardize each test within year and grade[7] using our empirical sample of Florida-born students to maintain consistency across years, but the results are robust to using measures available for a subset of our sample

[7] This is especially important due to several changes in the structure of the tests over the period covered by our analysis. The Florida Comprehensive Assessment Test (FCAT) was replaced by an updated version (FCAT 2.0) in 2010-11, and then by the Florida Standards Assessments (FSA) in 2014-15 (Florida Department of Education, n.d.b).

years that are standardized on the whole-state population. We use school years 2002-03 to 2013-14 and 2002-03 to 2016-17 for mathematics and reading, respectively. This discordance comes from the fact that, in mathematics, more advanced students were able to exercise more choice about which assessments to take starting in school year 2014-15; for instance, students taking Algebra I in eighth grade could take an algebra-specific examination rather than an examination on 8th grade math generally. We therefore exclude the years with less consistency in tests from our analysis. We also construct the measure of averaged mathematics and reading test scores for each student for school years 2002-03 to 2013-14.[8]

Another major contribution of our paper is to explore effects of competitive pressure on a novel set of behavioral outcomes: likelihood of being suspended and absence rates. Our suspension measure is an indicator variable for whether a student has ever been suspended in a given school year, while our absence measure captures the share of days that a student is reported absent net of days they are suspended. Thus, the former can be thought of as an indicator for more serious disciplinary problems while the latter is a measure of truancy. We observe suspension and absenteeism outcomes through the 2011-12 school-year.

Competition. Following Figlio & Hart (2014), we use five measures to capture the degree of competitive pressure that each school is likely to face. The "Density" measure captures the number of private schools serving the same grade range of students (i.e., elementary or middle school grades) within a five-mile radius of each public school. The "Distance" measure captures the distance between each public school and the nearest private competitor serving the same grade range; this measure is multiplied by -1 so that a positive sign on the

[8] In the main regression for years 2002-03 to 2013-14, 98.8 percent of observations have both math and reading scores, 0.3 percent have only math, and 0.9 percent have only reading. By comparison in years 2014-15 to 2016-17, 74.7 percent have both scores, 1.7 percent have only math, and 23.6 percent have only reading.

measure will indicate greater competitive pressure. The "Diversity" measure captures the number of different religious denominational categories represented among the private schools within a five-mile radius of each public school; we group each school into one of ten denominational categories (including non-religious) for this measure.[9] The "Slots" measure captures the number of private school students served in the same grade range within a five-mile radius, standardized by the number of grades served. The "Churches" measure captures the number of houses of worship in a five-mile radius. This measure captures the underlying religiosity of the community, which may be associated with demand for private religious education, as well as the possibility that private schools may co-locate in the buildings that serve as houses of worship (Figlio & Hart, 2014). This measure is related to others commonly used to capture demand for religious education in the literature on private school competition, such as the share of a population that is Catholic (Hoxby, 1994; Dee, 1998; Jepsen, 2002) or the density of Catholic churches in a locality (Jepsen, 2002), but captures religiosity across a greater number of faith traditions.

Because presenting all five measures is unwieldy for the purposes of robustness checks and heterogeneity analyses, we also construct a single composite "Competitive Pressure Index" measure based on a principal components analysis of the five aforementioned measures. The principal component analysis produced a single component with an eigenvalue greater than 1; the loadings for this component were used to generate the Competitive Pressure Index score for each school. The component loadings generated by the principal components analysis for the first two components are documented in Appendix Table A2.

We make two important decisions in assigning competitive pressure measures to schools. The first addresses the concern that the competitive pressure faced by a school in any given year

[9] The ten categories are Non-religious, Catholic, Protestant, Baptist, Evangelical, Non-denominational, Jewish, Islamic, Christian Other, and Religious Other.

may be endogenous to perceived school quality; for instance, private schools may be tempted to enter into markets where the public alternatives are of poorer quality (Arsen & Ni, 2008). In that case, competitive pressure would be conflated with other unobserved factors plausibly correlated with student outcomes. To avoid this problem, we measure the competitive pressure that each elementary and middle school in Florida faced using the competitive landscape in place in 2000, the last year before the voucher program was announced.

The second decision addresses the concern that students may move between public schools based on their perception of school quality. If these factors are correlated with competitive landscape of schools, our within-student estimates would be estimated partially off of the effects of these potentially endogenous school switches. To avoid this, we assign students to elementary and middle schools based on the school that they attend in first grade. This treats students as if they were in the same elementary school they entered in first grade, and thus abstracts from any potentially endogenous moves. For students' elementary-school years, the competitive pressure measures therefore capture the pre-policy competitive landscape of students' grade 1 schools. For the middle school years, we create a weighted average of the competitive landscapes that we would anticipate students to face based on 1) the flow of students empirically observed in our data from each elementary school (grade 1) to each middle school (grade 6) in Florida, and 2) the pre-program competitive landscape of Florida middle schools. Specifically, for each student attending a given grade-1 school, we observe the middle school that they actually attend, and we capture the pre-policy degree of competition faced by that middle school. We then weight these measures with empirical flows between elementary and middle schools to obtain the expected middle-school competition for each student based on the grade 1 school they attend. Importantly, in our estimating equation the fixed effect θ contains the

interaction of an individual fixed effect with an indicator for whether the child is in a middle school grade, so that the competitive pressure that we expect children to face as they progress from elementary to middle school is allowed to vary with expansion.[10]

Because our main interest is in whether effects from program expansion are more marked in schools that face greater competitive pressure, we dichotomize each competition measure to ease interpretation of the interaction terms. Thus, for each variable, the competitive pressure indicator captures whether the student's projected school is above or below the median on each competitive pressure measure. This median split is conducted at school- rather than individual-level and separately for grades 1 to 5 and 6 to 8.

Program expansion. Our main measure of program expansion captures the logged number of students participating in the FTC Scholarship program in a given year. The interaction of the logged expansion measure and the median-split measure of competitive pressure can therefore be interpreted as the relative effect of a one percent expansion in the number of students served in schools initially facing an above-median degree of competitive pressure, compared to the effects of expansion in locations with lower competitive pressure.

Student characteristics. We have a variety of student characteristics from birth records. In particular, we capture student sex, mother's race (White, African-American, other), mother's ethnicity (Hispanic vs. non-Hispanic), whether the child's mother was born in the US, mother's marital status at the time of birth, mother's years of education, whether the birth was paid for by Medicaid, maternal age at the time of birth, and information on the month and year of birth. These characteristics are time-invariant and are therefore captured by student fixed effects in our

[10] In robustness checks we also estimate models where we assign the pre-policy degree of competition faced by grade 1 and grade 6 schools separately, and likewise estimate models with student-by-school level fixed effects.

main estimating equation; however, we use some of them to provide extensive heterogeneity analysis to further the understanding of mechanisms at play.

In school records we also observe information on students' free and reduced price lunch status (FRPL), which we use as another stratifying characteristic in our heterogeneity tests. This measure varies within student across years but we focus our analysis on two groups of students: those that were never on free and reduced price lunch through their public schooling career; and those that were ever designated to be eligible for free or reduced price lunch.

Table 2 provides descriptive statistics for the full population of Florida births (column 1) and our empirical sample (column 2). The comparison between these two columns makes it clear that the set of children remaining in Florida to attend public school is negatively selected in terms of socioeconomic characteristics compared with all children born in Florida. We are more likely to observe children whose mothers are high school dropouts (24.9 percent vs. 20.9 percent), and less likely to observe children whose mothers are college graduates (14.7 percent vs. 20.2 percent). We are also more likely to observe African-American children (23.3 percent vs. 19.4 percent). At the same time, ethnicity, immigrant origin and maternal age at birth are comparable in these two samples.

In subsequent columns of Table 2 we investigate whether characteristics of students differ based on the degree of the pre-policy competitive pressure faced. Columns 3 and 4 provide statistics for children whose schools are above- or below-median in competitive pressure based on our Competitive Pressure Index. It appears that locations facing less competitive pressure prior to program's introduction have more White students (68.0 percent versus 37.3 percent for schools with above median competition) and have markers suggestive of higher socioeconomic status (66.6 percent ever on free or reduced price lunch vs. 75.8 percent; 65.1

percent of parents married at the time of birth vs. 54.6 percent). Interestingly, however, composition of parental education is relatively similar across these locations. Panel B of this table presents mean values of our five measures of competitive pressure and the combined Competitive Pressure Index based on principal components analysis while panel C shows the five outcome variables. Descriptively, it appears that children attending schools with above median competitive pressure have, on average, worse outcomes than children attending schools with below median competitive pressure. For example, mathematics and reading test scores are 9.1 and 10.8 percent of a standard deviation lower in the former as compared to the latter sample, respectively. These patterns may be because lower-SES families (who tend to have lower test scores on average) are more likely than higher-SES households to live in more densely-populated urban areas (which tend to have higher degrees of competition). Regardless of the explanation, these cross-sectional differences underscore the importance of our empirical strategy that identifies competitive pressure effects based on the roll-out of the voucher program, controlling for student fixed effects.

III. Results

A. Main Results

We find consistent evidence that as the voucher program scales-up, students in areas with more pre-program competitive pressure see a significantly greater improvement in outcomes than do students in areas with lighter pre-program competitive pressure (Table 3). While the magnitude of the coefficients varies across competition pressure measures, this pattern is consistent across all five underlying measures of competitive pressure (Panels A-E) for four of the five outcomes that we measure: averaged math and reading (Column 1), reading (Column 3),

suspension rates (Column 4), and absence rates (Column 5). When we combine all the measures into our single Competitive Pressure Index measure (Panel F), we likewise find statistically significant increases in test scores and reductions in behavioral problems. Recall that the outcomes are multiplied by 100, and thus, our results for the Diversity measure in Panel A suggest that a 10 percent increase in the number of students participating in the voucher program is associated with a 0.4 percent of a standard deviation greater improvement in combined math/reading scores for students in schools with above-median density of private competitors, compared to improvements for students schools facing lower degrees of competitive pressure. This effect is larger for reading (about 0.7 percent of a standard deviation) as compared to math (about 0.2 percent of standard deviation). We also find reductions in both suspensions and absences, indicating that behavioral outcomes in the schools facing higher competition likewise improve. The former declines by 0.4 percent and the latter by 0.5 percent compared to their means in specifications using the Diversity measure. While the Density, Diversity, Churches, and Slots measures suggest that expanded competitive pressure is significantly and positively associated with math achievement, results are non-significant when the Distance measure is used (Column 2). Nonetheless, large negative effects on math are unlikely: Based on a 95 percent confidence interval, we can rule out negative effects of up to 0.2 percent of a standard deviation.

The Competitive Pressure Index estimates presented in panel F – the estimates using our preferred measure – imply that a 10 percent increase in the number of students participating in the voucher program, in schools with above-median as compared to those with below-median baseline competitive pressure, increases math test scores by 0.3 percent of a standard deviation. Reading gains are larger at over 0.7 percent of a standard deviation. At the same time, suspensions decline by 0.9 percent relative to the baseline rate of 13.7% of students being

suspended, while absence rates decline by 0.6 percent relative to the baseline absence rate of 5.0%. Since this measure combines all the competitive inputs into a single index, for transparency and clarity we present subsequent heterogeneity and robustness analyses using the Competitive Pressure Index measure, but we make results for all other measures available in the Appendix as well.

Our results are similar in pattern if we simply look at the changes in effects of the initial level of competitive pressure over time (as the program was expanding) rather than as a function of a measure of program usage. In other words, the key terms here are interactions of initial competitive pressure level-by-year rather than interactions of initial competitive pressure level-by-logged number of participants. Figure 1 shows the year-by-year marginal effects for being located in a market with more baseline competitive pressure for each outcome using our preferred measure of competition pressure– the Competitive Pressure Index. In this analysis, we standardize all five outcome variables to have mean zero and standard deviation of 100 in their respective empirical samples, thus making the direct comparison of magnitudes across multiple outcomes feasible.[11]

The graph suggests that schools located in markets with more competitive pressure saw a roughly 14 percent of a standard deviation greater increase in combined mathematics and reading scores by 2014 relative to schools in markets with less-competitive pressure. These gains are initially modestly larger for mathematics but this pattern changes in favor of reading by 2013. Given that by 2014, the program had expanded by nearly 300 percent compared to its original size, this coefficient is strongly consistent with our main table.[12] With respect to behavioral

[11] This figure standardizes each outcome across the entire timespan studied, but versions of the graph that standardize each outcome by year produce similar results.

[12] More specifically, given our estimate of 5.111 (on a logged competitive pressure measure where the combined math and reading test score dependent variable was multiplied by 100 to show significant digits), the 284 percent

outcomes, improvements emerge later in the time period. Students attending schools in markets with more competitive pressure saw consistently greater reductions in suspensions (absences) relative to peers in schools facing less competitive pressure starting in 2006 (2009). We present corresponding graphs for each competitive pressure measure and each outcome (unstandardized) separately in the Appendix (Figures A1 to A3). The bottom panel of Figure 1 further illustrates estimates for families ever versus never on free or reduced price lunch, and shows that our effects are more pronounced for the former set. We come back to this point in our heterogeneity analysis below.

B. Benchmarking Effect Sizes

One benchmark to contextualize the size of our causal estimates is the extent to which expansion of the voucher program is associated with closing the gap between schools located in high-competitive pressure and low-competitive pressure areas. Our descriptive statistics in Table 2 suggest that students in schools facing higher competitive pressures tend to have poorer outcomes across all measures except for absences, which are similar in both groups. Those poorer outcomes include lower math scores (gap of 9.1 percent of a standard deviation), reading scores (gap of 10.8 percent of a standard deviation), combined scores (gap of 10.0 percent of a standard deviation), and higher likelihood of suspensions (gap of 1.4 percentage points or 4.1 percent of a standard deviation). Given these figures and the effect sizes presented above, a ten percent increase in the size of a voucher program would be expected to close between 2.9 and 6.8 percent of the test score gaps, and 9.2 percent of the gap in suspensions. The closure of these gaps is especially meaningful because students with poorer average academic outcomes

increase in the size of the program would be associated with an 14.52 percent of a standard deviation increase in the combined math and reading test scores; this is very close to the estimate of 14.46 percent of a standard deviation presented for 2014 in Figure 1.

(including African-American students, Latino students, and students using free and reduced price lunch) tend to be over-represented in schools facing higher degrees of competitive pressure.

However, it is worth highlighting that, as illustrated in Table 1, the program has expanded by much more than ten percent; the number of scholarship users was nearly seven times higher in 2017-18 than in 2002-03—and thus the realized gains should actually be much larger. If we use our preferred Competitive Pressure Index estimates from Table 3 (Panel F), we would expect a more conservative doubling of the program size to result in a greater differential improvement of 5.1 percent of a standard deviation in combined math/reading scores in areas with high competitive pressure vs. in areas with low competitive pressure, and a 1.3 percentage point greater decline in suspension rates (9.4 percent relative to the sample mean); these effect sizes represent a meaningful change in the gaps in outcomes between schools facing higher and lower degrees of competitive pressure.

A potentially more intuitive set of benchmarks may also help contextualize the size of our reduced-form results. For example, when considering combined math and reading, the mean test score difference between students ever observed on free or reduced price lunch and those never observed with that status is 73 percent of a standard deviation. Likewise, the test score gap between children with high school graduate and college graduate mothers is over 60 percent of a standard deviation, while the gap between African-American and White children is 62 percent of a standard deviation. The same gaps in likelihood of being suspended are 12.2, 8.5, and 14.4 percentage points, respectively. Thus, it is worth noting that the program effects, even when considering maximum observed scale-up, are still relatively modest compared to these larger, long-standing gaps.

To further asses the magnitude of our estimates, it is also helpful to compare them to other estimates in the education and human capital formation literature, and in particular to those obtained using data from Florida. For instance, Figlio and Hart (2014), who looked at the introduction of the program, found that a one standard deviation increase in the pre-program competition pressure predicted a differential improvement in test scores of 1.5 to 2.7 percent of a standard deviation in combined math and reading standardized scores. We can most directly compare the present results to those if we draw on specifications that use continuous (rather than median-split) measures of competitive pressures, which we show in robustness check presented in Table 5, panel B. Those figures suggest that our conservative quantification, assuming doubling (tripling) of the program, would result in a roughly 1.5 (2.9) percent of a standard deviation increase in combined math and reading standardized scores. Thus, a doubling or tripling of the program size is required to increase the salience of competition pressure—with attendant benefits to public school student test performance – to the same degree as the initial introduction of competitive pressures.

A doubling of the program based on the preferred Competitive Pressure Index estimates of 5.1 percent of a standard deviation effect from Table 3 is also comparable to or larger, depending on the exact outcome and specification, when contrasted with effect sizes from charter expansion studied by Ridley & Terrier (2018) and Gilraine, Petronijevis, & Singleton (2019). It is further about the same as the effect of a 10-percent increase in birth weight (Figlio et al., 2014), a quarter of the size of the effect of school entry cutoff on cognitive development (Dhuey et al., 2019), and 80% of the birth order gap in reading scores (Breining et al., 2020). More generally, it is equivalent to roughly 10-percent of the effect of child care subsidies on children's GPA (Black et al., 2014). Note, however, that all these papers estimate total effects of

the treatment while our estimates represent gains in high-competition relative to low-competition areas (thus providing lower bound on the total effect), where competitiveness of an area is measured prior to program's introduction. Nonetheless, we feel that these comparisons should help readers in understanding the magnitude of our coefficients.

We view benchmarking the effect sizes off of a doubling of the program (i.e., a relative measure of increase in size) as useful because no other program has expanded to the same extent in absolute terms; however, as documented in Figure 1 the effective gains of Florida students between first year of program operation (2002-2003 school year) and end of our data span are much larger at 12 and 17 percent of a standard deviation in math (school year 2013-2014) and reading (school year 2016-2017), respectively. Importantly, it appears that even after almost seven-fold expansion and coverage of about 4 percent of K-12 population the test score gains do not decline.

C. Heterogeneity

We next address questions of whether different types of students differentially benefit from increased competitive pressure, running our reduced-form regressions separately for each subsample of students in turn. We present the results using the composite Competitive Pressure Index measure for all five outcome variables (Table 4), and show the results for each underlying competitive pressure measure separately in the Appendix (Tables A3 to A7).

Results are generally consistent in pattern across all subgroups; however, the exact magnitudes and statistical significance vary somewhat. Lower socioeconomic status students– measured by free or reduced price lunch designation– see larger effects across all outcomes. Within test score outcomes, these differences are more pronounced for reading than for math.

When we compare how the effects unfold over time as the competitive pressure increases from year-to-year for individuals ever and never on free and reduced price lunch (panels B and C of Figure 1), it becomes clear that irrespective of the exact outcome, the gains are larger in former group than in the latter. Given that more affluent children should never have been eligible for the program, however, the fact that the expansion of the program was nonetheless associated with improvements for this group of children in more competitive landscapes suggests that the benefits of competitive pressure are diffuse and extend, albeit to a lesser degree, to children that the public schools are not at risk of losing to private schools due to the voucher program. Importantly, the results are also similar when we stratify the sample by maternal education: Effects are consistently smaller for children of mothers who graduated from college than for children whose mothers attained less education.

Similarly, we also can divide families into deciles of socio-economic status, using a measure introduced by Autor et al. (2019) using the same Florida data that we use here. The socioeconomic status composite index is created through a principal components analysis, similar to our competitive pressure index. Specifically, the principal components analysis generates factor loadings based on mother's marital status, age, and years of education at birth as well as an indicator for whether the birth was Medicaid-funded and median zip code-level neighborhood income at the time of birth; and uses these to construct the composite socioeconomic status gradient index. We then separate the sample by SES deciles to observe competitive effects across the SES gradient.

Appendix Figure A4 presents coefficients associated with increasing competitive pressure on students by SES decile. It shows that these effects are strongest for those families in the bottom six deciles, but that expanded competitive pressure is associated with benefits for all

families except for the very top SES decile. Taken together, these patterns of results suggest that voucher expansion may work partly through particularly stimulating competition near lower-SES schools, competition-induced additional focus on lower-SES students (regardless of school SES), or a combination of both.

We also observe differences in magnitudes across racial and ethnic groups as well as by immigration status. First, test score gains are very similar for African-American and White children. The former group, however, does not experience significant behavioral benefits in terms of absences or suspensions. Second, Hispanics experience larger gains in reading compared to the other two racial/ethnic groups but smaller and statistically insignificant increases in mathematics. It also appears that increased competitive pressure is particularly beneficial for Hispanic students in terms of reductions in suspensions. Third, students with foreign born mothers see a pattern of results comparable to that of Hispanic students. This is not surprising given that Latino children are disproportionately likely to have non-native born mothers in our sample (68 percent of Hispanic children in our sample has mother born outside of the U.S., compared to only 9 percent for non-Hispanic children).

Some of these results are somewhat sensitive to measures. For instance, the Distance measure generally yields the weakest effects on student outcomes (Appendix Table A5). Consistent with this weaker relationship, groups that experience more modest effects using our preferred Competitive Pressure Index measure often show non-significant benefits to expanded competitive pressure when we use the Distance measure. For instance, White students only experience significant benefits to absence rates using that measure; their estimated effects on other outcomes are of the expected sign but non-significant. Nonetheless, the pattern of results is mostly similar in terms of direction—if not in magnitude or significance—using this measure.

One exception from this rule are estimates for mathematics where six out of ten estimates using the Distance measure have negative sign, and one – for children of foreign-born mothers – is statistically significant. Interestingly, for these same children we estimate positive gains in reading that are comparable to those in panel C of Table 3. Overall, aside from the Distance measure, most other measures show similar results in terms of pattern and significance to the preferred results presented in Table 4.

D. Robustness

While our results vary somewhat across outcomes and different sociodemographic groups, they are largely robust to different modeling decisions. We present a set of robustness checks, using our preferred competitive pressure measure (the Competitive Pressure Index) in Table 5 while robustness tests for each competitive pressure measure separately are presented in the Appendix (Tables A8 to A12). Our main results from Table 3 are reproduced in Panel A for ease of comparisons.

A first set of tests looks at whether our results are sensitive to different constructions of the competitive pressures and expansion measures that underlie the interaction term. Panel B uses the underlying, continuous Competitive Pressure Index measure of pre-program competitive pressures in our interaction terms rather than the median-split term. The pattern of results remains the same, although consistent with the fact that this measure has a different underlying distribution, the magnitudes of the coefficients are predictably different compared to results in panel A. Panel C uses an alternate measure of program expansion, substituting a logged measure of program funding in place of the logged measure of student participation. Results are strongly consistent with our main analysis. As further documented in panel D, our results are likewise

unchanged when we assign the middle-school pre-policy competition measures based on the actual grade 6 (middle) schools initially attended by each student, thus potentially allowing for endogenous selection into middle school based on its quality. This suggests that our conservative decision to anchor each student to the competition that we project based on grade 1-school (proxied by empirically-observed transitions into middle school) is not consequential for the results.

Finally, our main results use all available data to generate flows of students between elementary and middle schools that we then use to generate the aforementioned expected competitive pressure measures faced by children in the latter set of institutions. This approach may, however, be subject to endogeneity if these flows are affected by program expansion differentially in areas with higher vs. lower baseline competition pressures. Thus, in panel E, we investigate if our results remain unchanged when we generate our expected competitive pressure measures for middle school-aged students using pre-policy announcement flows between elementary and middle schools. Since our data do not go back far enough to track children from grade 1 to grade 6 using only pre-announcement cohorts, we cannot execute this analysis based on first attended grade but rather utilize transitions between grades 5 and 6 for school years 1998/99 to 2000/01. As evidenced from Table 5 this refinement is inconsequential and, if anything, test scores coefficients increase in magnitude. Since this analysis necessarily treats grade five schools as if they were first observed as grade one schools we chose to include it as a robustness check rather than the main analysis.

The next set of tests in Table 5 checks for sensitivity of results to the exact samples used. Our earlier work (Figlio & Hart, 2014) used a sample restricted to students attending schools with at least one competitor within five miles. In panel F we impose a similar restriction, and we

reconstruct the median split to reflect the exclusion of the schools that do not have a competitor with 5 miles from the analysis. We find that the preferred results are robust to this sample change. We next test whether our results change if we limit all models to the set of years for which we have all outcomes available, ending our sample in 2011-12, after which we no longer observe suspension and absenteeism outcomes (panel G). Results are likewise robust to this sample limitation. Panel H uses the subsample of children for whom we observe student for six consecutive years, when a normal progression would have taken them from grade 3 to grade 8. Again, results are similar for this subsample.

The final two panels (I and J) address the concern that our test measures in the main specifications are standardized using the Florida-born sample rather than a statewide sample. This decision allowed us to use a greater set of years, because we only have data to standardize scores using the entire state population through the 2012/13 school-year. Panel I replicates our main specification using our sample-standardized outcome measure and dropping years after 2012/13 school-year where we lack population-level standardized measures. Panel J substitutes in the set of test score measures standardized using the whole state population. Again, these results are very similar to our primary findings. The stability of these results across different specifications gives us greater confidence in the reliability of our findings. Results are also generally stable across different competitive pressure measures (Appendix Tables A8-A12), although consistent with the generally weaker results overall, we see more sensitivity to results in the Distance measure or for mathematics when considering Churches measure.

A final concern for our identification may be that there are secular improvements over time that happen to be more pronounced in areas with high competition, but that are occurring regardless of the voucher policies. We refer readers to our earlier work on the introduction of the

FTC program (Figlio & Hart, 2014), where we show that trends in school performance prior to the introduction of the FTC program were unrelated to the baseline competition levels; only after the voucher program was introduced did schools in areas with higher baseline competition begin to improve disproportionately. Thus, we are confident that our results reflect effects of the scale up of the program itself differentially affecting schools in higher-baseline competition areas, rather than reflecting any prior differential trends.

IV. Alternative Explanations

So far, we have suggested that the results found in this paper are due to increased competitive pressure on public schools associated with voucher program expansion. However, there may be alternative explanations to these findings. For instance, voucher programs may change the composition of students remaining in the public schools, and these changes might be related to the degree of voucher competition individual schools face. Voucher programs also influence the resources that public schools have, and these resource effects might work in opposite directions. On the one hand, the voucher program reduces funding to school districts that lose state funding allocations for students attending private school. On the other hand, if the vouchers lead to fewer students per grade, class sizes might plausibly go down in the public schools. To the extent possible, we investigate these alternative explanations in this section.

A. Peer Composition

First, consider the possibility that our results are due to changes in school composition brought about by differing degrees of voucher competition. These composition changes could result in observed impacts through peer effects associated with who remains in the public

schools. For instance, if students who leave public schools to use the voucher program tend to be lower-achieving on average, then the loss of those peers to the private sector could leave behind an easier-to-educate core of students, and result in positive impacts on student learning. These compositional changes could produce benefits even if schools exert no more effort in response to the competitive pressure caused by vouchers.

To investigate this, we carry out an analysis to see whether schools facing increased competitive pressure have students remaining in the school who would have higher predicted test scores, all else equal, based solely on their background characteristics. Panels A1 to A3 of Table 6 and Figure 2 present the results of an analysis that is parallel to our main specification, but with two key changes. First, this analysis is conducted at the school level rather than the student level and accordingly uses school fixed effects rather than student fixed effects. Second, the dependent variable is the average predicted test score in each school, with the predicted values based solely on background characteristics of the students enrolled. That is, we regress individual-level test scores onto student background characteristics measured at birth (child's month and year of birth, sex, birth weight, gestational age, birth order, prenatal care start, abnormal conditions at birth and congenital anomalies; as well as mother's education, race, ethnicity, place of birth outside of U.S., state of birth other than Florida, health problems, age, marital status and Medicaid-paid birth), and use the resulting coefficients to predict test scores for each student. These predicted test scores are then collapsed to create a school-by-year level averages, which generate an indexed value of the average test score that would be predicted in the school and year, given only student background data. If we see that schools with more competition also have student cohorts with greater predicted scores enrolled over time, this

would provide evidence that changes in student composition, rather than any effort by the school, may explain the effects we documented above.

Indeed, it appears that when we pool all years together the estimated competition effects are all positive, however, they are statistically significant in only 7 out of 18 specifications and they are generally small in magnitude. In fact, the largest estimated coefficient, for reading using diversity measure, is more than 20 times smaller than the equivalent estimate for actual reading test scores in Table 3. The estimates for our preferred competition measure - Competitive Pressure Index – are even more negligible quantitatively at 18 to 35 times smaller than corresponding coefficients in Table 3. Furthermore, Figure 2 documents that schools facing greater competitive pressure under the expansion of the voucher program had student populations that were likely to have somewhat *lower* test scores in non-trivial number of years and specifications. This non-linearity could explain why many estimates in Table 6 are small, statistically insignificant, and close to zero.

The largest positive estimate in Figure 2 – corresponding to the Diversity measure for the average reading score of the school in 2017– is 0.25 percent of a standard deviation (b=2.51). This coefficient is six times smaller than the corresponding estimate of competition pressure measured by Diversity on student reading in 2017 (b=15.05) in Figure A1. The largest negative estimate in Figure 2 – corresponding to churches measure for mathematics in 2011 – is 0.20 percent of a standard deviation (b=-1.96). This coefficient is about 4.5 times smaller, in absolute value, than the corresponding positive estimate of competition pressure measured by Churches on student math in 2011 (b=8.67) in Figure A1.

Finally, it is worth noting that the non-linear pattern observed for predicted test scores is in stark contrast to persistently positive and, if anything, growing over time effects documented

in Figure A1 for actual test scores. This increase in test scores is further notable prior to school years 2015/16 and 2016/17 where we estimate positive and for some measures statistically significant coefficients in Figure 2. When it comes to small negative estimates documented in Figure 2, these would lead to underestimation of the effects of competition per se (net of composition effects), and thus, if anything, our results for mathematics where this phenomenon is more pronounced should be treated as lower bound. Overall, all this suggests that our results are unlikely to be driven primarily by changes in student composition associated with increased voucher utilization and competitive pressure effects, in our view, are the more plausible explanation for the estimated effects.

B. Resources

The voucher program could also have induced changes in resources received by affected schools. We lack reliable data on most measures of school resources, and candidate measures such as school-level measures of expenditures, when they exist, largely reflect either teacher experience levels or student participation in special education (Lankford, & Wyckoff, 1995). But arguably the most salient resource indicator – average class size in a school – is measurable in the Florida data. We therefore explore whether our estimates of the competitive effects of voucher scale-up are potentially due to changes in class size associated with increases in competitive pressure. These may occur mechanically, to the extent that voucher programs draw students away from the public schools they would have attended, or may be the result of strategic decisions by principals to make the school more attractive to students and parents. We draw on class size archives from 2006-2007 through 2016-2017 posted by the Florida Department of Education, which report the average class size for each school separately for students in grades

PK-3, grades 4-8, and grades 9-12.[13] For each school and year, we calculate the weighted average class size at the school level based on the grade range served by each school. This class size variable, varying at school-by-year level, is then used as an outcome in regressions akin to those used to produce our peer effect estimates.

Table 6, panels B1 to B3, shows the results of this exercise. Several patterns of note emerge. Schools facing landscapes with more initial competitive pressure did have somewhat smaller class sizes as the program expanded, though this relationship is only statistically distinct from zero in one measure of competition (Diversity). If we look at the yearly differential effects on class size based on initial competitive landscape, we see a more pronounced effects on class size reduction (Figure 3). Coefficients are negative and statistically significant for most measures of competitive pressure in most years. But, unlike Figures A1-A3, which show an increasingly positive (negative) relationship between initial competitive pressure and student testing (suspension and absences) outcomes as the program expands, the class size effects remain fairly static over time.

Moreover, the magnitudes of the coefficients are too small to realistically explain away much of our main cognitive and behavioral effects. Consider the largest estimate of effects of competition on class size – using the Diversity measure. Here, we estimate a 0.0042 student reduction in class size per 10 percent increase in the program. To contextualize the expected effects on test scores of a reduction of this magnitude, we make use of the 22 percent of a standard deviation increase in test scores associated with a roughly seven-student reduction in class size effect implied by the Tennessee STAR experiment (Krueger, 1999); this estimated

[13] Ideally, we would observe class size information starting in 2002/2003 school year, however, data for these earlier cohorts are not available online. At the same time, we cannot construct class size using our registry data because these are limited only to individuals born in the state of Florida.

effect is similar in magnitude to those found by Angrist & Lavy (1999), Chetty et al. (2011), Fredriksson et al. (2013), and Lindahl (2005) in related studies. If we assume the same proportionate effect would apply to competition-induced class size decrease, a 0.0042 reduction in class size would imply expected test score improvements scarcely different from 0 (0.0042*22/7=0.013 percent of a standard deviation). This would account for only about 3% of the effect of competition on test scores that we estimated in Table 3, where our coefficient on the Diversity Measure implied a 0.42 percent of a standard deviation increase predicted for a 10 percent program expansion. For the other measures of competition, the explained share is even smaller.[14] Given that the range of results in the extant literature (e.g., Jepsen & Rivkin, 2009; see Chingos, 2013 for a thorough review of the class size literature) estimate that class size reductions of about 10 students produce improvements in test scores of between 0.05 and 0.22 standard deviations, and other papers (e.g., Leuven & Løkken, forthcoming) sometimes find very small class size effects on other meaningful outcomes, the share of our estimated effects that can be explained by class size reductions may be even smaller still. Thus, class size changes would explain only a relatively small portion of the observed effects of program expansion presented in Figures 1 and A1. While class size may be contributing to the observed positive effects of competitive pressure on cognitive and behavioral outcomes, it seems unlikely to be a major driver of the findings.

[14] Even the upper bound of the 95 percent confidence interval in the Diversity measure would imply that less than 6 percent of the estimated competitive effects could be due to class size reductions. For the other measures of competition, the upper bound estimate, based on the 95 percent confidence interval of the estimates of competition expansion on class sizes, ranges from 2.9 to 4.7 percent of the estimated competitive effects.

V. Conclusions

School choice programs have been growing in the United States and worldwide over the past two decades, and thus there is considerable interest in how these policies affect students remaining in public schools. Although we now have relatively comprehensive knowledge on the immediate short-run effects stemming from the introduction of such programs, the evidence on the effects of these programs as they scale up is virtually non-existent. Here, we investigate this question using data from the state of Florida where, over the course of our sample period, the voucher program participation increased nearly seven-fold.

We find consistent evidence that as the program grows in size, students in public schools that faced higher competitive pressure levels see greater gains from the program expansion than do those in locations with less competitive pressure. Importantly, we find that these positive externalities extend to behavioral outcomes— absenteeism and suspensions—that have not been well-explored in prior literature on school choice from either voucher or charter programs. Our preferred competition measure, the Competitive Pressure Index, produces estimates implying that a 10 percent increase in the number of students participating in the voucher program increases test scores by 0.3 to 0.7 percent of a standard deviation and reduces behavioral problems by 0.6 to 0.9 percent. We show that these results are robust to alternative plausible ways of measuring competition and expansion, as well as to different modeling choices. Our results are also consistent with past work showing modest benefits to the initial introduction of voucher programs (e.g., Hoxby, 2003; Figlio & Hart, 2014; Egalite, 2016; Egalite & Wolf, 2016; Figlio & Karbownik, 2016), while extending upon these findings to show the persistence and growth of these positive effects as the program scaled up.

Finally, we find that public school students who are most positively affected come from comparatively lower socioeconomic background, which is the set of students that schools should be most concerned about losing under the Florida Tax Credit Scholarship program. However, smaller effects remain statistically significant – in most cases – even for students who are very unlikely to be targeted by vouchers themselves, suggesting that benefits may come partially through generalized school improvements rather than through improvements targeted solely at voucher-eligible students.

References

Angrist, J.D., & Lavy, V. (1999). Using Maimonides' Rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics,* 114(2), 533-575.

Arsen, D., & Ni, Y. (2008). *The competitive effect of school choice policies on traditional public schools.* Arizona State University/University of Colorado at Boulder. Tempe, AZ: Education Public Interest Center.

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman (2019). Family disadvantage and the gender gap in behavioral and educational outcomes. *American Economic Journal: Applied Economics,* 11(3), 338-381.

Black, S.E., Devereux, P.J., Løken, K.V., & Salvanes, K.G. (2014). Care or cash? The effect of child care subsidies on student performance. *Review of Economics and Statistics,* 96(5), 824-837.

Breining, S., Doyle, J., Figlio, D., Karbownik, K., & Roth, J. (2020). Birth order and delinquency: Evidence from Denmark and Florida. *Journal of Labor Economics,* 38(1): 95-142.

Chakrabarti, R. (2008). Can increasing private school participation and monetary loss in a voucher program affect public school performance? Evidence from Milwaukee. *Journal of Public Economics,* 92, 1371-1393.

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *Quarterly Journal of Economics,* 126(4), 1593-1660.

Chingos, M.M. (2013). Class size and student outcomes: Research and policy implications. *Journal of Policy Analysis and Management,* 32(2), 411-438.

Cohodes, S., Setren, E., & Walters, C. (2019). Can successful schools replicate? Scaling up Boston's charter school sector. NBER working paper 25796.

Dee, T.S. (1998). Competition and the quality of public schools. *Economics of Education Review,* 17(4), 419-427.

Dhuey, E., Figlio, D., Karbownik, K., & Roth, J. (2019). School starting age and cognitive development. *Journal of Policy Analysis and Management,* 38(3), 538-578.

EdChoice. (2019, February 12). *School choice in American*. Retrieved from EdChoice: https://www.edchoice.org/school-choice/school-choice-in-america/#

Egalite, A. J. (2016). *Competitive impacts of means-tested vouchers on public school performance: Evidence from Louisiana and Indiana.* Cambridge, MA: Program on Education and Governance Working Paper Series, Harvard Kennedy School. Retrieved

February 23, 2019, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.696.8049&rep=rep1&type=pdf

Egalite, A. J., & Wolf, P. J. (2016). A review of the empirical research on private school choice. *Peabody Journal of Education,* 91(4), 441-454.

Epple, D., & Romano, R.E. (2008). Educational vouchers and cream-skimming. *International Economic Review,* 49(4), 1395-1435.

Epple, D., Romano, R. E., & Urquiola, M. (2017). School vouchers: A survey of the economics literature. *Journal of Economic Literature,* 55(2), 441-492.

Figlio, D. N., & Hart, C. M. (2014). Competitive effects of means-tested school vouchers. *American Economic Journal: Applied Economics,* 6(1), 133-156.

Figlio, D., & Karbownik, K. (2016). *Evaluation of Ohio's EdChoice Scholarship Program: Selection, competition, and performance effects.* Washington, DC: Thomas B. Fordham Institute.

Figlio, D., Guryan, J., & Karbownik, K., Roth, J. (2014). The effects of poor neonatal health on children's cognitive development. *American Economic Review,* 104(12), 3921-3955.

Florida Department of Education. (2018, September). *Fact Sheet: Florida Tax Credit Scholarship Program*. Retrieved February 13, 2019, from http://www.fldoe.org/core/fileparse.php/5606/urlt/FTC-Sept-2018.pdf

Florida Department of Education. (n.d.a). *Student Enrollment: State Level.* Retrieved January 31, 2019 from https://edstats.fldoe.org/SASWebReportStudio/gotoReportSection.do?sectionNumber=1

Florida Department of Education. (n.d.b). *FCAT Historical*. Retrieved February 3, 2019, from http://www.fldoe.org/accountability/assessments/k-12-student-assessment/archive/fcat/

Florida Statute 220.187 (2001).

Fredriksson, P., Öckert, B., & Oosterbeek, H. (2013). Long-term effects of class size. *Quarterly Journal of Economics,* 128(1), 249-285.

Gilraine, M., Petronijevic, U., & Singleton, J.D. (2019). Horizontal differentiation and the policy effect of charter schools. Unpublished manuscript.

Greene, J. P., & Winters, M. A. (2007). An evaluation of the effect of DC's voucher program on public school achievement and racial integration after one year. *Catholic Education: A Journal of Inquiry and Practice,* 11(1), 83-101.

Hoxby, C.M. (1994). Do private schools provide competition for public schools? National Bureau of Economic Research (Cambridge, MA) Working Paper No. 4978,

Hoxby, C. M. (2003). School choice and school productivity: Could school choice be a tide that lifts all boats? In C. M. Hoxby, *The Economics of School Choice* (pp. 287-241). Chicago, IL: University of Chicago Press.

Jepsen, C. (2002). The role of aggregation in estimating the effects of private school competition on student achievement. *Journal of Urban Economics,* 52: 477-500.

Jepsen, C., & Rivkin, S. (2009). Class size reduction and student achievement: The potential tradeoff between teacher quality and class size. *Journal of Human Resources,* 44(1), 223-250.

Krueger, A.B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics,* 114(2), 497-532.

Lankford, H., & Wyckoff, J. (1995). Where has the money gone? An analysis of school district spending in New York. *Educational Evaluation and Policy Analysis,* 17(2), 195-218.

Leuven, E., & Løkken, S.A. (Forthcoming). Long-term effects of class size in compulsory school. *Journal of Human Resources.*

Lindahl, M. (2005). Home versus school learning: A new approach to estimating the effect of class size on achievement. *Scandanavian Journal of Economics,* 107(2), 375-394.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica,* 49(6), 1417-1426.

Ridley, M., & Terrier, C. (2018). Fiscal and education spillovers from charter school expansion. NBER working paper 25070.

Rouse, C. E., Hannaway, J., Goldhaber, D., & Figlio, D. (2013). Feeling the Florida heat? How low-performing schools respond to voucher and accountability pressure. *American Economic Journal: Economic Policy, 5*(2), 251-281.

Urquiola, M. (2016). Chapter 4. Competition among schools: Traditional public and private schools. In E. A. Hanushek, S. Machin, & L. Woessmann, *Handbook of the Economics of Education* (pp. 209-237). Oxford: Elsevier.

Appendix A. Composition of Students in Florida Public Schools

Here we explore the extent to which competitive pressures affect the composition of students ever appearing in Florida public schools. Recall that, during this paper's study period, voucher participants must either have spent the previous year in Florida public schools or been entering kindergarten or first grade, and the latter would never be observed in the sample. Since the empirical strategy in our paper relies on student fixed effects, would-be peers never observed in the public schools will not contribute to changes in students' schooling environments and thus our coefficients should remain unbiased. Who is in the sample, however, could affect the external validity and interpretation of our results.

To address this question, we analyze whether the voucher program's roll-out affected which children were present in the birth records but not in the school records. To do so, we proxy for the zip code of birth's level of competition pressure in any given year by re-weighting our measures of competition pressure (introduced in Section II.C.) for birth cohorts expected to enter grade one after the program started (September 1995 to December 2002 births) with empirically observed flows of students born in any given zip code to all possible grade one schools as observed for birth cohorts entering schooling before the program started (January 1994 to August 1995 births). Table A1 shows how the voucher program roll-out affected the probability that a child would ultimately appear in the Florida public school data, both overall (panel A) and stratified for samples with a given characteristic (e.g. child of high school dropout mother in panel D or child of immigrant mother in panel J).

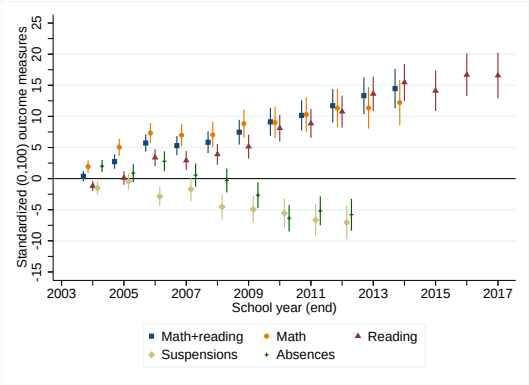
We observe that, unsurprisingly, as the program expanded fewer students born in communities with greater competitive pressures ended up in public schools, meaning that locales with more competition pressure straightaway were the places sending more children to private schools as the voucher program expanded. These results are concentrated in the set of children

whose births were funded by Medicaid and those with relatively poorly-educated mothers which makes sense since the program supports vouchers for low-income families. At the same time, however, as the program expanded, Black children and the children of immigrants were disproportionately likely to never appear in public schooling in communities with greater competitive pressures. It is also worth highlighting that the estimates from Table A1 are very modest in magnitude with effect sizes not exceeding 6 percent of sample mean. In summary, the voucher program led to a public school sector that is modestly more affluent with higher parental education. These composition changes, albeit small in magnitude, underscore the importance of gauging heterogeneity in the effects of competition pressure, as we report in Section III.C. On the one hand, such selection could reduce the estimated competitive effects if higher-SES individuals are less responsive to the effects of competition. On the other hand, it could increase them if there is complementarity between school-level student ability and competitive pressure. In neither case the estimates will be biased given our identification strategy.

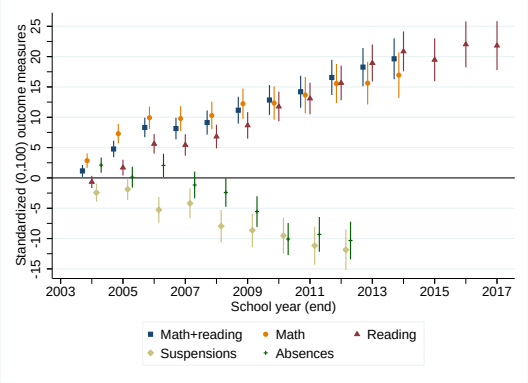
Figures

Figure 1: Effects of voucher expansion over school years for standardized outcomes

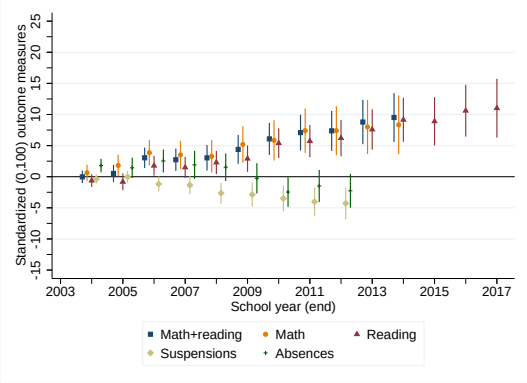
A. Pooled



B. Children ever on free or reduced price lunch



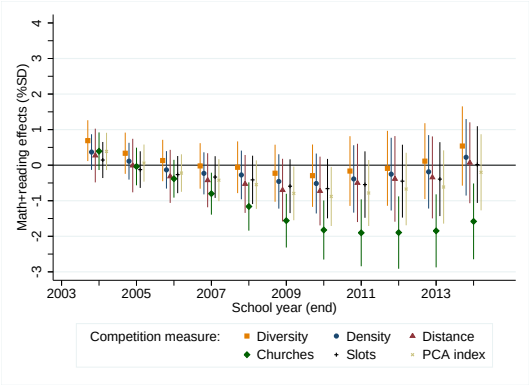
C. Children never on free or reduced price lunch



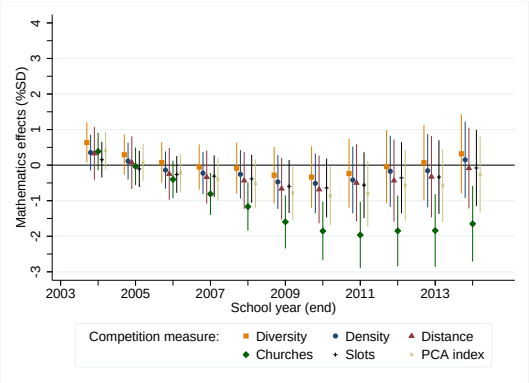
Note: These figures plot modified estimates from the main specification estimated in panel F of Table 3 and from heterogeneity analysis from panels A and B of Table 4, where instead of interaction between composite competition measure and log number of scholarships we plot composite competition measure interacted with school years, and with baseline omitted year 2002/2003. Panel A presents results for pooled sample (Table 3) while panels B and C divide the sample by free and reduced price lunch status of a child (Table 4). Outcomes are averaged test scores in mathematics and reading (navy squares); mathematics test scores (orange circles); reading test scores (maroon triangles); likelihood of being suspended (khaki diamonds); and absence rate (green pluses). Each outcome variable is standardized in its empirical sample to have mean zero and standard deviation of 100. Spikes present 95% confidence intervals based on standard errors clustered at grade one school level.

Figure 2: Estimating potential peer effects

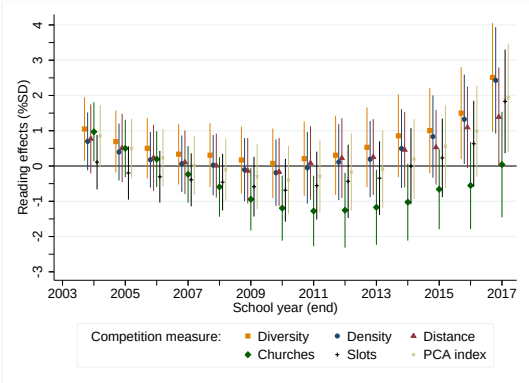
A. School-level predicted averaged math + reading



B. School-level predicted mathematics

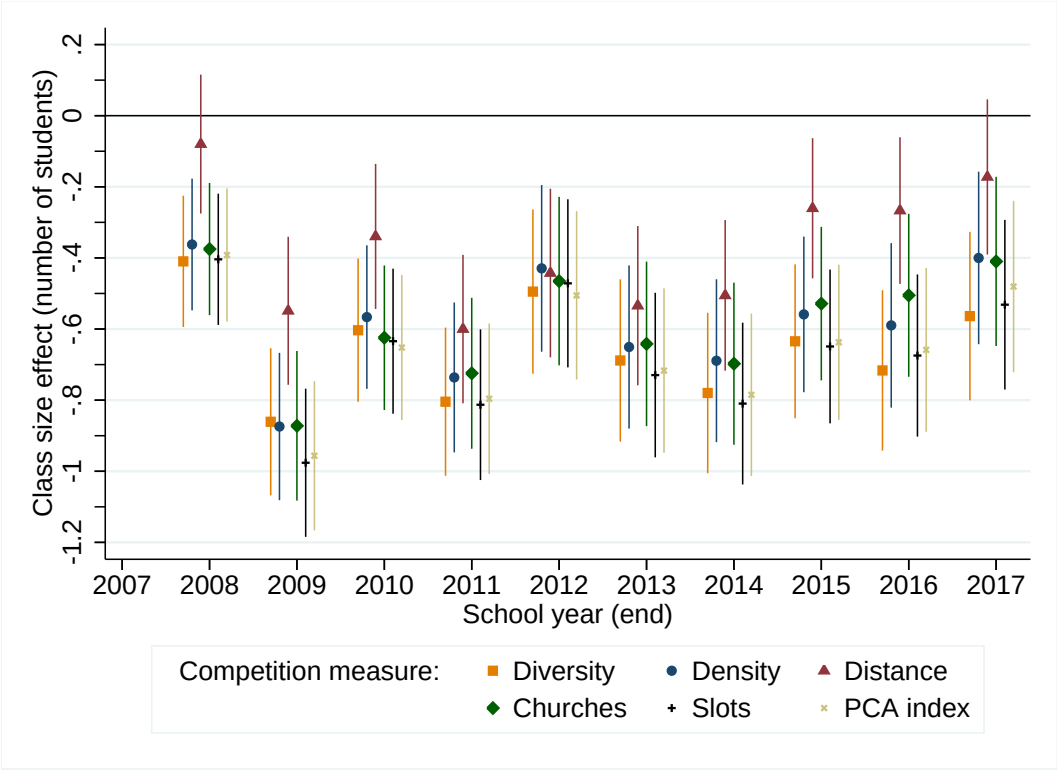


C. School-level predicted reading



Note: These figures plot potential peer effect estimates where the dependent variables are predicted rather than actual test scores. Predicted tests scores are based on predicted values from a regression of actual test scores on year and month of birth dummies, gender, birth weight, maternal years of education dummies, gestational age dummies, marital status, mother's place of birth, race, ethnicity, maternal age at birth, prior number of births to mother, month prenatal care began, complications of labor and delivery, abnormal conditions at birth, congenital anomalies, maternal health problems and Medicaid paid birth. We then aggregate these data at grade one school by year level and measure competition according to panels A to F in Table 3. Regressions include interaction between competition measure and school year along school year and grade one school fixed effects. Effects are presented relative to baseline (omitted) 2002/03 school year. Panel A presents results for predicted school-level averaged mathematics and reading scores, panel B presents results for predicted school-level mathematics scores, and panel C presents results for predicted school-level reading scores. Spikes present 95% confidence intervals based on standard errors clustered at grade one school level.

Figure 3: Estimating potential class size effects



Note: This figure plots estimates where the dependent variable is class size measured at school-by-year level. Data on class size for school years 2006/2007 to 2016/2017 are based on reports provided by Florida Department of Education (http://www.fldoe.org/finance/budget/class-size/class-size-reduction-averages.stml) separately for grades PK to 3, 4 to 8 and 9 to 12. For each school we weight these reported class sizes according to actual grades served e.g., if school is serving grades PK to 8 then we compute school-level class size as $CS = 0.5CS_{PK-3} + 0.5CS_{4-8} + 0CS_{9-12}$. These are then merged at grade one school by year level to competition measures used in panels A to F in Table 3. Regressions include interaction between competition measure and school year along school year and grade one school fixed effects, and school year 2006/2007 is a reference (omitted) category. Spikes present 95% confidence intervals based on standard errors clustered at grade one school level.

Tables

Table 1: Program expansion

School year	(1) Designated state funds	(2) Realized spending	(3) Number of participating private schools	(4) Number of scholarship enrollments	(5) Maximum amount granted per student	(6) Maximum annual family income allowed
2002/03	50,000,000	50,000,000	N/A	15,585	3,500	< 185% FPL
2003/04	88,000,000	40,000,000	924	11,550	3,500	< 185% FPL
2004/05	88,000,000	36,655,500	973	10,549	3,500	< 185% FPL
2005/06	88,000,000	46,745,482	895	15,123	3,500	< 185% FPL
2006/07	88,000,000	59,300,655	948	17,819	3,750	< 200% FPL
2007/08	88,000,000	73,450,691	933	21,493	3,750	< 200% FPL
2008/09	118,000,000	88,626,463	1,002	24,871	3,950	< 200% FPL
2009/10	118,000,000	106,049,940	1,033	28,927	3,950	< 200% FPL
2010/11	140,000,000	129,474,868	1,114	34,550	4,106	< 230% FPL
2011/12	175,000,000	147,481,308	1,216	40,248	4,011	< 230% FPL
2012/13	229,000,000	206,974,102	1,338	51,075	4,335	< 230% FPL
2013/14	286,000,000	274,495,570	1,429	59,822	4,880	< 230% FPL
2014/15	357,800,000	344,887,372	1,533	69,950	5,272	< 230% FPL
2015/16	447,000,000	418,693,458	1,602	78,664	5,677	< 230% FPL
2016/17	559,000,000	539,252,526	1,733	98,936	5,886	< 260% FPL
2017/18	698,000,000	641,024,651	1,818	108,098	7,208	< 260% FPL

Note: This table presents Florida Tax Credit (FTC) Scholarship Program expansion between school years 2002/03 and 2017/18. Column 1 shows total amount of tax credits which may be a granted in given year; column 2 shows realized spending in the program; column 3 shows number of participating private schools; column 4 shows number of students enrolled through the scholarship program; column 5 shows maximum amount of scholarship that can be awarded; column 6 shows maximum annual family income allowed. Columns 1, 5 and 6 are based on Florida Statues 220.187 for years 2002/03 to 2009/10 and Florida Statues 1002.395 for years 2010/11 to 2017/18. Columns 2 to 4 are based on Florida Department of Education reports (Choice Facts, 2008; 2010; 2011; 2014; 2018).

Table 2: Descriptive statistics

	(1)	(2)	(3)	(4)
			Competition index	
	All births	Empirical sample	Above median	Below median
		Panel A. Sociodemographic characteristics		
White	55.4	50.9	37.3	68.0
African-American	19.4	23.3	30.2	14.5
Hispanic	22.8	23.9	30.5	15.6
Mother foreign born	23.2	23.0	29.7	14.5
Male	51.2	51.1	51.1	51.1
Mother HS dropout	20.9	24.9	25.0	24.6
Mother HS graduate	58.8	60.4	60.7	60.1
Mother college graduate	20.2	14.7	14.3	15.2
Mother age at birth	27.1	26.6	26.6	26.5
Parents married at birth	64.9	59.2	54.6	65.1
Ever on FRPL	N/A	71.8	75.8	66.6
		Panel B. Competition measures		
Diversity	N/A	5.1	6.8	2.9
Density		15.6	24.0	5.0
Distance		-1.9	-1.0	-3.0
Number of churches		143.0	207.9	61.0
Number of slots		2.9	4.7	0.7
Competition index (PCA)		0.3	1.6	-1.5
		Panel C. Outcomes		
Math+reading score	N/A	0.0	-4.4	5.6
Math score		0.0	-4.1	5.0
Reading scores		0.0	-4.8	6.0
Likelihood suspended		13.7	14.3	12.9
Absence rate		5.0	5.0	5.2
Maximum # observations	2,028,798	6,971,914	3,890,161	3,081,753
Maximum # children	2,028,798	1,255,084	755,254	609,646

Note: Panel A presents means of sociodemographic variables (all indicator variables multiplied by 100); panel B presents means of competition measures with distance reverse coded (more positive values indicate higher competition); panel C presents outcome variables (all multiplied by 100). Column 1 presents characteristics of full sample of births between 1992 and 2002; column 2 presents characteristics of our preferred empirical sample for school years between 2002/03 to 2016/17; columns 3 and 4 divide sample from column 2 into two mutually exclusive categories based on median of the PCA competition index.

Table 3: Effects of voucher expansion by baseline competition measures

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
			Panel A. Diversity		
Expansion × above median competition	4.233***	1.937***	6.539***	-0.504*	-0.265***
	(0.599)	(0.736)	(0.618)	(0.267)	(0.052)
			Panel B. Density		
Expansion × above median competition	5.293***	2.817***	7.566***	-1.109***	-0.258***
	(0.586)	(0.728)	(0.608)	(0.268)	(0.052)
			Panel C. Distance		
Expansion × above median competition	1.648***	-0.308	3.542***	-0.430*	-0.151***
	(0.590)	(0.712)	(0.622)	(0.261)	(0.052)
			Panel D. Churches nearby		
Expansion × above median competition	3.917***	1.643**	5.966***	-1.428***	-0.223***
	(0.598)	(0.727)	(0.626)	(0.265)	(0.051)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
# children	1,222,165	1,222,912	1,223,799	1,228,461	1,228,461
			Panel E. Slots per grade		
Expansion × above median competition	6.064***	3.361***	8.684***	-1.749***	-0.290***
	(0.590)	(0.732)	(0.604)	(0.266)	(0.052)
			Panel F. Competitive Pressure Index		
Expansion × above median competition	5.111***	2.639***	7.389***	-1.282***	-0.281***
	(0.593)	(0.738)	(0.611)	(0.267)	(0.052)
Mean [SD] of Y	0.000 [93.085]	-0.038 [99.977]	-0.017 [99.984]	13.666 [34.349]	5.041 [5.790]
Observations	6,160,525	6,104,889	6,584,014	5,427,985	5,427,985
# children	1,221,023	1,220,753	1,223,123	1,225,713	1,225,713

Note: Sample is based on individual-level observations in grades 3 to 8 for students attending Florida public schools between 2002/03 to 2016/17 and born between 1992 and 2002. Each child has to be observed at least in grade 1 so that we can assign them school-level competition measures which are based on Figlio and Hart (2014); these are assigned to individuals for the schools they attend in grades 1 and 6. Thus, there are up to two values of competition observed for each individual. Expansion is measured at annual level between 2002/03 and 2016/17 as logarithm of number of scholarships awarded. Test scores are based on FCAT developmental scores for years 2000/2001 to 2013/2014 and on FSA developmental scores for years 2014/2015 to 2016/2017, and we standardize them in-sample by years and grade to have mean 0 and standard deviation of 1. These standardized scores are then multiplied by 100. Averaged mathematics and reading as well as mathematics test scores are available up to school year 2013/2014 while reading test scores are available up to school year 2016/2017. Suspensions (indicator for ever being suspended in a given year) and absences (absence rate in a given year net of suspension days) are measured for years 2002/03 to 2011/2012, and they are multiplied by 100. Each column represents a separate outcome variable. Competition measures are: number of denominational types represented (panel A); number of local private schools (panel B); miles to nearest private school competitor (panel C); number of churches, synagogues, and mosques (panel D); number of private school slots per grade (panel E); and principal components analysis competition index ("Competitive Pressure Index") based on five measures from panels A to E (panel F). Regression table presents interactions between competition measure (dummy for competition above median in the full sample of schools) and log of expansion measure, and all regressions include student-by-school level FE and grade-by-school year FE. School level is defined as indicator for grade 6 to 8 vs. 3 to 5. Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table 4: Heterogeneity in the effects of voucher expansion

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
			Panel A. Ever on free or reduced price lunch		
Expansion × above median competition	6.504***	3.660***	9.187***	-1.946***	-0.431***
	(0.624)	(0.770)	(0.666)	(0.338)	(0.063)
Mean [SD] of Y	-21.409 [89.396]	-21.402 [97.315]	-21.010 [96.385]	17.305 [37.829]	5.587 [6.314]
Observations	4,362,211	4,324,143	4,696,426	3,803,417	3,803,417
			Panel B. Never on free or reduced price lunch		
Expansion × above median competition	3.971***	2.781***	4.860***	-0.917***	-0.144***
	(0.759)	(0.970)	(0.771)	(0.254)	(0.055)
Mean [SD] of Y	51.933 [80.552]	51.840 [86.435]	52.216 [89.054]	5.147 [22.095]	3.763 [4.044]
Observations	1,798,314	1,780,746	1,887,588	1,624,568	1,624,568
			Panel C. Mother high school dropout		
Expansion × above median competition	5.256***	2.956***	7.504***	-2.159***	-0.338***
	(0.762)	(0.953)	(0.852)	(0.479)	(0.096)
Mean [SD] of Y	-43.261 [88.539]	-42.496 [97.764]	-43.506 [95.362]	21.774 [41.271]	6.460 [7.175]
Observations	1,504,461	1,492,865	1,609,399	1,334,914	1,334,914
			Panel D. Mother high school graduate		
Expansion × above median competition	5.675***	2.900***	8.297***	-1.113***	-0.337***
	(0.597)	(0.752)	(0.620)	(0.287)	(0.052)
Mean [SD] of Y	1.324 [87.359]	0.979 [94.291]	1.434 [94.682]	12.655 [33.247]	4.870 [5.415]
Observations	3,739,944	3,709,186	3,989,457	3,304,238	3,304,238
			Panel E. Mother college graduate		
Expansion × above median competition	3.163***	1.863	3.812***	-0.776***	-0.060
	(0.902)	(1.146)	(0.940)	(0.292)	(0.057)
Mean [SD] of Y	65.640 [82.760]	65.993 [88.952]	65.155 [91.140]	4.182 [20.017]	3.356 [3.792]
Observations	916,120	902,838	985,158	788,833	788,833
			Panel F. White mother		
Expansion × above median competition	4.682***	3.803***	5.269***	-1.116***	-0.228***
	(0.677)	(0.860)	(0.657)	(0.295)	(0.064)
Mean [SD] of Y	22.894 [88.067]	22.368 [93.933]	23.256 [96.293]	10.145 [30.193]	5.257 [5.730]
Observations	3,156,514	3,132,289	3,348,248	2,815,187	2,815,187
			Panel G. African-American mother		
Expansion × above median competition	4.319***	3.693***	5.277***	-0.121	-0.021
	(0.844)	(1.122)	(0.834)	(0.650)	(0.094)
Mean [SD] of Y	-48.816 [88.572]	-49.807 [98.517]	-47.479 [94.430]	25.486 [43.578]	5.311 [6.483]
Observations	1,414,642	1,403,085	1,513,010	1,248,570	1,248,570
			Panel H. Hispanic mother		
Expansion × above median competition	5.488***	1.557	9.126***	-2.452***	-0.170**
	(0.989)	(1.274)	(1.093)	(0.423)	(0.084)
Mean [SD] of Y	-6.380 [88.165]	-4.972 [95.255]	-7.608 [95.334]	10.518 [30.678]	4.441 [5.190]
Observations	1,469,512	1,451,296	1,593,065	1,261,615	1,261,615
			Panel I. Mother born in the U.S.		
Expansion × above median competition	3.607***	2.142***	4.744***	-0.898***	-0.288***
	(0.591)	(0.759)	(0.564)	(0.296)	(0.059)
Mean [SD] of Y	-1.281 [93.900]	-1.947 [100.685]	-0.724 [100.883]	14.870 [35.579]	5.374 [6.008]
Observations	4,739,008	4,701,848	5,049,671	4,198,015	4,198,015
			Panel J. Foreign born mother		
Expansion × above median competition	2.690***	-0.945	6.103***	-1.631***	-0.127*
	(0.926)	(1.223)	(1.018)	(0.390)	(0.071)
Mean [SD] of Y	4.272 [90.182]	6.359 [97.294]	2.311 [96.932]	9.558 [29.402]	3.907 [4.805]
Observations	1,421,517	1,403,041	1,534,343	1,229,970	1,229,970

Note: Specifications are based on those in panel F of Table 3 with the baseline sample split by child's free or reduced price lunch history (panels A and B), maternal education (panels C to E), race/ethnicity (panels F to H), and maternal immigration status (panels I and J). Outcome variables are averaged mathematics and reading test scores (column 1), mathematics test scores (column 2), reading test scores (column 3), likelihood of being suspended (column 4), and absence rate (column 5). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table 5: Robustness of the preferred estimates

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
			Panel A. Baseline		
Expansion × above median competition	5.111***	2.639***	7.389***	-1.282***	-0.281***
	(0.593)	(0.738)	(0.611)	(0.267)	(0.052)
Mean [SD] of Y	0.000 [93.085]	-0.038 [99.977]	-0.017 [99.984]	13.666 [34.349]	5.041 [5.790]
Observations	6,160,525	6,104,889	6,584,014	5,427,985	5,427,985
			Panel B. Continuous competition measure		
Expansion × competition	1.453***	0.620***	2.296***	-0.422***	-0.086***
	(0.157)	(0.196)	(0.161)	(0.068)	(0.014)
Mean [SD] of Y	0.000 [93.085]	-0.038 [99.977]	-0.017 [99.984]	13.666 [34.349]	5.041 [5.790]
Observations	6,160,525	6,104,889	6,584,014	5,427,985	5,427,985
			Panel C. Log funding expansion measure		
Expansion × above median competition	4.747***	2.547***	6.536***	-1.358***	-0.311***
	(0.555)	(0.695)	(0.552)	(0.268)	(0.052)
Mean [SD] of Y	0.000 [93.085]	-0.038 [99.977]	-0.017 [99.984]	13.666 [34.349]	5.041 [5.790]
Observations	6,160,525	6,104,889	6,584,014	5,427,985	5,427,985
			Panel D. Competition measure unweighted with elementary to middle school flows		
Expansion × above median competition	4.337***	1.955***	6.528***	-1.031***	-0.280***
	(0.574)	(0.722)	(0.588)	(0.250)	(0.048)
Mean [SD] of Y	-0.003 [92.865]	-0.002 [99.755]	-0.072 [99.844]	13.223 [33.874]	5.011 [5.706]
Observations	5,761,773	5,714,711	6,123,884	5,117,781	5,117,781
			Panel E. Weights based solely on pre-program transitions between grades 5 and 6		
Expansion × above median competition	5.633***	3.111***	8.209***	-1.333***	-0.228***
	(0.578)	(0.718)	(0.594)	(0.265)	(0.052)
Mean [SD] of Y	0.002 [93.084]	-0.036 [99.975]	-0.014 [99.984]	13.653 [34.336]	5.038 [5.793]
Observations	6,071,801	6,016,952	6,487,847	5,351,967	5,351,967
			Panel F. Limiting the sample to schools with at least one competitor within 5 miles		
Expansion × above median competition	4.271***	2.221***	6.131***	-1.129***	-0.193***
	(0.618)	(0.767)	(0.642)	(0.284)	(0.052)
Mean [SD] of Y	0.003 [93.124]	-0.042 [99.976]	-0.019 [99.984]	13.636 [34.317]	5.008 [5.781]
Observations	5,703,761	5,650,965	6,092,805	5,030,501	5,030,501
			Panel G. School years 2002/03 to 2011/12		
Expansion × above median competition	4.584***	2.644***	6.578***	-1.282***	-0.281***
	(0.589)	(0.738)	(0.615)	(0.267)	(0.052)
Mean [SD] of Y	-0.219 [92.915]	-0.040 [99.974]	-0.019 [99.981]	13.666 [34.349]	5.041 [5.790]
Observations	5,336,140	5,323,917	5,323,137	5,427,985	5,427,985
			Panel H. Balanced panel (6-years)		
Expansion × above median competition	5.463***	3.496***	6.933***	-1.427***	-0.364***
	(0.660)	(0.825)	(0.683)	(0.306)	(0.055)
Mean [SD] of Y	0.095 [92.933]	-0.037 [99.974]	-0.015 [99.986]	14.524 [35.234]	4.969 [5.551]
Observations	3,958,889	3,919,656	5,303,632	2,845,185	2,845,185
			Panel I. In sample standardized test scores		
Expansion × above median competition	4.825***	2.484***	7.165***	N/A	N/A
	(0.586)	(0.733)	(0.603)		
Mean [SD] of Y	0.061 [92.587]	0.190 [99.823]	-0.069 [99.805]		
Observations	5,756,691	5,756,691	5,756,691		
			Panel J. In population standardized scores sample		
Expansion × above median competition	4.913***	2.729***	7.097***	N/A	N/A
	(0.555)	(0.700)	(0.573)		
Mean [SD] of Y	4.669 [89.064]	4.657 [96.371]	4.681 [95.610]		
Observations	5,756,691	5,756,691	5,756,691		

Note: Robustness checks based on estimates from panel F of Table 3. Panel A replicates the main result from panel F of Table 3; panel B replaces dummy indicator for above median pre-reform competition with continuous measure; panel C replaces logarithm of number of scholarships expansion measure with logarithm of funding; panel D assigns the middle-school pre-policy competition measures based on the actual grade 6 (middle) school initially attended by each student; panel E generates expected competitive pressure measures for middle school-aged students using only pre-policy announcement flow between elementary and middle schools; panel F limits the initial sample to only schools with at least one competitor within 5 miles; panel G restricts the sample to school years 2002/03 to 2011/12 where we observe all five outcomes; panel H restricts the sample to 6-year panel of observations starting with grade 3 and within school years available for a given variable; panels I and J restrict the sample to school years 2002/03 to 2012/13 where we observe test scores that are standardized for the full population of Florida students - panel I presents our in-sample standardization while panel J population-level standardization. Outcome variables are averaged mathematics and reading test score (column 1), mathematics test score (column 2), reading test score (column 3), likelihood of being suspended (column 4), and absence rate (column 5). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table 6: Alternative explanations: Effects of voucher program expansion on peer composition and class size

	(1) Diversity	(2) Density	(3) Distance	(4) Churches	(5) Slots	(6) PCA
			Panel A. Potential "peer effects"			
			A1. School-level predicted averaged mathematics and reading scores			
Expansion × above median competition	0.253**	0.199*	0.063	0.053	0.042	0.148
	(0.119)	(0.112)	(0.073)	(0.102)	(0.104)	(0.107)
Observations	20,009	20,009	20,009	20,009	19,986	19,986
			A2. School-level predicted mathematics scores			
Expansion × above median competition	0.218*	0.185	0.061	0.045	0.037	0.148
	(0.118)	(0.113)	(0.073)	(0.103)	(0.103)	(0.108)
Observations	20,009	20,009	20,009	20,009	19,986	19,986
			A3. School-level predicted reading scores			
Expansion × above median competition	0.294***	0.276***	0.125**	0.116	0.106	0.245**
	(0.105)	(0.104)	(0.060)	(0.092)	(0.092)	(0.103)
Observations	24,930	24,930	24,930	24,930	24,907	24,907
			Panel B. Potential class size effects			
			B1. Main class size estimates			
Expansion × above median competition	-0.042**	-0.011	-0.003	-0.012	-0.020	-0.027
	(0.018)	(0.019)	(0.011)	(0.018)	(0.018)	(0.020)
			B2. Effect on number of students per 10% increase in the progam			
Estimate	-0.0042	-0.0011	-0.0003	-0.0012	-0.002	-0.0027
			B3. Implied effect sizes based on 22% SD 3-year STAR effect (in % of SD)			
Implied effect	0.013	0.003	0.001	0.004	0.006	0.008
Observations	17,747	17,747	17,747	17,747	17,739	17,739

Note: Panel A presents potential peer effect estimates where the dependent variables are predicted rather than actual test scores. Predicted tests scores are based on predicted values from a regression of actual test scores on year and month of birth dummies, gender, birth weight, maternal years of education dummies, gestational age dummies, marital status, mother's place of birth, race, ethnicity, maternal age at birth, prior number of births to mother, month prenatal care began, complications of labor and delivery, abnormal conditions at birth, congenital anomalies, maternal health problems and Medicaid paid birth. These are then aggregate at grade one school by year level. RRegressions include interaction between competition measure (dummy for competition above median in the full sample of schools) and log of expansion measure along school year and grade one school fixed effects. Panel A1 presents results for predicted school-level averaged mathematics and reading scores, panel A2 presents results for predicted school-level mathematics scores, and panel A3 presents results for predicted school-level reading scores. Panel B presents estimates where the outcome variable is class size measured at school-by-year level. Data on class size for school years 2006/2007 to 2016/2017 are based on reports provided by Florida Department of Education (http://www.fldoe.org/finance/budget/class-size/class-size-reduction-averages.stml) separately for grades PK to 3, 4 to 8 and 9 to 12. For each school we weight these reported class sizes according to actual grades served e.g., if school is serving grades PK to 8 then we compute school-level class size as $CS = 0.5CS_{PK-3} + 0.5CS_{4-8} + 0CS_{9-12}$. These are then merged at grade one school by year level to competition measures used in panels A to F in Table 3. Regressions include interaction between competition measure (dummy for competition above median in the full sample of schools) and log of expansion measure, year fixed effects, and grade one school fixed effects. Panel B1 presents the effects of voucher program expansion on class size; panel B2 scales this estimate into effect size of 10% increase in voucher program scholarships; and panel B3 computes the implied potential gain in test scores from reduced class size based on estimates from project STAR (Krueger, 1999). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Appendix

Table A1: Voucher program expansion and likelihood of being observed in matched birth-public school records

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: Probability of being matched to school records (*100)					
Competition measures	PCA index	Diversity	Density	Distance	Churches	Slots
	Panel A: Overall (Mean = 79.8; N = 1,279,009)					
Expansion × above median	-1.055***	-1.218***	-1.238***	-1.364***	-0.963***	-1.185***
weighted competition	(0.356)	(0.351)	(0.354)	(0.385)	(0.359)	(0.356)
Implied % effect	-1.3	-1.5	-1.6	-1.7	-1.2	-1.5
	Panel B. Non-Medicaid paid birth (Mean = 74.7; N = 709,570)					
Expansion × above median	-0.115	-0.301	-0.258	-1.090**	-0.290	-0.125
weighted competition	(0.505)	(0.494)	(0.506)	(0.525)	(0.502)	(0.506)
Implied % effect	-0.2	-0.4	-0.3	-1.5	-0.4	-0.2
	Panel C. Medicaid paid birth (Mean = 86.1; N = 569,438)					
Expansion × above median	-2.342***	-2.433***	-2.435***	-1.687***	-2.036***	-2.506***
weighted competition	(0.385)	(0.384)	(0.384)	(0.415)	(0.390)	(0.378)
Implied % effect	-2.7	-2.8	-2.8	-2.0	-2.4	-2.9
	Panel D. Mother high school dropout (Mean = 87.1; N = 250,565)					
Expansion × above median	-1.451***	-1.714***	-1.581***	-1.114*	-1.370**	-1.428***
weighted competition	(0.536)	(0.528)	(0.534)	(0.597)	(0.547)	(0.529)
Implied % effect	-1.7	-2.0	-1.8	-1.3	-1.6	-1.6
	Panel E. Mother high school graduate (Mean = 82.2; N = 746,382)					
Expansion × above median	-0.769**	-0.863**	-0.818**	-1.256***	-0.686*	-0.765*
weighted competition	(0.389)	(0.381)	(0.387)	(0.409)	(0.388)	(0.391)
Implied % effect	-0.9	-1.0	-1.0	-1.5	-0.8	-0.9
	Panel F. Mother college graduate (Mean = 66.9; N = 282,062)					
Expansion × above median	-0.189	-0.397	-0.421	-0.641	-0.322	-0.683
weighted competition	(0.853)	(0.817)	(0.834)	(0.776)	(0.773)	(0.834)
Implied % effect	-0.3	-0.6	-0.6	-1.0	-0.5	-1.0
	Panel G. White, non-Hispanic, non-immigrant (Mean = 77.4; N = 640,193)					
Expansion × above median	0.034	-0.083	-0.265	-0.704	-0.232	-0.053
weighted competition	(0.433)	(0.435)	(0.436)	(0.452)	(0.440)	(0.433)
Implied % effect	0.0	-0.1	-0.3	-0.9	-0.3	-0.1
	Panel H. Black, non-Hispanic, non-immigrant (Mean = 89.5; N = 213,720)					
Expansion × above median	-2.230***	-2.566***	-2.374***	-2.439***	-1.351**	-2.255***
weighted competition	(0.597)	(0.573)	(0.590)	(0.609)	(0.620)	(0.578)
Implied % effect	-2.5	-2.9	-2.7	-2.7	-1.5	-2.5
	Panel I. Hispanic, non-immigrant (Mean = 81.4; N = 107,344)					
Expansion × above median	-0.742	-0.639	0.095	-0.493	-0.031	-1.252
weighted competition	(1.049)	(1.031)	(1.038)	(0.856)	(1.022)	(1.053)
Implied % effect	-0.9	-0.8	0.1	-0.6	0.0	-1.5
	Panel J. Immigrant mother (Mean = 77.6; N = 317,752)					
Expansion × above median	-4.296***	-4.289***	-4.231***	-2.830***	-3.508***	-4.335***
weighted competition	(0.815)	(0.810)	(0.825)	(0.872)	(0.830)	(0.779)
Implied % effect	-5.5	-5.5	-5.5	-3.6	-4.5	-5.6
	Panel K: Females (Mean = 80.2; N = 624,677)					
Expansion × above median	-1.290***	-1.313***	-1.466***	-1.326***	-0.998**	-1.221***
weighted competition	(0.419)	(0.414)	(0.418)	(0.444)	(0.426)	(0.420)
Implied % effect	-1.6	-1.6	-1.8	-1.7	-1.2	-1.5
	Panel L. Males (Mean = 79.5; N = 654,332)					
Expansion × above median	-0.816*	-1.118**	-1.006**	-1.400***	-0.910**	-1.125**
weighted competition	(0.444)	(0.438)	(0.445)	(0.460)	(0.442)	(0.438)
Implied % effect	-1.0	-1.4	-1.3	-1.8	-1.1	-1.4

Note: This table presents estimates where the outcome variable is an indicator for being matched between birth and school records multiplied by 100. Panel A presents overall probability while panels B to L present results for various subsamples. Independent variable of interest is an interaction between annual voucher expansion and weighted competition at zip code level. Analysis is based on data for cohorts entering grade one after the program stated (September 1995 and later) while weights are created based on pre-program grade one cohorts (January 1994 to August 1995). Weighting is based on observed flows of individuals born in a given zip code to all possible schools. Models further include zip code level and year fixed effects. Standard errors are clustered at zip code level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A2: Construction of Competitive Pressure Index based on principal components analysis

	(1)	(2)	(3)	(4)
	Grades 1 to 5		Grades 6 to 8	
	First component	Second component	First component	Second component
Diversity	0.474	0.144	0.474	0.169
Density	0.499	-0.214	0.494	-0.215
Distance	0.295	0.903	0.331	0.873
Number of churches	0.46	-0.229	0.455	-0.271
Number of slots	0.477	-0.257	0.463	-0.301
Eigenvalue	3.614	0.813	3.829	0.733

Note: This table reports the results of a principal components analysis of number of denominational types represented (diversity), number of local private schools (density), miles to nearest private school competitor (distance), number of churches, synagogues, and mosques, and number of private school slots per grade. The eigenvectors associated with the first (columns 1 and 3) and second (columns 2 and 4) components are reported separately for grades 1 to 5 and 6 to 8, as well as their associated eigenvalues.

Table A3: Heterogeneity in the effects of voucher expansion: Diversity measure

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
			Panel A. Ever on free or reduced price lunch		
Expansion × above median competition	5.877***	3.101***	8.604***	-1.054***	-0.370***
	(0.627)	(0.764)	(0.669)	(0.336)	(0.064)
Mean [SD] of Y	-21.420 [89.403]	-21.405 [97.323]	-21.028 [96.392]	17.271 [37.800]	5.585 [6.310]
Observations	4,379,915	4,341,810	4,714,149	3,820,075	3,820,075
			Panel B. Never on free or reduced price lunch		
Expansion × above median competition	2.180***	1.304	3.142***	-0.226	-0.170***
	(0.773)	(0.970)	(0.796)	(0.255)	(0.055)
Mean [SD] of Y	51.995 [80.557]	51.918 [86.455]	52.259 [89.053]	5.124 [22.048]	3.762 [4.042]
Observations	1,807,648	1,790,068	1,896,918	1,633,578	1,633,578
			Panel C. Mother high school dropout		
Expansion × above median competition	4.747***	2.044**	7.438***	-1.307***	-0.249***
	(0.760)	(0.935)	(0.833)	(0.484)	(0.096)
Mean [SD] of Y	-43.279 [88.547]	-42.509 [97.770]	-43.525 [95.370]	21.728 [41.240]	6.455 [7.169]
Observations	1,511,217	1,499,602	1,616,155	1,341,389	1,341,389
			Panel D. Mother high school graduate		
Expansion × above median competition	4.712***	2.249***	7.206***	-0.259	-0.318***
	(0.601)	(0.751)	(0.629)	(0.286)	(0.052)
Mean [SD] of Y	1.336 [87.366]	1.001 [94.301]	1.438 [94.687]	12.627 [33.215]	4.869 [5.413]
Observations	3,754,631	3,723,846	4,004,160	3,318,165	3,318,165
			Panel E. Mother college graduate		
Expansion × above median competition	1.706*	0.896	2.423**	-0.360	-0.103*
	(0.916)	(1.151)	(0.981)	(0.291)	(0.058)
Mean [SD] of Y	65.699 [82.758]	66.072 [88.959]	65.192 [91.139]	4.160 [19.967]	3.355 [3.789]
Observations	921,715	908,430	990,752	794,099	794,099
			Panel F. White mother		
Expansion × above median competition	3.130***	2.353***	3.873***	-0.115	-0.233***
	(0.694)	(0.867)	(0.674)	(0.304)	(0.063)
Mean [SD] of Y	22.957 [88.068]	22.440 [93.939]	23.308 [96.293]	10.113 [30.150]	5.253 [5.726]
Observations	3,170,846	3,146,594	3,362,580	2,828,887	2,828,887
			Panel G. African-American mother		
Expansion × above median competition	4.246***	3.501***	5.361***	0.767	0.079
	(0.817)	(1.073)	(0.810)	(0.640)	(0.095)
Mean [SD] of Y	-48.852 [88.588]	-49.836 [98.533]	-47.516 [94.446]	25.448 [43.557]	5.310 [6.478]
Observations	1,420,501	1,408,927	1,518,874	1,254,177	1,254,177
			Panel H. Hispanic mother		
Expansion × above median competition	4.934***	1.178	8.515***	-2.174***	-0.209**
	(0.979)	(1.232)	(1.095)	(0.429)	(0.086)
Mean [SD] of Y	-6.396 [88.173]	-4.972 [95.269]	-7.637 [95.336]	10.491 [30.643]	4.440 [5.187]
Observations	1,475,615	1,457,395	1,599,177	1,267,259	1,267,259
			Panel I. Mother born in the U.S.		
Expansion × above median competition	2.722***	1.379*	3.934***	0.035	-0.267***
	(0.594)	(0.756)	(0.572)	(0.298)	(0.058)
Mean [SD] of Y	-1.250 [93.922]	-1.907 [100.711]	-0.701 [100.901]	14.836 [35.545]	5.371 [6.004]
Observations	4,759,439	4,722,232	5,070,104	4,217,517	4,217,517
			Panel J. Foreign born mother		
Expansion × above median competition	2.082**	-1.261	5.484***	-1.457***	-0.138*
	(0.929)	(1.202)	(1.026)	(0.389)	(0.076)
Mean [SD] of Y	4.284 [90.192]	6.389 [97.308]	2.307 [96.939]	9.529 [29.361]	3.905 [4.802]
Observations	1,428,124	1,409,646	1,540,963	1,236,136	1,236,136

Note: Specifications are based on those in panel A of Table 3 with the baseline sample split by child's free or reduced price lunch history (panels A and B), maternal education (panels C to E), race/ethnicity (panels F to H), and maternal immigration status (panels I and J). Outcome variables are averaged mathematics and reading test scores (column 1), mathematics test scores (column 2), reading test scores (column 3), likelihood of being suspended (column 4), and absence rate (column 5). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A4: Heterogeneity in the effects of voucher expansion: Density measure

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
			Panel A. Ever on free or reduced price lunch		
Expansion × above median competition	6.818***	3.959***	9.495***	-1.677***	-0.375***
	(0.613)	(0.756)	(0.658)	(0.339)	(0.063)
Mean [SD] of Y	-21.420 [89.403]	-21.405 [97.323]	-21.028 [96.392]	17.271 [37.800]	5.585 [6.310]
Observations	4,379,915	4,341,810	4,714,149	3,820,075	3,820,075
			Panel B. Never on free or reduced price lunch		
Expansion × above median competition	3.613***	2.323**	4.564***	-0.833***	-0.150***
	(0.753)	(0.961)	(0.773)	(0.256)	(0.054)
Mean [SD] of Y	51.995 [80.557]	51.918 [86.455]	52.259 [89.053]	5.124 [22.048]	3.762 [4.042]
Observations	1,807,648	1,790,068	1,896,918	1,633,578	1,633,578
			Panel C. Mother high school dropout		
Expansion × above median competition	5.417***	3.125***	7.684***	-1.824***	-0.259***
	(0.746)	(0.929)	(0.843)	(0.483)	(0.096)
Mean [SD] of Y	-43.279 [88.547]	-42.509 [97.770]	-43.525 [95.370]	21.728 [41.240]	6.455 [7.169]
Observations	1,511,217	1,499,602	1,616,155	1,341,389	1,341,389
			Panel D. Mother high school graduate		
Expansion × above median competition	5.784***	3.028***	8.371***	-0.948***	-0.311***
	(0.591)	(0.744)	(0.617)	(0.287)	(0.051)
Mean [SD] of Y	1.336 [87.366]	1.001 [94.301]	1.438 [94.687]	12.627 [33.215]	4.869 [5.413]
Observations	3,754,631	3,723,846	4,004,160	3,318,165	3,318,165
			Panel E. Mother college graduate		
Expansion × above median competition	3.133***	1.583	3.928***	-0.633**	-0.081
	(0.901)	(1.147)	(0.945)	(0.290)	(0.057)
Mean [SD] of Y	65.699 [82.758]	66.072 [88.959]	65.192 [91.139]	4.160 [19.967]	3.355 [3.789]
Observations	921,715	908,430	990,752	794,099	794,099
			Panel F. White mother		
Expansion × above median competition	4.587***	3.666***	5.187***	-0.975***	-0.224***
	(0.676)	(0.853)	(0.660)	(0.299)	(0.063)
Mean [SD] of Y	22.957 [88.068]	22.440 [93.939]	23.308 [96.293]	10.113 [30.150]	5.253 [5.726]
Observations	3,170,846	3,146,594	3,362,580	2,828,887	2,828,887
			Panel G. African-American mother		
Expansion × above median competition	5.015***	4.500***	5.826***	0.345	0.070
	(0.807)	(1.072)	(0.808)	(0.649)	(0.094)
Mean [SD] of Y	-48.852 [88.588]	-49.836 [98.533]	-47.516 [94.446]	25.448 [43.557]	5.310 [6.478]
Observations	1,420,501	1,408,927	1,518,874	1,254,177	1,254,177
			Panel H. Hispanic mother		
Expansion × above median competition	5.472***	1.236	9.355***	-2.479***	-0.176**
	(0.984)	(1.245)	(1.072)	(0.414)	(0.082)
Mean [SD] of Y	-6.396 [88.173]	-4.972 [95.269]	-7.637 [95.336]	10.491 [30.643]	4.440 [5.187]
Observations	1,475,615	1,457,395	1,599,177	1,267,259	1,267,259
			Panel I. Mother born in the U.S.		
Expansion × above median competition	3.829***	2.422***	4.909***	-0.684**	-0.264***
	(0.583)	(0.748)	(0.558)	(0.299)	(0.058)
Mean [SD] of Y	-1.250 [93.922]	-1.907 [100.711]	-0.701 [100.901]	14.836 [35.545]	5.371 [6.004]
Observations	4,759,439	4,722,232	5,070,104	4,217,517	4,217,517
			Panel J. Foreign born mother		
Expansion × above median competition	2.513***	-1.358	6.100***	-1.570***	-0.109
	(0.932)	(1.194)	(1.016)	(0.389)	(0.070)
Mean [SD] of Y	4.284 [90.192]	6.389 [97.308]	2.307 [96.939]	9.529 [29.361]	3.905 [4.802]
Observations	1,428,124	1,409,646	1,540,963	1,236,136	1,236,136

Note: Specifications are based on those in panel B of Table 3 with the baseline sample split by child's free or reduced price lunch history (panels A and B), maternal education (panels C to E), race/ethnicity (panels F to H), and maternal immigration status (panels I and J). Outcome variables are averaged mathematics and reading test scores (column 1), mathematics test scores (column 2), reading test scores (column 3), likelihood of being suspended (column 4), and absence rate (column 5). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A5: Heterogeneity in the effects of voucher expansion: Distance measure

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
			Panel A. Ever on free or reduced price lunch		
Expansion × above median competition	3.208***	1.042	5.365***	-1.101***	-0.233***
	(0.632)	(0.750)	(0.682)	(0.326)	(0.063)
Mean [SD] of Y	-21.420 [89.403]	-21.405 [97.323]	-21.028 [96.392]	17.271 [37.800]	5.585 [6.310]
Observations	4,379,915	4,341,810	4,714,149	3,820,075	3,820,075
			Panel B. Never on free or reduced price lunch		
Expansion × above median competition	0.016	-1.053	0.869	0.003	-0.151***
	(0.749)	(0.945)	(0.780)	(0.250)	(0.053)
Mean [SD] of Y	51.995 [80.557]	51.918 [86.455]	52.259 [89.053]	5.124 [22.048]	3.762 [4.042]
Observations	1,807,648	1,790,068	1,896,918	1,633,578	1,633,578
			Panel C. Mother high school dropout		
Expansion × above median competition	3.109***	0.966	5.404***	-1.589***	-0.193**
	(0.762)	(0.923)	(0.833)	(0.477)	(0.095)
Mean [SD] of Y	-43.279 [88.547]	-42.509 [97.770]	-43.525 [95.370]	21.728 [41.240]	6.455 [7.169]
Observations	1,511,217	1,499,602	1,616,155	1,341,389	1,341,389
			Panel D. Mother high school graduate		
Expansion × above median competition	1.775***	-0.213	3.740***	-0.252	-0.176***
	(0.598)	(0.728)	(0.640)	(0.281)	(0.051)
Mean [SD] of Y	1.336 [87.366]	1.001 [94.301]	1.438 [94.687]	12.627 [33.215]	4.869 [5.413]
Observations	3,754,631	3,723,846	4,004,160	3,318,165	3,318,165
			Panel E. Mother college graduate		
Expansion × above median competition	1.161	-0.324	2.009**	-0.113	-0.153***
	(0.875)	(1.104)	(0.927)	(0.286)	(0.056)
Mean [SD] of Y	65.699 [82.758]	66.072 [88.959]	65.192 [91.139]	4.160 [19.967]	3.355 [3.789]
Observations	921,715	908,430	990,752	794,099	794,099
			Panel F. White mother		
Expansion × above median competition	0.592	0.015	1.077	-0.095	-0.163**
	(0.692)	(0.865)	(0.683)	(0.296)	(0.066)
Mean [SD] of Y	22.957 [88.068]	22.440 [93.939]	23.308 [96.293]	10.113 [30.150]	5.253 [5.726]
Observations	3,170,846	3,146,594	3,362,580	2,828,887	2,828,887
			Panel G. African-American mother		
Expansion × above median competition	1.087	0.867	1.721**	0.879	0.136
	(0.833)	(1.054)	(0.865)	(0.594)	(0.090)
Mean [SD] of Y	-48.852 [88.588]	-49.836 [98.533]	-47.516 [94.446]	25.448 [43.557]	5.310 [6.478]
Observations	1,420,501	1,408,927	1,518,874	1,254,177	1,254,177
			Panel H. Hispanic mother		
Expansion × above median competition	2.828***	-0.552	6.013***	-1.736***	-0.055
	(0.895)	(1.120)	(1.005)	(0.389)	(0.075)
Mean [SD] of Y	-6.396 [88.173]	-4.972 [95.269]	-7.637 [95.336]	10.491 [30.643]	4.440 [5.187]
Observations	1,475,615	1,457,395	1,599,177	1,267,259	1,267,259
			Panel I. Mother born in the U.S.		
Expansion × above median competition	0.358	-0.763	1.312**	-0.056	-0.161***
	(0.588)	(0.740)	(0.578)	(0.293)	(0.059)
Mean [SD] of Y	-1.250 [93.922]	-1.907 [100.711]	-0.701 [100.901]	14.836 [35.545]	5.371 [6.004]
Observations	4,759,439	4,722,232	5,070,104	4,217,517	4,217,517
			Panel J. Foreign born mother		
Expansion × above median competition	0.602	-2.469**	3.567***	-0.967***	-0.018
	(0.852)	(1.090)	(0.950)	(0.365)	(0.061)
Mean [SD] of Y	4.284 [90.192]	6.389 [97.308]	2.307 [96.939]	9.529 [29.361]	3.905 [4.802]
Observations	1,428,124	1,409,646	1,540,963	1,236,136	1,236,136

Note: Specifications are based on those in panel C of Table 3 with the baseline sample split by child's free or reduced price lunch history (panels A and B), maternal education (panels C to E), race/ethnicity (panels F to H), and maternal immigration status (panels I and J). Outcome variables are averaged mathematics and reading test scores (column 1), mathematics test scores (column 2), reading test scores (column 3), likelihood of being suspended (column 4), and absence rate (column 5). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A6: Heterogeneity in the effects of voucher expansion: Churches measure

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
			Panel A. Ever on free or reduced price lunch		
Expansion × above median competition	5.335***	2.900***	7.698***	-2.330***	-0.378***
	(0.640)	(0.768)	(0.694)	(0.338)	(0.062)
Mean [SD] of Y	-21.420 [89.403]	-21.405 [97.323]	-21.028 [96.392]	17.271 [37.800]	5.585 [6.310]
Observations	4,379,915	4,341,810	4,714,149	3,820,075	3,820,075
			Panel B. Never on free or reduced price lunch		
Expansion × above median competition	3.926***	2.423***	4.789***	-1.125***	-0.110**
	(0.738)	(0.928)	(0.756)	(0.255)	(0.053)
Mean [SD] of Y	51.995 [80.557]	51.918 [86.455]	52.259 [89.053]	5.124 [22.048]	3.762 [4.042]
Observations	1,807,648	1,790,068	1,896,918	1,633,578	1,633,578
			Panel C. Mother high school dropout		
Expansion × above median competition	4.594***	2.705***	6.433***	-2.832***	-0.305***
	(0.779)	(0.945)	(0.870)	(0.487)	(0.096)
Mean [SD] of Y	-43.279 [88.547]	-42.509 [97.770]	-43.525 [95.370]	21.728 [41.240]	6.455 [7.169]
Observations	1,511,217	1,499,602	1,616,155	1,341,389	1,341,389
			Panel D. Mother high school graduate		
Expansion × above median competition	4.648***	2.085***	7.097***	-1.346***	-0.273***
	(0.602)	(0.743)	(0.637)	(0.286)	(0.050)
Mean [SD] of Y	1.336 [87.366]	1.001 [94.301]	1.438 [94.687]	12.627 [33.215]	4.869 [5.413]
Observations	3,754,631	3,723,846	4,004,160	3,318,165	3,318,165
			Panel E. Mother college graduate		
Expansion × above median competition	3.281***	1.512	3.943***	-0.761***	-0.051
	(0.883)	(1.127)	(0.929)	(0.288)	(0.057)
Mean [SD] of Y	65.699 [82.758]	66.072 [88.959]	65.192 [91.139]	4.160 [19.967]	3.355 [3.789]
Observations	921,715	908,430	990,752	794,099	794,099
			Panel F. White mother		
Expansion × above median competition	4.031***	3.013***	4.683***	-1.382***	-0.094
	(0.665)	(0.828)	(0.652)	(0.283)	(0.066)
Mean [SD] of Y	22.957 [88.068]	22.440 [93.939]	23.308 [96.293]	10.113 [30.150]	5.253 [5.726]
Observations	3,170,846	3,146,594	3,362,580	2,828,887	2,828,887
			Panel G. African-American mother		
Expansion × above median competition	3.406***	3.159***	4.197***	-1.322**	-0.043
	(0.872)	(1.132)	(0.880)	(0.672)	(0.095)
Mean [SD] of Y	-48.852 [88.588]	-49.836 [98.533]	-47.516 [94.446]	25.448 [43.557]	5.310 [6.478]
Observations	1,420,501	1,408,927	1,518,874	1,254,177	1,254,177
			Panel H. Hispanic mother		
Expansion × above median competition	4.666***	1.455	7.624***	-1.686***	-0.145*
	(0.975)	(1.210)	(1.082)	(0.421)	(0.080)
Mean [SD] of Y	-6.396 [88.173]	-4.972 [95.269]	-7.637 [95.336]	10.491 [30.643]	4.440 [5.187]
Observations	1,475,615	1,457,395	1,599,177	1,267,259	1,267,259
			Panel I. Mother born in the U.S.		
Expansion × above median competition	2.481***	1.130	3.490***	-1.239***	-0.231***
	(0.591)	(0.746)	(0.570)	(0.293)	(0.058)
Mean [SD] of Y	-1.250 [93.922]	-1.907 [100.711]	-0.701 [100.901]	14.836 [35.545]	5.371 [6.004]
Observations	4,759,439	4,722,232	5,070,104	4,217,517	4,217,517
			Panel J. Foreign born mother		
Expansion × above median competition	2.476***	-0.838	5.590***	-1.104***	-0.080
	(0.931)	(1.176)	(1.012)	(0.386)	(0.067)
Mean [SD] of Y	4.284 [90.192]	6.389 [97.308]	2.307 [96.939]	9.529 [29.361]	3.905 [4.802]
Observations	1,428,124	1,409,646	1,540,963	1,236,136	1,236,136

Note: Specifications are based on those in panel D of Table 3 with the baseline sample split by child's free or reduced price lunch history (panels A and B), maternal education (panels C to E), race/ethnicity (panels F to H), and maternal immigration status (panels I and J). Outcome variables are averaged mathematics and reading test scores (column 1), mathematics test scores (column 2), reading test scores (column 3), likelihood of being suspended (column 4), and absence rate (column 5). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A7: Heterogeneity in the effects of voucher expansion: Slots measure

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
			Panel A. Ever on free or reduced price lunch		
Expansion × above median competition	7.481***	4.384***	10.515***	-2.334***	-0.427***
	(0.614)	(0.755)	(0.655)	(0.337)	(0.063)
Mean [SD] of Y	-21.409 [89.396]	-21.402 [97.315]	-21.010 [96.385]	17.305 [37.829]	5.587 [6.314]
Observations	4,362,211	4,324,143	4,696,426	3,803,417	3,803,417
			Panel B. Never on free or reduced price lunch		
Expansion × above median competition	4.392***	2.949***	5.655***	-1.325***	-0.133**
	(0.774)	(0.981)	(0.775)	(0.255)	(0.054)
Mean [SD] of Y	51.933 [80.552]	51.840 [86.435]	52.216 [89.054]	5.147 [22.095]	3.763 [4.044]
Observations	1,798,314	1,780,746	1,887,588	1,624,568	1,624,568
			Panel C. Mother high school dropout		
Expansion × above median competition	6.144***	3.651***	8.629***	-2.533***	-0.355***
	(0.749)	(0.924)	(0.835)	(0.480)	(0.096)
Mean [SD] of Y	-43.261 [88.539]	-42.496 [97.764]	-43.506 [95.362]	21.774 [41.271]	6.460 [7.175]
Observations	1,504,461	1,492,865	1,609,399	1,334,914	1,334,914
			Panel D. Mother high school graduate		
Expansion × above median competition	6.464***	3.504***	9.376***	-1.549***	-0.324***
	(0.594)	(0.749)	(0.616)	(0.287)	(0.052)
Mean [SD] of Y	1.324 [87.359]	0.979 [94.291]	1.434 [94.682]	12.655 [33.247]	4.870 [5.415]
Observations	3,739,944	3,709,186	3,989,457	3,304,238	3,304,238
			Panel E. Mother college graduate		
Expansion × above median competition	3.459***	1.634	4.881***	-0.967***	-0.031
	(0.932)	(1.174)	(0.957)	(0.297)	(0.058)
Mean [SD] of Y	65.640 [82.760]	65.993 [88.952]	65.155 [91.140]	4.182 [20.017]	3.356 [3.792]
Observations	916,120	902,838	985,158	788,833	788,833
			Panel F. White mother		
Expansion × above median competition	5.523***	4.367***	6.528***	-1.717***	-0.255***
	(0.672)	(0.854)	(0.646)	(0.282)	(0.063)
Mean [SD] of Y	22.894 [88.067]	22.368 [93.933]	23.256 [96.293]	10.145 [30.193]	5.257 [5.730]
Observations	3,156,514	3,132,289	3,348,248	2,815,187	2,815,187
			Panel G. African-American mother		
Expansion × above median competition	5.482***	4.467***	6.961***	-0.810	0.008
	(0.816)	(1.076)	(0.806)	(0.660)	(0.097)
Mean [SD] of Y	-48.816 [88.572]	-49.807 [98.517]	-47.479 [94.430]	25.486 [43.578]	5.311 [6.483]
Observations	1,414,642	1,403,085	1,513,010	1,248,570	1,248,570
			Panel H. Hispanic mother		
Expansion × above median competition	5.986***	1.721	9.875***	-2.658***	-0.208**
	(0.993)	(1.241)	(1.081)	(0.435)	(0.087)
Mean [SD] of Y	-6.380 [88.165]	-4.972 [95.255]	-7.608 [95.334]	10.518 [30.678]	4.441 [5.190]
Observations	1,469,512	1,451,296	1,593,065	1,261,615	1,261,615
			Panel I. Mother born in the U.S.		
Expansion × above median competition	4.585***	2.904***	6.065***	-1.427***	-0.295***
	(0.588)	(0.754)	(0.558)	(0.292)	(0.059)
Mean [SD] of Y	-1.281 [93.900]	-1.947 [100.685]	-0.724 [100.883]	14.870 [35.579]	5.374 [6.008]
Observations	4,739,008	4,701,848	5,049,671	4,198,015	4,198,015
			Panel J. Foreign born mother		
Expansion × above median competition	3.259***	-0.614	6.952***	-1.859***	-0.141*
	(0.946)	(1.194)	(1.031)	(0.401)	(0.074)
Mean [SD] of Y	4.272 [90.182]	6.359 [97.294]	2.311 [96.932]	9.558 [29.402]	3.907 [4.805]
Observations	1,421,517	1,403,041	1,534,343	1,229,970	1,229,970

Note: Specifications are based on those in panel E of Table 3 with the baseline sample split by child's free or reduced price lunch history (panels A and B), maternal education (panels C to E), race/ethnicity (panels F to H), and maternal immigration status (panels I and J). Outcome variables are averaged mathematics and reading test scores (column 1), mathematics test scores (column 2), reading test scores (column 3), likelihood of being suspended (column 4), and absence rate (column 5). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A8: Robustness of the estimates: Diversity measure

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
			Panel A. Baseline		
Expansion × above median competition	4.233***	1.937***	6.539***	-0.504*	-0.265***
	(0.599)	(0.736)	(0.618)	(0.267)	(0.052)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
			Panel B. Continuous competition measure		
Expansion × competition	1.020***	0.534***	1.534***	-0.179***	-0.071***
	(0.129)	(0.166)	(0.127)	(0.053)	(0.011)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
			Panel C. Log funding expansion measure		
Expansion × above median competition	3.940***	1.945***	5.777***	-0.554**	-0.292***
	(0.561)	(0.693)	(0.559)	(0.268)	(0.052)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
			Panel D. Competition measure unweighted with elementary to middle school flows		
Expansion × above median competition	3.965***	1.960***	5.861***	-0.200	-0.197***
	(0.572)	(0.713)	(0.589)	(0.252)	(0.047)
Mean [SD] of Y	0.107 [92.883]	0.118 [99.757]	0.026 [99.869]	13.240 [33.892]	5.008 [5.703]
Observations	5,843,075	5,795,031	6,213,961	5,186,087	5,186,087
			Panel E. Weights based solely on pre-program transitions between grades 5 and 6		
Expansion × above median competition	4.317***	1.918***	6.792***	-0.368	-0.225***
	(0.579)	(0.712)	(0.603)	(0.264)	(0.052)
Mean [SD] of Y	0.027 [93.105]	0.000 [100.000]	0.000 [100.000]	13.620 [34.301]	5.035 [5.789]
Observations	6,098,353	6,043,455	6,514,413	5,377,233	5,377,233
			Panel F. Limiting the sample to schools with at least one competitor within 5 miles		
Expansion × above median competition	2.992***	1.064	4.810***	-0.387	-0.183***
	(0.604)	(0.753)	(0.628)	(0.274)	(0.053)
Mean [SD] of Y	0.033 [93.144]	0.000 [100.000]	0.000 [100.000]	13.600 [34.279]	5.006 [5.776]
Observations	5,730,114	5,677,273	6,119,166	5,055,539	5,055,539
			Panel G. School years 2002/03 to 2011/12		
Expansion × above median competition	3.911***	1.912***	5.976***	-0.504*	-0.265***
	(0.592)	(0.735)	(0.621)	(0.267)	(0.052)
Mean [SD] of Y	-0.190 [92.937]	0.000 [99.999]	0.000 [99.999]	13.633 [34.314]	5.039 [5.786]
Observations	5,361,423	5,349,158	5,348,374	5,453,653	5,453,653
			Panel H. Balanced panel (6-years)		
Expansion × above median competition	4.786***	3.053***	6.271***	-0.770**	-0.356***
	(0.666)	(0.825)	(0.690)	(0.306)	(0.055)
Mean [SD] of Y	0.121 [92.953]	0.001 [99.999]	0.001 [99.999]	14.491 [35.201]	4.967 [5.547]
Observations	3,975,445	3,936,192	5,325,819	2,857,336	2,857,336
			Panel I. In sample standardized test scores		
Expansion × above median competition	4.124***	1.767**	6.480***	N/A	N/A
	(0.590)	(0.731)	(0.611)		
Mean [SD] of Y	0.090 [92.607]	0.229 [99.846]	-0.049 [99.823]		
Observations	5,783,177	5,783,177	5,783,177		
			Panel J. In population standardized scores sample		
Expansion × above median competition	4.160***	1.946***	6.374***	N/A	N/A
	(0.559)	(0.698)	(0.581)		
Mean [SD] of Y	4.698 [89.085]	4.695 [96.395]	4.702 [95.627]		
Observations	5,783,177	5,783,177	5,783,177		

Note: Robustness checks based on those in panel A of Table 3. Panel A replicates the main result from panel A of Table 3; panel B replaces dummy indicator for above median pre-reform competition with continuous measure; panel C replaces logarithm of number of scholarships expansion measure with logarithm of funding; panel D assigns the middle-school pre-policy competition measures based on the actual grade 6 (middle) school initially attended by each student; panel E generates expected competitive pressure measures for middle school-aged students using only pre-policy announcement flow between elementary and middle schools; panel F limits the initial sample to only schools with at least one competitor within 5 miles; panel G restricts the sample to school years 2002/03 to 2011/12 where we observe all five outcomes; panel H restricts the sample to 6-year panel of observations starting with grade 3 and within school years available for a given variable; panels I and J restrict the sample to school years 2002/03 to 2012/13 where we observe test scores that are standardized for the full population of Florida students - panel I presents our in-sample standardization while panel J population-level standardization. Outcome variables are averaged mathematics and reading test score (column 1), mathematics test score (column 2), reading test score (column 3), likelihood of being suspended (column 4), and absence rate (column 5). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A9: Robustness of the estimates: Density measure

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
			Panel A. Baseline		
Expansion × above median competition	5.293***	2.817***	7.566***	-1.109***	-0.258***
	(0.586)	(0.728)	(0.608)	(0.268)	(0.052)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
			Panel B. Continuous competition measure		
Expansion × competition	0.209***	0.080***	0.342***	-0.063***	-0.014***
	(0.023)	(0.029)	(0.025)	(0.009)	(0.002)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
			Panel C. Log funding expansion measure		
Expansion × above median competition	4.909***	2.735***	6.666***	-1.179***	-0.290***
	(0.549)	(0.686)	(0.550)	(0.270)	(0.051)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
			Panel D. Competition measure unweighted with elementary to middle school flows		
Expansion × above median competition	4.982***	2.658***	7.100***	-0.712***	-0.271***
	(0.561)	(0.705)	(0.580)	(0.252)	(0.047)
Mean [SD] of Y	0.107 [92.883]	0.118 [99.757]	0.026 [99.869]	13.240 [33.892]	5.008 [5.703]
Observations	5,843,075	5,795,031	6,213,961	5,186,087	5,186,087
			Panel E. Weights based solely on pre-program transitions between grades 5 and 6		
Expansion × above median competition	5.789***	3.337***	8.213***	-1.151***	-0.224***
	(0.574)	(0.713)	(0.591)	(0.266)	(0.052)
Mean [SD] of Y	0.027 [93.105]	0.000 [100.000]	0.000 [100.000]	13.620 [34.301]	5.035 [5.789]
Observations	6,098,353	6,043,455	6,514,413	5,377,233	5,377,233
			Panel F. Limiting the sample to schools with at least one competitor within 5 miles		
Expansion × above median competition	4.324***	2.184***	6.317***	-1.295***	-0.178***
	(0.613)	(0.754)	(0.639)	(0.282)	(0.053)
Mean [SD] of Y	0.033 [93.144]	0.000 [100.000]	0.000 [100.000]	13.600 [34.279]	5.006 [5.776]
Observations	5,730,114	5,677,273	6,119,166	5,055,539	5,055,539
			Panel G. School years 2002/03 to 2011/12		
Expansion × above median competition	4.768***	2.700***	6.868***	-1.109***	-0.258***
	(0.582)	(0.729)	(0.610)	(0.268)	(0.052)
Mean [SD] of Y	-0.190 [92.937]	0.000 [99.999]	0.000 [99.999]	13.633 [34.314]	5.039 [5.786]
Observations	5,361,423	5,349,158	5,348,374	5,453,653	5,453,653
			Panel H. Balanced panel (6-years)		
Expansion × above median competition	5.813***	3.850***	7.243***	-1.277***	-0.361***
	(0.652)	(0.814)	(0.682)	(0.308)	(0.054)
Mean [SD] of Y	0.121 [92.953]	0.001 [99.999]	0.001 [99.999]	14.491 [35.201]	4.967 [5.547]
Observations	3,975,445	3,936,192	5,325,819	2,857,336	2,857,336
			Panel I. In sample standardized test scores		
Expansion × above median competition	5.017***	2.634***	7.399***	N/A	N/A
	(0.580)	(0.725)	(0.600)		
Mean [SD] of Y	0.090 [92.607]	0.229 [99.846]	-0.049 [99.823]		
Observations	5,783,177	5,783,177	5,783,177		
			Panel J. In population standardized scores sample		
Expansion × above median competition	5.040***	2.814***	7.266***	N/A	N/A
	(0.549)	(0.692)	(0.570)		
Mean [SD] of Y	4.698 [89.085]	4.695 [96.395]	4.702 [95.627]		
Observations	5,783,177	5,783,177	5,783,177		

Note: Robustness checks based on those in panel B of Table 3. Panel A replicates the main result from panel B of Table 3; panel B replaces dummy indicator for above median pre-reform competition with continuous measure; panel C replaces logarithm of number of scholarships expansion measure with logarithm of funding; panel D assigns the middle-school pre-policy competition measures based on the actual grade 6 (middle) school initially attended by each student; panel E generates expected competitive pressure measures for middle school-aged students using only pre-policy announcement flow between elementary and middle schools; panel F limits the initial sample to only schools with at least one competitor within 5 miles; panel G restricts the sample to school years 2002/03 to 2011/12 where we observe all five outcomes; panel H restricts the sample to 6-year panel of observations starting with grade 3 and within school years available for a given variable; panels I and J restrict the sample to school years 2002/03 to 2012/13 where we observe test scores that are standardized for the full population of Florida students - panel I presents our in-sample standardization while panel J population-level standardization. Outcome variables are averaged mathematics and reading test score (column 1), mathematics test score (column 2), reading test score (column 3), likelihood of being suspended (column 4), and absence rate (column 5). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A10: Robustness of the estimates: Distance measure

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
Panel A. Baseline					
Expansion × above median competition	1.648***	-0.308	3.542***	-0.430*	-0.151***
	(0.590)	(0.712)	(0.622)	(0.261)	(0.052)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
Panel B. Continuous competition measure					
Expansion × competition	0.887***	0.532***	1.246***	-0.126**	-0.071***
	(0.148)	(0.201)	(0.140)	(0.061)	(0.013)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
Panel C. Log funding expansion measure					
Expansion × above median competition	1.528***	-0.280	3.174***	-0.408	-0.167***
	(0.551)	(0.670)	(0.561)	(0.264)	(0.052)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
Panel D. Competition measure unweighted with elementary to middle school flows					
Expansion × above median competition	1.623***	0.223	3.009***	-0.672***	-0.109**
	(0.553)	(0.692)	(0.573)	(0.237)	(0.045)
Mean [SD] of Y	0.107 [92.883]	0.118 [99.757]	0.026 [99.869]	13.240 [33.892]	5.008 [5.703]
Observations	5,843,075	5,795,031	6,213,961	5,186,087	5,186,087
Panel E. Weights based solely on pre-program transitions between grades 5 and 6					
Expansion × above median competition	2.270***	-0.017	4.678***	-0.276	-0.136***
	(0.568)	(0.687)	(0.607)	(0.257)	(0.051)
Mean [SD] of Y	0.027 [93.105]	0.000 [100.000]	0.000 [100.000]	13.620 [34.301]	5.035 [5.789]
Observations	6,098,353	6,043,455	6,514,413	5,377,233	5,377,233
Panel F. Limiting the sample to schools with at least one competitor within 5 miles					
Expansion × above median competition	1.245**	0.145	2.345***	-0.700**	-0.050
	(0.602)	(0.733)	(0.641)	(0.273)	(0.051)
Mean [SD] of Y	0.033 [93.144]	0.000 [100.000]	0.000 [100.000]	13.600 [34.279]	5.006 [5.776]
Observations	5,730,114	5,677,273	6,119,166	5,055,539	5,055,539
Panel G. School years 2002/03 to 2011/12					
Expansion × above median competition	1.213**	-0.275	2.755***	-0.430*	-0.151***
	(0.590)	(0.717)	(0.629)	(0.261)	(0.052)
Mean [SD] of Y	-0.190 [92.937]	0.000 [99.999]	0.000 [99.999]	13.633 [34.314]	5.039 [5.786]
Observations	5,361,423	5,349,158	5,348,374	5,453,653	5,453,653
Panel H. Balanced panel (6-years)					
Expansion × above median competition	1.861***	0.232	3.022***	-0.530*	-0.170***
	(0.662)	(0.815)	(0.691)	(0.304)	(0.054)
Mean [SD] of Y	0.121 [92.953]	0.001 [99.999]	0.001 [99.999]	14.491 [35.201]	4.967 [5.547]
Observations	3,975,445	3,936,192	5,325,819	2,857,336	2,857,336
Panel I. In sample standardized test scores					
Expansion × above median competition	1.573***	-0.265	3.412***	N/A	N/A
	(0.584)	(0.708)	(0.616)		
Mean [SD] of Y	0.090 [92.607]	0.229 [99.846]	-0.049 [99.823]		
Observations	5,783,177	5,783,177	5,783,177		
Panel J. In population standardized scores sample					
Expansion × above median competition	1.720***	0.019	3.420***	N/A	N/A
	(0.555)	(0.677)	(0.589)		
Mean [SD] of Y	4.698 [89.085]	4.695 [96.395]	4.702 [95.627]		
Observations	5,783,177	5,783,177	5,783,177		

Note: Robustness checks based on those in panel C of Table 3. Panel A replicates the main result from panel C of Table 3; panel B replaces dummy indicator for above median pre-reform competition with continuous measure; panel C replaces logarithm of number of scholarships expansion measure with logarithm of funding; panel D assigns the middle-school pre-policy competition measures based on the actual grade 6 (middle) school initially attended by each student; panel E generates expected competitive pressure measures for middle school-aged students using only pre-policy announcement flow between elementary and middle schools; panel F limits the initial sample to only schools with at least one competitor within 5 miles; panel G restricts the sample to school years 2002/03 to 2011/12 where we observe all five outcomes; panel H restricts the sample to 6-year panel of observations starting with grade 3 and within school years available for a given variable; panels I and J restrict the sample to school years 2002/03 to 2012/13 where we observe test scores that are standardized for the full population of Florida students - panel I presents our in-sample standardization while panel J population-level standardization. Outcome variables are averaged mathematics and reading test score (column 1), mathematics test score (column 2), reading test score (column 3), likelihood of being suspended (column 4), and absence rate (column 5). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A11: Robustness of the estimates: Churches measure

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
			Panel A. Baseline		
Expansion × above median competition	3.917***	1.643**	5.966***	-1.428***	-0.223***
	(0.598)	(0.727)	(0.626)	(0.265)	(0.051)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
			Panel B. Continuous competition measure		
Expansion × competition	0.012***	0.003	0.021***	-0.006***	-0.001***
	(0.003)	(0.003)	(0.003)	(0.001)	(0.000)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
			Panel C. Log funding expansion measure		
Expansion × above median competition	3.639***	1.518**	5.336***	-1.401***	-0.245***
	(0.559)	(0.684)	(0.566)	(0.266)	(0.051)
Mean [SD] of Y	0.027 [93.104]	0.000 [100.000]	0.000 [100.000]	13.633 [34.314]	5.039 [5.786]
Observations	6,187,563	6,131,878	6,611,067	5,453,653	5,453,653
			Panel D. Competition measure unweighted with elementary to middle school flows		
Expansion × above median competition	3.650***	1.483**	5.517***	-1.411***	-0.242***
	(0.570)	(0.710)	(0.590)	(0.249)	(0.046)
Mean [SD] of Y	0.107 [92.883]	0.118 [99.757]	0.026 [99.869]	13.240 [33.892]	5.008 [5.703]
Observations	5,843,075	5,795,031	6,213,961	5,186,087	5,186,087
			Panel E. Weights based solely on pre-program transitions between grades 5 and 6		
Expansion × above median competition	3.557***	1.149	5.904***	-1.203***	-0.141***
	(0.576)	(0.707)	(0.607)	(0.260)	(0.051)
Mean [SD] of Y	0.027 [93.105]	0.000 [100.000]	0.000 [100.000]	13.620 [34.301]	5.035 [5.789]
Observations	6,098,353	6,043,455	6,514,413	5,377,233	5,377,233
			Panel F. Limiting the sample to schools with at least one competitor within 5 miles		
Expansion × above median competition	2.421***	0.479	4.108***	-1.290***	-0.120**
	(0.610)	(0.743)	(0.644)	(0.278)	(0.051)
Mean [SD] of Y	0.033 [93.144]	0.000 [100.000]	0.000 [100.000]	13.600 [34.279]	5.006 [5.776]
Observations	5,730,114	5,677,273	6,119,166	5,055,539	5,055,539
			Panel G. School years 2002/03 to 2011/12		
Expansion × above median competition	3.356***	2.022***	4.728***	-1.428***	-0.223***
	(0.593)	(0.730)	(0.627)	(0.265)	(0.051)
Mean [SD] of Y	-0.190 [92.937]	0.000 [99.999]	0.000 [99.999]	13.633 [34.314]	5.039 [5.786]
Observations	5,361,423	5,349,158	5,348,374	5,453,653	5,453,653
			Panel H. Balanced panel (6-years)		
Expansion × above median competition	4.304***	2.768***	5.107***	-1.475***	-0.231***
	(0.662)	(0.815)	(0.698)	(0.303)	(0.054)
Mean [SD] of Y	0.121 [92.953]	0.001 [99.999]	0.001 [99.999]	14.491 [35.201]	4.967 [5.547]
Observations	3,975,445	3,936,192	5,325,819	2,857,336	2,857,336
			Panel I. In sample standardized test scores		
Expansion × above median competition	3.530***	1.615**	5.445***	N/A	N/A
	(0.590)	(0.723)	(0.618)		
Mean [SD] of Y	0.090 [92.607]	0.229 [99.846]	-0.049 [99.823]		
Observations	5,783,177	5,783,177	5,783,177		
			Panel J. In population standardized scores sample		
Expansion × above median competition	3.757***	2.046***	5.469***	N/A	N/A
	(0.559)	(0.690)	(0.588)		
Mean [SD] of Y	4.698 [89.085]	4.695 [96.395]	4.702 [95.627]		
Observations	5,783,177	5,783,177	5,783,177		

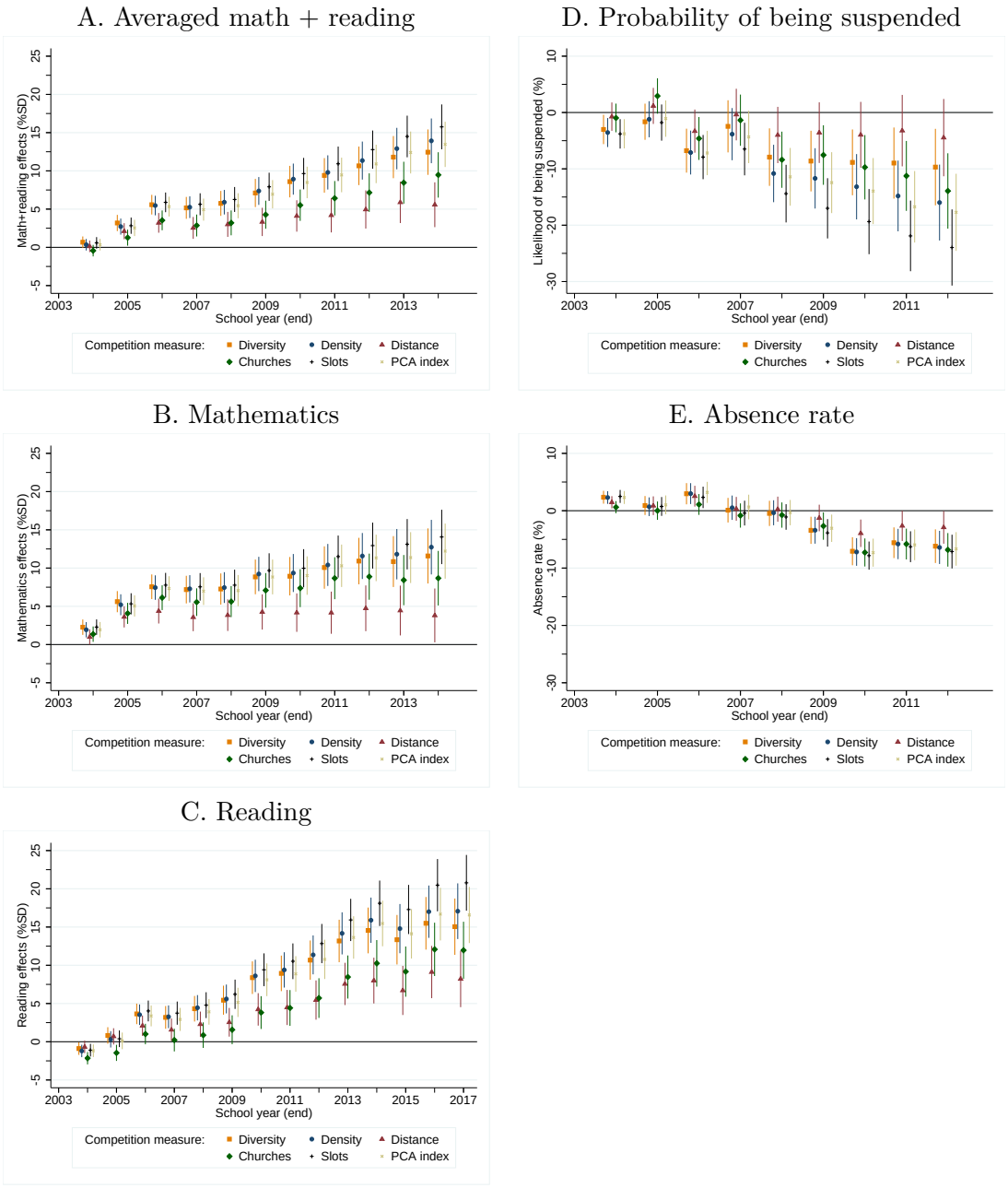
Note: Robustness checks based on those in panel D of Table 3. Panel A replicates the main result from panel D of Table 3; panel B replaces dummy indicator for above median pre-reform competition with continuous measure; panel C replaces logarithm of number of scholarships expansion measure with logarithm of funding; panel D assigns the middle-school pre-policy competition measures based on the actual grade 6 (middle) school initially attended by each student; panel E generates expected competitive pressure measures for middle school-aged students using only pre-policy announcement flow between elementary and middle schools; panel F limits the initial sample to only schools with at least one competitor within 5 miles; panel G restricts the sample to school years 2002/03 to 2011/12 where we observe all five outcomes; panel H restricts the sample to 6-year panel of observations starting with grade 3 and within school years available for a given variable; panels I and J restrict the sample to school years 2002/03 to 2012/13 where we observe test scores that are standardized for the full population of Florida students - panel I presents our in-sample standardization while panel J population-level standardization. Outcome variables are averaged mathematics and reading test score (column 1), mathematics test score (column 2), reading test score (column 3), likelihood of being suspended (column 4), and absence rate (column 5). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A12: Robustness of the estimates: Slots measure

	(1) Math + Reading	(2) Mathematics	(3) Reading	(4) Suspensions	(5) Absences
			Panel A. Baseline		
Expansion × above median competition	6.064***	3.361***	8.684***	-1.749***	-0.290***
	(0.590)	(0.732)	(0.604)	(0.266)	(0.052)
Mean [SD] of Y	0.000 [93.085]	-0.038 [99.977]	-0.017 [99.984]	13.666 [34.349]	5.041 [5.790]
Observations	6,160,525	6,104,889	6,584,014	5,427,985	5,427,985
			Panel B. Continuous competition measure		
Expansion × competition	1.164***	0.491***	1.871***	-0.349***	-0.060***
	(0.102)	(0.128)	(0.108)	(0.043)	(0.008)
Mean [SD] of Y	0.000 [93.085]	-0.038 [99.977]	-0.017 [99.984]	13.666 [34.349]	5.041 [5.790]
Observations	6,160,525	6,104,889	6,584,014	5,427,985	5,427,985
			Panel C. Log funding expansion measure		
Expansion × above median competition	5.631***	3.227***	7.728***	-1.859***	-0.314***
	(0.552)	(0.688)	(0.546)	(0.267)	(0.052)
Mean [SD] of Y	0.000 [93.085]	-0.038 [99.977]	-0.017 [99.984]	13.666 [34.349]	5.041 [5.790]
Observations	6,160,525	6,104,889	6,584,014	5,427,985	5,427,985
			Panel D. Competition measure unweighted with elementary to middle school flows		
Expansion × above median competition	4.700***	2.040***	7.305***	-1.315***	-0.266***
	(0.571)	(0.716)	(0.585)	(0.251)	(0.048)
Mean [SD] of Y	-0.003 [92.865]	-0.002 [99.755]	-0.072 [99.844]	13.223 [33.874]	5.011 [5.706]
Observations	5,761,773	5,714,711	6,123,884	5,117,781	5,117,781
			Panel E. Weights based solely on pre-program transitions between grades 5 and 6		
Expansion × above median competition	5.964***	3.140***	8.952***	-1.554***	-0.211***
	(0.571)	(0.709)	(0.589)	(0.264)	(0.053)
Mean [SD] of Y	0.002 [93.084]	-0.036 [99.975]	-0.014 [99.984]	13.653 [34.336]	5.038 [5.793]
Observations	6,071,801	6,016,952	6,487,847	5,351,967	5,351,967
			Panel F. Limiting the sample to schools with at least one competitor within 5 miles		
Expansion × above median competition	5.406***	2.873***	7.876***	-1.781***	-0.192***
	(0.612)	(0.757)	(0.631)	(0.284)	(0.053)
Mean [SD] of Y	0.003 [93.124]	-0.042 [99.976]	-0.019 [99.984]	13.636 [34.317]	5.008 [5.781]
Observations	5,703,761	5,650,965	6,092,805	5,030,501	5,030,501
			Panel G. School years 2002/03 to 2011/12		
Expansion × above median competition	5.485***	3.349***	7.670***	-1.749***	-0.290***
	(0.586)	(0.733)	(0.610)	(0.266)	(0.052)
Mean [SD] of Y	-0.219 [92.915]	-0.040 [99.974]	-0.019 [99.981]	13.666 [34.349]	5.041 [5.790]
Observations	5,336,140	5,323,917	5,323,137	5,427,985	5,427,985
			Panel H. Balanced panel (6-years)		
Expansion × above median competition	6.372***	4.225***	8.304***	-1.938***	-0.368***
	(0.659)	(0.821)	(0.678)	(0.304)	(0.056)
Mean [SD] of Y	0.095 [92.933]	-0.037 [99.974]	-0.015 [99.986]	14.524 [35.234]	4.969 [5.551]
Observations	3,958,889	3,919,656	5,303,632	2,845,185	2,845,185
			Panel I. In sample standardized test scores		
Expansion × above median competition	5.752***	3.223***	8.282***	N/A	N/A
	(0.582)	(0.728)	(0.596)		
Mean [SD] of Y	0.061 [92.587]	0.190 [99.823]	-0.069 [99.805]		
Observations	5,756,691	5,756,691	5,756,691		
			Panel J. In population standardized scores sample		
Expansion × above median competition	5.746***	3.357***	8.135***	N/A	N/A
	(0.551)	(0.695)	(0.568)		
Mean [SD] of Y	4.669 [89.064]	4.657 [96.371]	4.681 [95.610]		
Observations	5,756,691	5,756,691	5,756,691		

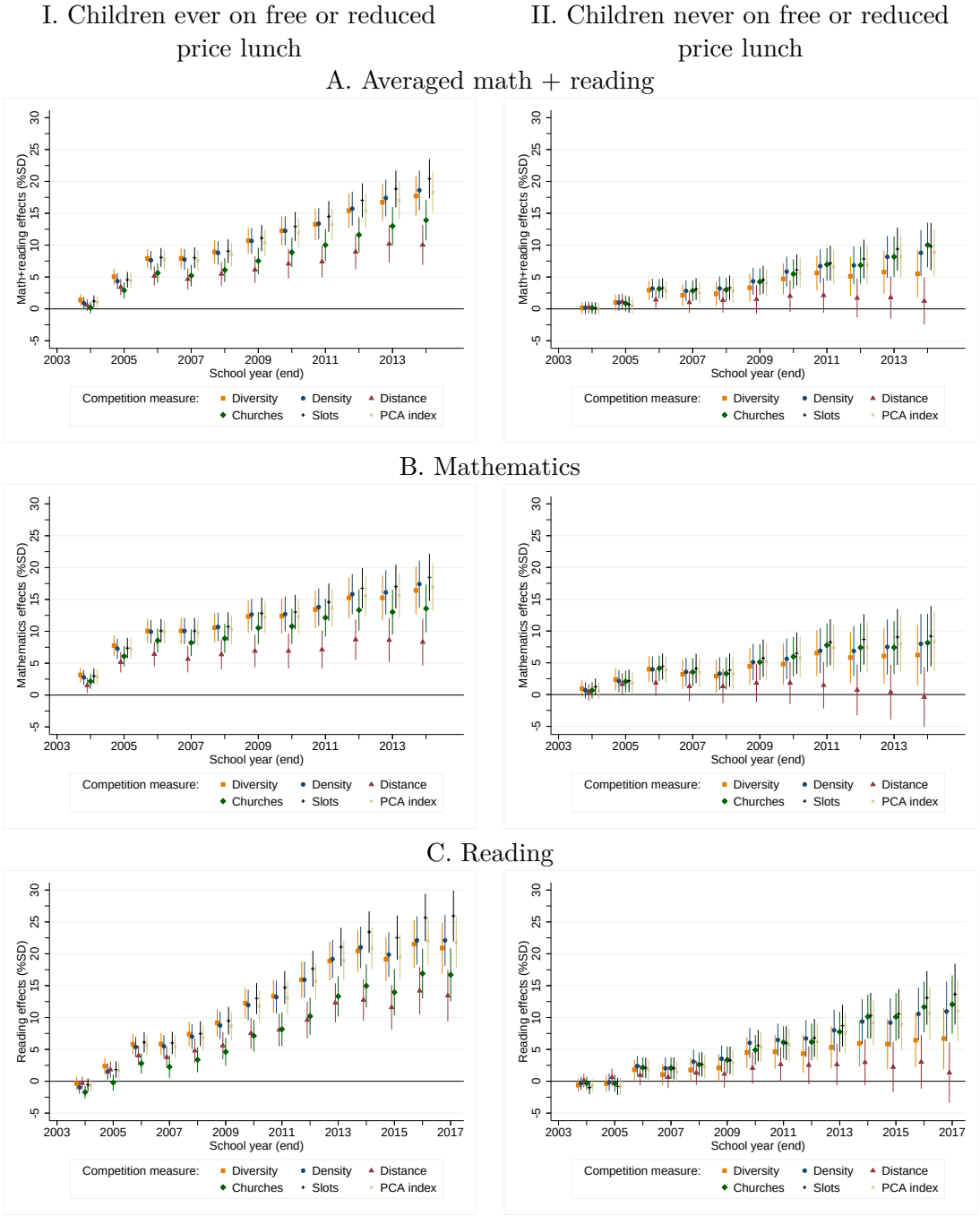
Note: Robustness checks based on those in panel E of Table 3. Panel A replicates the main result from panel E of Table 3; panel B replaces dummy indicator for above median pre-reform competition with continuous measure; panel C replaces logarithm of number of scholarships expansion measure with logarithm of funding; panel D assigns the middle-school pre-policy competition measures based on the actual grade 6 (middle) school initially attended by each student; panel E generates expected competitive pressure measures for middle school-aged students using only pre-policy announcement flow between elementary and middle schools; panel F limits the initial sample to only schools with at least one competitor within 5 miles; panel G restricts the sample to school years 2002/03 to 2011/12 where we observe all five outcomes; panel H restricts the sample to 6-year panel of observations starting with grade 3 and within school years available for a given variable; panels I and J restrict the sample to school years 2002/03 to 2012/13 where we observe test scores that are standardized for the full population of Florida students - panel I presents our in-sample standardization while panel J population-level standardization. Outcome variables are averaged mathematics and reading test score (column 1), mathematics test score (column 2), reading test score (column 3), likelihood of being suspended (column 4), and absence rate (column 5). Standard errors are clustered at grade one school level. Point estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Figure A1: Effects of voucher expansion over school years



Note: These figures plot estimates from the specifications estimated in panels A to F of Table 3 where instead of interaction between competition measures and log number of scholarships we plot competition measures interacted with school years, and with baseline omitted year 2002/2003. Outcomes are averaged test scores in mathematics and reading (panel A); mathematics test scores (panel B); and reading test scores (panel C); likelihood of being suspended (panel D); and absence rate (panel E). Competition measures are: number of denominational types represented (orange square); number of local private schools (navy circle); miles to nearest private school competitor (maroon triangles); number of churches, synagogues, and mosques (green diamonds); number of private school slots per grade (black pluses); and composite index of all five measures (khaki exes). Spikes present 95% confidence intervals based on standard errors clustered at grade one school level.

Figure A2: Effects of voucher expansion over school years by family socioeconomic status: Test scores



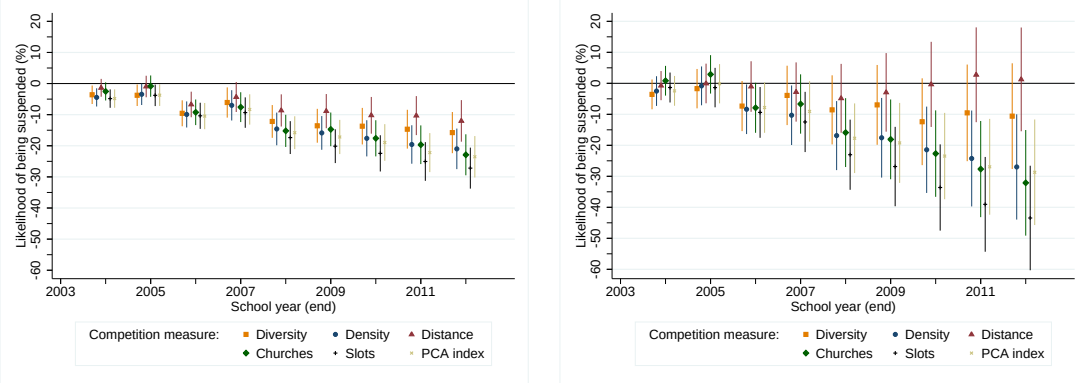
Note: These figures plot estimates from the specifications estimated in panels A and B of Tables 4 and A3 to A7, where instead of interaction between competition measures and log number of scholarships we plot competition measures interacted with school years, and with baseline omitted year 2002/2003. Outcomes are averaged test scores in mathematics and reading (panel A); mathematics test scores (panel B); and reading test scores (panel C). Competition measures are: number of denominational types represented (orange square); number of local private schools (navy circle); miles to nearest private school competitor (maroon triangles); number of churches, synagogues, and mosques (green diamonds); number of private school slots per grade (black pluses); and composite index of all five measures (khaki exes). Sample is divided into children from families ever observed on free or reduced price lunch (column 1) and those never observed on free or reduced price lunch (column 2). Spikes present 95% confidence intervals based on standard errors clustered at grade one school level.

Figure A3: Effects of voucher expansion over school years by family socioeconomic status: Disciplinary outcomes

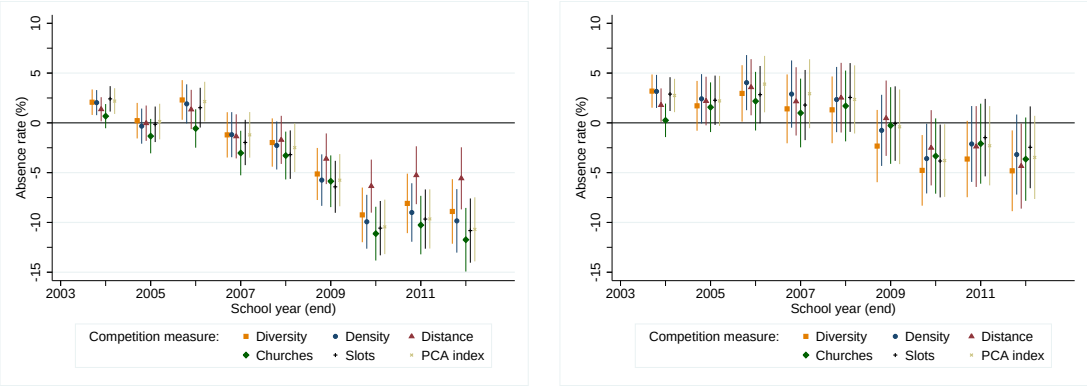
I. Children ever on free or reduced price lunch

II. Children never on free or reduced price lunch

A. Probability of being suspended

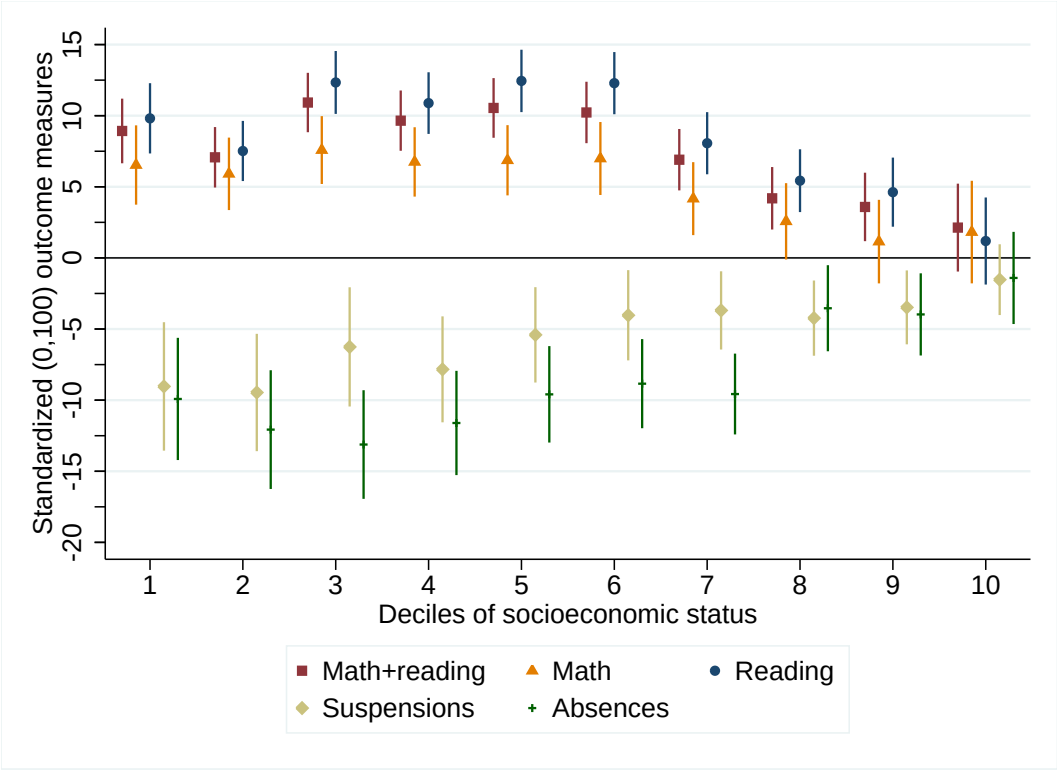


B. Absence rate



Note: These figures plot estimates from the specifications estimated in panels A and B of Tables 4 and A3 to A7, where instead of interaction between competition measures and log number of scholarships we plot competition measures interacted with school years, and with baseline omitted year 2002/2003. Outcomes are likelihood of being suspended (panel A) and absence rate (panel B). Competition measures are: number of denominational types represented (orange square); number of local private schools (navy circle); miles to nearest private school competitor (maroon triangles); number of churches, synagogues, and mosques (green diamonds); number of private school slots per grade (black pluses); and composite index of all five measures (khaki exes). Sample is divided into children from families ever observed on free or reduced price lunch (column 1) and those never observed on free or reduced price lunch (column 2). Spikes present 95% confidence intervals based on standard errors clustered at grade one school level.

Figure A4: Effects of voucher expansion: Heterogeneity by socioeconomic status index



Note: This figure plots heterogeneity estimates for the main specification estimated in panel F of Table 3. These are computed separately for each outcome and each decile of socioeconomic status distribution. SES index is computed as first component from Principal Components Analysis (PCA) of maternal years of education, marital status, maternal age at birth, indicator for Medicaid paid birth, and zip code neighborhood income at the time of birth. Sample is restricted to births between 1994 and 2002. Outcomes are averaged test scores in mathematics and reading (maroon squares); mathematics test scores (orange triangles); reading test scores (navy circles); likelihood of being suspended (khaki diamonds); and absence rate (green pluses). Each outcome variable is standardized in its empirical sample to have mean zero and standard deviation of 100. Spikes present 95% confidence intervals based on standard errors clustered at grade one school level.