# Targeted Acceleration in Middle School Math: Impacts on College Entry, Degree Completion, and STEM

David Card and Laura Giuliano*

October 2024

**Abstract**

The traditional course sequence for mathematics in middle and high school leaves many students ill-prepared to enter selective engineering and science programs. Attempts in several states to push all students through an accelerated mathematics curriculum starting in middle school have faltered; the latest Common Core curriculum postpones the teaching of algebra until 9th grade. We study a *targeted* multi-year accelerated curriculum that enables students to finish Algebra I and Geometry before high school. Entry to the program is based on 5th-grade scores in the statewide math test, with a cutoff at the 80th percentile, enabling a highly credible regression discontinuity analysis of the impacts of the program. Following students for 13 years after entry, we show that the accelerated curriculum has significant positive effects on the probabilities that marginally eligible girls enter selective colleges, complete a bachelor's degree within 5 years, and graduate in STEM or business/economics. The share of girls with a degree in STEM nearly doubles (from 6% to 11%), with a comparable effect on degrees in business and economics. In contrast, we find negligible effects on boys, implying that the accelerated program brings girls up to the level of similar-scoring boys in completion of degrees in the most highly-paid fields.

---

* Card: University of California, Berkeley and NBER, card@berkeley.edu; Giuliano: University of California, Santa Cruz and NBER, lgiulian@ucsc.edu. We are extremely grateful to Cynthia Park for assistance in accessing and interpreting the data used in this study, and to Tatiana Reyes for extraordinary research assistance. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through grant R305A190175 to the NBER. The opinions expressed are those of the authors and do not represent views of the U.S. Department of Education.

The issues of how math courses should be sequenced, and at what grade to begin algebra, are highly controversial. They also matter: UC Berkeley's College of Engineering, for example, tells prospective students that "*Four years of mathematics, especially including calculus*" demonstrates the necessary aptitude in science and math to be considered for admission.[1] Unfortunately, the traditional U.S. high school math sequence of Algebra I, Geometry, Algebra II, and Pre-calculus delays exposure to advanced math until 12th grade, and omits calculus altogether. The conflict between the "hidden curriculum" required by many selective college programs and the traditional math curriculum has led to much debate and wide policy swings. In the 2000's several states including California began promoting "Algebra for All" in 8th grade.[2] Just a few years later, this effort was abandoned. More recently some districts have tried to ***delay*** algebra for all students until 9th grade, citing concerns that tracking exacerbates inequities based on race and family income (see e.g., Huffaker et al. 2024 and Huffaker, 2023). Nevertheless, demand for more advanced math in middle school is strong (Yoder, 2024), and enrollment in Algebra and Geometry in 7th and 8th grades remains relatively high in many states.[3]

In this paper we study the impacts of a targeted accelerated math curriculum available in the middle schools of a large urban school district in Florida ("the District"). The curriculum, known as "Great Explorations in Mathematics" (GEM) includes pre-algebra in 6th grade, high-school algebra in 7th grade, and high-school geometry in 8th grade, effectively accelerating students 2 years beyond the traditional curriculum, and 1 year beyond the other "advanced" tracks offered in many schools, including those in the District and in other districts in Florida. Eligibility for the GEM program is based on scores in the state's 5th-grade math test, with a threshold at roughly the 80th percentile. Thus, marginally eligible participants – for whom the causal effects of the program are identified in a regression discontinuity (RD) analysis – are

[1] This information is from the front page of the website https://engineering.berkeley.edu/admissions/undergraduate-admissions/ retrieved August 20, 2024.
[2] The National Mathematics Advisory Panel (2008) recommended that "School districts should ensure that all prepared students have access to an authentic algebra course and should prepare more students than at present to enroll in such a course by Grade 8. See Domina et al. (2015) for an analysis of the California initiative, which is widely perceived as having hurt lower-ability students. Similar findings are reported by Clotfelter et al. (2015) from an analysis of an 8th grade algebra initiative in North Carolina in the early 2000s.
[3] For example, Peters and Carter (2023, Table 2) report that average enrollment rates in Geometry in 8th grade in 2017-28 were around 18% in Utah, 13% in Virginia, and 8-9% in Florida, Maryland, Minnesota, and Washington. At the opposite extreme the rate was below 0.5% in Arkansas, Alabama, Louisiana, Mississippi, and West Virginia.

relatively well qualified.[4] Indeed, most students who barely miss the GEM threshold will enter an advanced math track in middle school.

We use information on middle and high school courses, combined with information from the National Student Clearinghouse on college enrollment and degree completion, to study the effects of entry to the GEM accelerated mathematics sequence in 6th grade. Focusing on students whose 5th-grade scores are within a narrow bandwidth ($\pm 0.5\sigma's$) around the entry threshold, we follow 16,000 students for at least 13 years after 5th grade (i.e., allowing up to 6 years after they would normally graduate from high school). This relatively large sample allows us to conduct an informative RD analysis and make meaningful comparisons of the effect of the GEM program on college entry and degree completion along two key dimensions that have been identified in previous work: gender, and participation in the free or reduced price lunch (FRL) program.[5]

For students as a whole, we find that entry to the GEM accelerated program leads to large gains in the probability of taking algebra and geometry in middle school (+20 percentage points (ppt) and +43 ppt respectively); a smaller but still significant +7 ppt rise in the probability of taking and passing calculus in high school; and a marginally significant +4.5 ppt effect on the probability of obtaining a degree in STEM or business/economics fields within 5 years after high school. Importantly, ***all*** of the effects on college-related outcomes are attributable to girls. Entry to GEM leads to a near doubling of the probability that girls complete a BA in a STEM field (from around 5% to 10%), pushing their rate up to that of boys. There is a similar-sized effect on degrees in business or economics (from 6% to 10%), again catching up to the level of boys. In contrast, we find only small impacts on college-related outcomes of boys, centered on zero. We also see very similar effects of entry to GEM for FRL versus non-FRL students.

A potential concern with an RD-based analysis of GEM is that students who barely miss the threshold experience a causal *reduction* in their academic performance, arising from the removal of better-performing peers from their math classes, or from discouragement effects associated with missing the GEM cutoff.[6] To address this, we use data from another large urban

[4] The closest antecedent to our work is a recent study by Dougherty et al. (2017), which evaluates test-based admission to Algebra in 8th grade for students in Wake County NC. The threshold for this program is set at about the 20th percentile of the state test.
[5] Dougherty et al. (2017) find that admission to algebra in 8th grade has a relatively large and significant effect on the probability of taking pre-calculus in 11th grade for girls, but no effect for boys. They also find that the gains are concentrated among higher income students and are negligible for lower-income students.
[6] It is also possible that students who barely miss the threshold experience a causal increase in performance arising from their new status as the top students in their math class – see Denning et al. (2023) for example.

school district in the same state that has no GEM program. Comparing boys and girls with similar 5th grade math scores in the two districts, we find no evidence of negative peer effects or discouragement effects for girls (or boys) with scores just below the GEM threshold in the District, but gains for girls who score just above the threshold relative to their peers in the comparison district, consistent with our RD estimates.

Our empirical findings contribute to two main literatures. The first is a relatively small set of studies that have tried to identify the causal effects of math acceleration on student outcomes. Research on the Algebra-for-All initiative in California (Domina et al, 2015) and on a related initiative in North Carolina (Clotfelter et al., 2015) showed that pushing all students into algebra in 8th grade was potentially helpful for better qualified students but definitely harmful for weaker students. Subsequent research by Dougherty et al. (2017) showed that a more targeted approach – leaving out the lowest-scoring quintile of students – leads to a positive effect on math course-taking of girls, including enrollment in pre-calculus in 11th grade, with little effect (positive or negative) on boys. They do not follow students to college. Cohodes (2020) examines an accelerated curriculum program for high-achievers in 4th-6th grades and finds positive effects on high school graduation and college entry. She does not follow students long enough to measure degree completion, nor does she present results separately by gender.[7]

We also contribute to the growing literature on gender differences in college major choice (see McNally 2020 for a recent review). Women and men choose different fields of study in college, with a lower share of females graduating in science, technology, engineering and math (STEM) fields, and in business-related fields. As noted by Altonji et al. (2016), graduates of these fields are among the most highly paid college workers, and observational comparisons suggest that differences in field of study can explain about one fifth of the gender earnings gap among college graduates in the US.[8] While there appears to be a correlation between math preparation in high school and entry to STEM fields in college (see e.g., Speer, 2017; Card and Payne, 2021), there is almost no causal evidence that *increases* in math preparation lead to a *rise*

[7] An older study by Altonji (1996) studied the effects of high school course taking on college completion and wages. He used the mean shares of students taking different courses in a students' high school as instrumental variables, potentially identifying the effects of access to course options. He found relatively small effects of math or science courses on college outcomes or wages.

[8] See Appendix Table A1, where we present a simple analysis of data on 31-35 year old college graduates in the 2010-2018 American Community Survey. We show that adding two dummies that indicate whether a college graduate has a degree in STEM or business-related fields reduces the gender gap in earnings by 14%.

in STEM majoring by young women.[9] Our results suggest that exposure to an accelerated math curriculum in middle school can help close the gender gap in both STEM and business-related degree attainment.

## I. Background: District Setting and the GEM Program

The District is one of the 10 largest school districts in the country, serving around 20,000 students per grade. Its student body is racially diverse, with roughly equal numbers of Hispanic, Black and White students, just under one-half of whom are eligible for the FRL program. During the past two decades the District operated approximately 140 elementary schools with grades K-5, about 80 mainstream middle schools offering grades 6-8, and 70 conventional high schools. It also has a wide variety of charter and special service schools.

The District's mainstream middle schools offer three primary tracks of mathematics instruction (taught in separate classrooms): regular, advanced and GEM. There is also a course stream with additional instructional time for weaker students. As shown in the flowchart in Figure 1, both the regular and advanced streams follow the standard math curriculum prescribed by the State of Florida in grades 6-8.[10] Students in the advanced track who perform well in 7th-grade math, however, can take Algebra I in 8th grade. In addition, since the 2000-2001 school year, all middle schools have offered the Great Explorations in Mathematics (GEM) track (also taught in separate classrooms), which compresses the standard 6th and 7th-grade course content into 6th grade, and offers Algebra I in 7th grade and Geometry in 8th grade. The availability of GEM in all its mainstream middle schools (including those located in lower-income and minority neighborhoods) reflects the District's long-term effort to expand access to advanced academic programs for all children.

To qualify for GEM, students must have a 5th-grade score of 380 or above in Math and 300 or above in Reading on the Florida Comprehensive Assessment Tests (FCAT's).[11] Participation also requires a teacher's recommendation and parental consent. Students have to

[9] Owen (2023) conducts an RCT among college students informing them of their relative performance on STEM-antecedent courses but finds no effect on major choice.
[10] The content of the regular and advanced courses is based on a curricular planning system operated by the State of Florida, known as CPALMS. The "advanced" track courses in the Flowchart (which was assembled by the District) correspond to "accelerated" courses in current CPALMS terminology.
[11] The statewide tests have changed over time but the GEM cutoff in math has remained at about the 80th percentile of the math test, while the minimum standard for the reading test has remained at around the median.

maintain at least a B grade to stay in the program in later years, so entry to GEM does not guarantee that they will complete Geometry before high school. It does, however, provide an alternate pathway to completing Algebra 1 by the end of middle school, as shown in Figure 1. In contrast, students who miss the cutoff for GEM but enter the advanced track in sixth grade can be accelerated by at most one grade level to Algebra 1 in 8th grade.

## II. Empirical Framework

### *a. Regression Discontinuity Model*

We use a "fuzzy RD" approach to analyze the effect of entering GEM in 6th grade on subsequent student outcomes (Lee and Lemieux, 2010). Specifically, we estimate a locally linear first-stage model for entering GEM, with the 5th-grade math score as a running variable and a potential jump at the GEM cutoff. We estimate a parallel set of reduced-form models for outcomes like college enrollment and degree attainment. Under standard assumptions (e.g., Hahn, Todd and van der Klaauw, 2001), the ratio of the estimated discontinuity in the outcome variable at the GEM cutoff to the estimated discontinuity in the fraction of students who enter GEM provides a causal estimate of the effect of GEM entry on students whose GEM entry status switches as their math score crosses the 380 point threshold.

Specifically, our first-stage model takes the form:

$$GEM6_i = \pi \cdot 1[M_i \geq 380] + g(M_i) + u_i, \qquad (1)$$

where $GEM6_i$ is an indicator that student $i$ entered GEM in 6th grade and $g(M_i)$ is a piece-wise linear function of the student's math score $(M_i)$ that allows a slope change at $M_i = 380$:

$$g(M_i) = g_0 + g_1(M_i - 380) + g_2(M_i - 380) \times 1[M_i \geq 380].$$

We estimate (1) using data for students with math scores in a symmetric bandwidth BW around the 380 point threshold, i.e., using students with scores in the interval $[380 - BW, 380 + BW]$.

Likewise, we estimate reduced-form models for the jump in an outcome of interest, $Y_i$, at the 380-point threshold of the form:

$$Y_i = \delta \cdot 1[M_i \geq 380] + h(M_i) + v_i, \qquad (2)$$

where $h(M_i)$ is a piece-wise linear function analogous to $g(M_i)$. Our IV estimator of the treatment effect of gifted status on the mean outcome of the compliers is $\beta = \delta/\pi$. We construct this using a two-stage least squares (2SLS) estimator for the second-stage model with GEM entry in 6th grade on the right-hand side:

$$Y_i = \beta \cdot GEM6_i + f(M_i) + e_i, \qquad (3)$$

treating the indicator $1[M_i \geq 380]$ as an instrumental variable, and parameterizing $f(M_i)$ in the same way as $g(.)$ and $h(.)$. Our baseline models have no controls other than cohort fixed effects (4 dummies) to allow for trends over time in GEM participation and college outcomes, but we show that combinations of baseline characteristics that predict our outcomes trend smoothly at the 380-point threshold, and we also present models with controls.

For our main analysis, we focus on a bandwidth of 25 points, which is equivalent to $\pm 0.5\sigma's$ (standard deviation units) of the statewide math test. The first-stage and reduced-form relationships in our data are all relatively close to linear, so (as we show below) our estimates of $\delta$ and $\pi$ are relatively stable across bandwidths. We also present estimates using alternative bandwidths, including the bandwidth selected by the algorithm of Calonico, Cattaneo and Titiunik (2014) (hereafter, CCT): in most cases these are close to 25.

The entry rules for GEM specify that the student has a score of 380 or higher on the $5^{th}$-grade statewide math test <u>and</u> 300 or more on the $5^{th}$-grade reading test. The reading requirement introduces a source of "fuzziness" to the first stage model (1), since not all students who pass the 380 point math threshold have an eligible reading score, but it does not invalidate our basic approach. In fact, over 90% of students who barely satisfy the math test criterion have reading scores of 300 or above, so the requirement excludes relatively few students. It is slightly more constraining for boys than girls, and for FRL students than non-FRL students, leading to slightly smaller first stage effects of passing the math threshold for boys and FRL recipients.

***b. Compliers***

The IV estimator in our fuzzy RD design identifies the average treatment effect of entry to GEM for the subset of *compliers* whose GEM entry status changes at the math score cutoff. We measure the characteristics of the compliers using the method of Abadie (2002), applied to an RD setting. Specifically, we estimate the mean of characteristic $X$ among the compliers by the coefficient $\mu_x$ from a 2SLS regression of the form:

$$X_i \cdot GEM6_i = \mu_x \cdot GEM6_i + k(M_i) + \zeta_i, \quad (4)$$

using $1[M_i \geq 380]$ as an instrumental variable for $GEM6_i$, and including a piece-wise linear control function $k(M_i)$ similar to the ones included in equations (1)-(3). We use the same approach to estimate the mean potential outcomes $Y_i(1)$ and $Y_i(0)$ for compliers. In the former case we fit a 2SLS model with $Y_i \cdot GEM6_i$ as the dependent variable, and interpret the estimated coefficient of the endogenous regressor $GEM6_i$ as an estimate of the mean of $Y_i(1)$. In the latter case we use $Y_i \cdot (1 - GEM6_i)$ as the dependent variable and switch the sign of the 2SLS estimate of the coefficient on $GEM6_i$ to estimate the mean of $Y_i(0)$.

## III. Data and Analysis Sample

### *a. Data Sources*

We have access to anonymized student-level data from the District's centralized record system for children who completed 5$^{th}$ grade between the 2003 and 2007 academic years. These records contain information on gender, age, race/ethnicity, and FRL/ELL status. They also include state-wide achievement test scores in reading and math for grades 3-8; course enrollments for grades 5-12, and course grades for middle and high school.

We obtain information on college enrollment and degree completion from data reported to the District by the National Student Clearinghouse (NSC). These data are only available for students who completed high school in the District. As we discuss in the next subsection, however, the probability that students stay in the District long enough to have college data is quite smooth through the 380 point math score threshold, so there is no evidence of a sample selection bias in our RD estimates of the effect of GEM entry.

### *b. Analysis Sample*

Appendix Table A2 summarizes the derivation of our analysis sample. We start with 98,709 children who were in 5$^{th}$ grade in the District in the academic years 2003-2007. As shown in column 1 of the table, these children are about one-half female, 43% participated in the FRL program, and their racial composition is 35% White non-Hispanic ("White"), 34% Black non-Hispanic ("Black"), and 24% Hispanic. We then eliminate the 20% of grade 5 attendees who were not also in grades 4 and 6 in the District (column 2) – which has little effect on the characteristics of the sample. Next, in column 3, we eliminate students who did not attend a traditional middle school in the District. (These students were mainly attending charter schools). This cuts the sample by another 7% and again leaves the sample means largely unchanged.

Next we eliminate the roughly 30% of students who were not still in the District 6 years after entering 6 grade (column 4). This has a relatively small effect on the sample characteristics, and leaves us with around 51,000 students. Finally we drop the roughly 3,800 gifted students in the sample, yielding an analysis sample of 47,439 (column 5).[12] Aside from having a slightly higher share of girls (52% vs. 49%), the characteristics of this sample are very similar, on average, to those of the overall pool. About 14% participated in GEM and 31% participated in another advanced math track math class in 6th grade.

A potential concern with our analysis sample is that students who get admitted to GEM may be more likely to stay in the District and have college data, leading to differential selection biases on the two sides of the 380 point threshold for GEM entry. We investigate this issue in Appendix Figure A1, where we plot the probability of being included in the analysis sample from among the population of all 5th graders against their 5th-grade math scores. Reassuringly, the probability of sample inclusion shows no jump at the threshold: the estimated discontinuity at 380 points is -0.013 with a standard error of 0.010.

For our "*RD analysis sample*", shown in column 6, we focus on students who were in the analysis sample described in column 5 and had a 5th-grade math score within 25 points of the GEM threshold of 380 points. This subset of 16,357 students has a nearly 50-50 gender composition but its 31% FRL rate is lower than the District average of 43%. The sample over-represents White students and over-represents Black students relative to the District averages, reflecting racial disparities in test scores. Not surprisingly, students with scores near the GEM cutoff had higher math (+0.56σ's) and reading scores (+0.44σ's) in 4th grade than others in the analysis sample. About 28% ultimately entered GEM in 6th grade. Most of the others (in all, 55% of the RD analysis sample) were in the advanced track math class in 6th grade; only 18% were in the basic math stream.

The last four columns of the table show the mean characteristics of students in our RD analysis sample separately by gender and FRL status. Most characteristics are similar across gender and FRL lines, though 4th-grade math scores are slightly higher for boys and 4th-grade reading scores are higher for girls, while test scores in both domains are lower for FRL students.

[12] We exclude gifted students because, as discussed in Card, Chyn and Giuliano (2024) these students are often allowed to participate in GEM even if they miss the 380 point math threshold. In addition, they receive other services that could confound the interpretation of their entry to GEM.

## IV. Empirical Results

### *a. Validity of RD Design and First Stage*

We begin our main empirical analysis with Figure 2. The left panel of this figure probes the potential validity of our RD design by plotting the *predicted* probability of on-time high school graduation and completion of a college degree (associate or bachelor's) within 5 years against 5$^{th}$-grade math scores of students in our RD analysis sample. We use 1-point bins in this and all subsequent plots, reflecting the granularity of the underlying FCAT scores. In the RD analysis sample we have around 300 observations per bin, so the confidence interval for a bin mean extends $\pm 6$ percentage points (ppt) when the mean is close to 0.5. The prediction model uses pre-5$^{th}$-grade characteristics of students, including their 4$^{th}$-grade math and reading scores on the statewide tests, their demographics (gender, age as of 5$^{th}$ grade, FRL status) and the average characteristics of students in their 5$^{th}$-grade school (average test scores in 4$^{th}$-grade and average FRL status).

As discussed by Lee (2008), in a valid RD design there should be no discontinuity in the distribution of predetermined characteristics of sample members at the RD threshold: thus, we are looking for any evidence of a jump in the predicted degree completion rate at 380 points. The predicted degree completion rate rises steadily with 5$^{th}$-grade scores, suggesting that there is power to detect abnormal shifts in student characteristics around the 380 point threshold. The estimated local linear polynomial fit, however, shows virtually no jump. We also constructed similar figures by gender and FRL status. These are reported in Appendix Figure A2, and show no indication of discontinuous changes in student characteristics at the 380-point threshold.

We also looked for evidence of any discontinuity in the density of test scores around the threshold, as suggested by McCrary (2006). Plots of the density, shown in Appendix Figure A3, show no sign of bunching or other manipulation of the scores, which is expected given that the test is administered statewide and graded electronically.

The right panel of Figure 2 shows the first-stage relationship between 5$^{th}$-grade math scores and GEM entry. For reference, we also plot the relationship between test scores and the probability that a student either enters GEM ***or*** takes advanced math in 6$^{th}$ grade. Focusing on the first stage for GEM, we note three things. First, the relationships to the right and left of the 380 point cutoff are nearly linear, apart from a small upward trend in GEM participation for students with scores within 5 points of the threshold.[13] Second, among the students with scores near the cutoff, there are both 'GEM always takers' (as evidenced by the positive rate of GEM entry for students with scores just below the 380 cutoff) and 'GEM never takers' (as evidenced by the roughly 40 percent of students with scores just above the 380 cutoff who do not enter GEM). Third, there is a large jump in the probability of GEM entry – around 52 ppt in magnitude – implying that our RD design has a powerful first stage ($t \approx 15$ for the pooled sample).

Comparing the first stage for GEM entry to the RD model for either advanced math or GEM shows that most of the students who move into GEM when their score crosses the 380 point threshold come from the advanced math track. Only about 7 ppt of the 52 ppt total jump in the GEM enrollment share comes from students who would have been in the regular track.[14]

Appendix Figure A4 presents a parallel set of first-stage plots by gender and FRL status. All four of these figures look very similar to Figure 2B, with relatively low rates of GEM participation to the left of the 380 point threshold and large jumps in the probability of entering GEM at the cutoff, of roughly 50-55 ppt in magnitude. We also see that for all four subgroups, there is a much smaller rise (5-10 ppt) in the probability of GEM or advanced track math in 6th grade, implying that most of the students who enter GEM would otherwise have enrolled in the advanced track.

***b. Characteristics of Compliers***

Table 1 shows the characteristics of the compliers for the GEM entry cutoff, estimated using the approach of equation (4). Panel A shows baseline (pre-GEM) characteristics, while Panels B and C present estimated potential outcomes without the GEM treatment (i.e., the means of $Y_i(0)$), in middle and high school and in college, respectively. These mean potential outcomes – which we refer to the "untreated complier means" – are analogous to the means for the *control group* in a simple 2-arm RCT: they represent the counterfactual outcomes for the compliers who

---

[13] This appears to be driven by a practice of filling the 6th-grade GEM classroom with nearly-eligible students at schools where relatively few students meet the 380 threshold.

[14] It is possible that passing the 380 point threshold causes some students who would have been assigned to the regular track to be assigned to the advanced track (rather than being pushed all the way to GEM). This would violate the identification assumptions underlying the interpretation of equation (3), since it would mean that passing the threshold has some potential effect on students whose GEM status remains constant (the GEM never takers). The maximum size of the shift from regular to advanced track is 7 ppt (under the assumption that all the students moved out of regular classes moved to advanced classes at the 380 point threshold), which is 13% of the jump in GEM enrollment. If the causal effect of moving from regular to advanced track were the same size as the effect of moving to GEM this would lead to an upward bias of about 13% in our estimated causal effects of GEM entry.

are admitted to GEM and have realized outcomes $Y_i(1)$ – which we refer to as "treated complier means." For reference, the first column of the table shows the means of the same variables for all the students in our RD analysis sample (i.e., all those who graduated on time from the District and had a 5$^{th}$-grade math score between 355 and 405 points).

Looking first at the baseline characteristics in Panel A we see that on average, the compliers are broadly similar to students in the overall analysis sample. Looking within the complier population, however, there are some interesting differences by gender: female compliers appear to be somewhat less advantaged, with a higher shares of Black and FRL students (but a lower share of Hispanics). The girls also have nearly 0.1 $\sigma's$ higher 4$^{th}$-grade reading scores than the boys, while the boys are 0.05 $\sigma's$ ahead in their 4$^{th}$-grade math scores. Likewise, comparisons between FRL and non-FRL compliers show relatively big differences in their racial composition, reading scores, and school peers.

Turning to Panel B, we first show the 6$^{th}$-grade math course enrollments of the untreated compliers. As expected given our discussion of Figure 2B, 83% are in the advanced track math and 17% in the basic math track, with rates that do not vary too much by subgroup (though interestingly boys are more likely to be in the basic math track than girls despite their higher average 4$^{th}$ grade math scores). About 60% of untreated compliers complete algebra by the end of middle school, but only 2% also complete geometry, reflecting the lack of a pathway to geometry in the absence of GEM. Untreated compliers have fairly strong math GPA's (close to 3.5 in 6$^{th}$ grade, falling to 3.0 by 8$^{th}$ grade). As is often the case, girls have higher course grades than boys even though by construction, all the compliers have very similar 5$^{th}$-grade math scores.

About 23% of the untreated compliers took a calculus class in high school, with fairly similar rates for girls and boys and for non-FRL versus FRL. Among the calculus takers, a higher fraction of girls achieved a grade of B- or better (0.18/0.24=75%) than males (0.14/0.23=61%). Another 14% of the untreated compliers took AP calculus (or calculus in the International Baccalaureate, IB, program), with fairly similar rates across subgroups.

Nearly all high school students in the District take the PSAT, and about 90% of those in our RD analysis sample also took the SAT or ACT. Looking at PSAT math scores, the untreated compliers have an average score (in national percentile units) of 68, with a slightly higher mean for the boys than the girls (+2.6 percentiles).[15] Average percentile scores in the PSAT reading test are a little lower and are also 5 percentiles higher for girls.

Finally, turning to the post-secondary outcomes in Panel C, we see that 97% of the untreated compliers graduated high school on time, with similar rates by gender but a lower rate for FRL students. (Of course, our analysis sample requires that students were *enrolled* in 12th grade 6 years after entering 6th grade, so these high rates are partly mechanical.) 80% graduate on time and enter college the next year, with similar rates by gender but a relatively large gap (9 ppt) between FRL and non-FRL students. Just over 60% graduate on time and enter *selective* colleges, with a slightly higher rate for girls than boys, but a 13 ppt gap by FRL status. When we look at on-time college entry and persistence for at least a semester beyond the first year, we see the emergence of a relatively large gender gap (67% for girls versus 61% for boys) but no difference by FRL status.[16]

Our main outcomes of interest are degree completion and field of study. Among the untreated compliers, rates of degree completion and majoring in different fields closely parallel the overall means for our RD analysis sample: 51% complete a college degree within 5 years, 38% compete a bachelor's degree, and 16% complete a bachelor's degree in a STEM or business-related field. The degree completion rates are substantially higher for untreated complier females than males, and for non-FRL than FRL participants. Despite their +16 ppt advantage in completion of a bachelor's degree, however, girls are less likely to complete a degree in STEM or business-related fields. Indeed, only 28% (=0.13/0.46) of the untreated female compliers who complete a bachelor's degree within 5 years graduate in a STEM or business field, versus 70% (=0.21/0.30) of the untreated male compliers.

### *c. Impacts of GEM Entry on Math Courses in Middle and High School*

With this background, we turn to an analysis of the impacts of GEM entry on subsequent student outcomes, using the framework of equation (3). Figure 3 shows the reduced-form

[15] Note that boys have higher math PSAT's than girls, even though they have lower grades in their math classes (and are somewhat less likely to take advanced math classes in high school). This pattern is also present in other samples, such as the sample of gifted students from the District studied in Card, Chyn and Giuliano (2024), and has interesting implications for college admission policies that down-weight or ignore standardized test scores.

[16] The gender gap in college persistence is a national phenomenon. The 5-year degree completion rate for the 2015 cohort of men who started at all 4-year institutions was 57.5%, while the rate for women in the same cohort was 64% (Digest of Education Statistics, 2022, Table 326.1). The 3 year degree completion rate for the 2015 cohort of men who started at 2-year institutions was 29.6%, while the rate for women in the same cohort was 34.7% (Digest of Education Statistics, 2022, Table 326.2).

relationships between two main pre-college outcomes: whether the student completed algebra before high school, and whether she or he took calculus in high school. In panel A we see very clear jumps in the probability of taking algebra in middle school, and taking and passing the class, of around 10-12 percentage points. In panel B we see and somewhat less visually striking jumps in the probability of taking (or taking and passing) calculus, of around 5 percentage points. Figure 4 shows the same outcomes, focusing on successful completion of each course and cutting the sample by gender and FRL status. We see a slightly bigger reduced-form rise in completing algebra by 8$^{th}$ grade for girls than boys, and a much bigger gender disparity in the reduced forms for calculus. Indeed, all of the gains in calculus for the sample as a whole appear to be driven by girls. The reduced form models for both algebra before high school and calculus show somewhat larger effects for FRL participants than non-participants.

Table 2 presents a series of estimated RD models for our pooled sample and by gender and FRL status. Each column shows estimates for a different outcome, while the rows pertain to different subgroups. For each sample or subsample we show estimates from models that include only 4 fifth-grade cohort dummies, and from models that include the full set of control variables used in the construction of our predicted probability of on-time degree attainment (graphed in Figure 2A). We cluster standard errors by 6$^{th}$-grade middle school to allow for potential correlations in outcomes arising from school- or neighborhood-specific factors.

Column 1 presents the discontinuities in the *predicted* probability of on-time degree attainment. These are all small in magnitude, as expected given the patterns in Figure 2A and Appendix Figure A2, though there is a marginally significant 7/10$^{ths}$ ppt jump in the predicted probability of degree attainment for girls ($t \approx 1.7$). Column 2 presents the estimated first stage effects of passing the 380 point threshold on entering GEM in 6$^{th}$ grade (i.e., estimates of the parameter $\pi$ in equation 1). Consistent with the plots of the first stage relationships, these range from 0.48 to 0.55 in magnitude, and are highly significant. Even for our smallest subsample of FRL participants (N=5,076) we have a t-statistic for the first stage of about 15.8 (or an F-statistic of 249), well above the threshold proposed by Lee et al. (2022) that allows an analyst to use conventional confidence intervals for 2SLS models with no adjustment for the imprecision of the first stage.

Columns 3-13 present 2SLS estimates of the effect of entry to GEM on course participation and course grades. As shown in the first row of column 3, the overall effect on the probability of taking early algebra is 20 ppt. Given the 60% probability of taking algebra early for the untreated compliers, this implies that 80% of the treated compliers enrolled in algebra by 8th grade. The other 20% entered GEM class but failed to earn an A or B grade in 6th grade and were moved back to the advanced track course in 7th grade, where they did not perform well enough to enter algebra in 8th grade (see Figure 1). Comparing results by gender we see that the treatment effect of GEM entry on early algebra is higher for girls (+22 ppt) than boys (+16 ppt). Taken together with the means of the untreated compliers, these impacts imply that only 14% of girls who entered GEM failed to enroll in algebra by 8th grade, compared to 28% of boys.

The results in column 4 show that the overall effect of GEM entry on enrolling in geometry in 8th grade is 43 ppt, with a significantly larger effect for girls (52 ppt) than boys (33 ppt). Combing these treatment effects with the means in Table 1 and the effects of GEM on algebra enrollment, it is straightforward to calculate the distributions of middle school enrollment outcomes for the untreated and treated compliers across three categories: never enrolled in algebra, enrolled in algebra but not geometry, enrolled in both algebra and geometry.[17] These are summarized graphically in Figure 5a. Entry to GEM has a notably bigger effect on the middle school math experiences of complying girls (54% of whom end up enrolling in geometry after successfully completing algebra, and 32% of whom end up enrolling in algebra but not geometry) than complying boys (34% of whom end up enrolling in both classes and 38% of whom end up enrolling only in algebra).

Column 5 presents estimates of the effect of GEM entry on the probability of taking calculus in high school. For the overall sample we see a significant effect of +6.8 ppt (t=2.6). Comparing impacts by subgroups, the effect is concentrated among girls (+10.4 ppt), with only a small, insignificant effect on boys (+2.5 ppt). In contrast the gains in calculus enrollment are similar for FRL participants (+6.4 ppt) and non-participants (+6.8 ppt).

Columns 6-9 present 2SLS estimates in which the outcome variable is the event of taking high school algebra (columns 6-7) or high school calculus (columns 8-9) ***and*** achieving a certain grade. In interpreting these models it is important to recognize that entry to GEM can increase the share of students taking a given course (the extensive margin) and also affect the grades of students who would take the course regardless of GEM (an intensive margin effect on the

[17] Assuming that students who take geometry in 8th grade took algebra in 7th grade, the share of students who only enrolled in algebra equals the share who enrolled in algebra minus the share who enrolled in geometry.

"algebra always takers").[18] Thus the effect of GEM entry on the share of students achieving a certain grade (or better) in a given course can actually exceed the effect on the fraction taking the course. Indeed, the effect of GEM entry on the overall share of students taking algebra in middle school and earning a C- or better is +21.5 ppt, slightly larger than the 19.5 ppt effect on the share of students taking algebra. The 2 ppt gap implies that entry to GEM must have had some effect on grades for the 60% of students with scores near the GEM cutoff who would have taken algebra by 8th grade even if they narrowly missed the GEM threshold. Comparing the estimates in column 7 and column 3 we conclude that entry to GEM had intensive margin effects on algebra performance for all four subgroups. One explanation for this is that entrants to GEM who earn less than a B on their first attempt at algebra in 7th grade get a second chance (see Figure 1).

The presence of intensive margin effects is also apparent for high school calculus. Recall from Table 1 that 23% of the untreated compliers would take calculus. Entry to GEM raises that by 6.8 ppt on average, but raises the fraction who take calculus and earn a C- or better by 6.9 ppt. Similarly, there is a larger effect on taking calculus and earning C- or better than on just taking the class for boys and non-FRL participants. For FRL students, however, the effect on earning C- or better is slightly smaller than the effect on just enrolling in the class. Assuming that entry to GEM did not <u>lower</u> the course grades of FRL's who would have taken calculus even without GEM, these estimates imply that at most 90% (=0.058/0.064) of the extensive margin FRL's (those who took calculus because of GEM) finished the course with a C- or better.[19] While this might be seen as evidence of "mismatch," we note that the share of the *untreated* FRL compliers who took calculus and earned a C- or better is also 90% (see Table 1), so the extensive margin group may have performed similarly.

Figure 5b summarizes the relative impacts on high school calculus course taking and grade outcomes for female and male compliers. As in Figure 5a, we show the distributions for the untreated and treated compliers. The effect of GEM entry on girls is larger across the board, with a gain of 10 ppt in the share of girls with scores around the 380 threshold who complete calculus with a C- or better versus a 3 ppt for boys, and a gain of 5 ppt in the share of girls who have a B- or better, versus 3 ppt for boys.

---

[18] As noted, the untreated GEM compliers have about a 60% enrollment rate in algebra.
[19] This is an upper bound because participation in GEM may have raised the share of calculus always takers who earned C- or better.

Finally, columns 10-13 show effects of GEM entry on math GPA's in 6th through 8th grades, and in high school. For the overall sample and all subgroups there are large negative impacts of entering GEM on 6th grade math GPA's, on the order of −0.7 of a letter grade. There are also negative, but smaller, effects on 7th and 8th grade math GPA's. These impacts presumably arise because teachers award similar distributions of grades regardless of class difficulty: thus an untreated complier will tend to get a higher grade in advanced track math (where she is close to the top of the distribution of math ability) than a treated complier in the GEM class (where she is closer to the bottom). Importantly, by high school these negative impacts have largely disappeared. Even for boys, the effect is small in magnitude and far from statistical significance (t=0.3).

We have also investigated the effects of GEM entry in 6th grade on a variety of other high school-related outcomes, including enrollment in *any* advanced high school math class (pre-calculus, calculus, or statistics); enrollment in AP or IB calculus; completion of AP or IB calculus with a score of 3 or better (which is required for college credit); overall high school GPA; PSAT math and verbal scores; and whether the student took the SAT or ACT. The estimates for these outcomes are reported in Appendix Table 2.

In brief, we find effects on advanced math that are closely aligned with the effects on calculus, with a large positive boost for girls (+9.3 ppt, t=2.6) but a smaller, insignificant effect for boys (+5 ppt); and a large positive effect for FRL participants (+13 ppt, t=2.8) with a smaller boost for non-FRL participants (+6 ppt, t=2.2). For AP/IB calculus enrollment we find a modest positive overall effect (+4.7 ppt, t=2.2) that is a little larger for girls, but for completion of the class with a grade of 3 or better we find only small effects of around 2-3 ppt. Likewise, we find a small marginally significant positive effect on high school GPA for girls but a small negative effect for boys, similar to the effects on their GPA's in math.

Interestingly, we also find a significant and relatively large positive effect of GEM entry on girls' PSAT math scores (+3.5 percentiles, t=2.4). For boys, the effect on math PSAT is also positive (+1.4 percentiles) but insignificant. With these impacts, the PSAT math scores of treated complier girls are virtually the same as for treated complier boys, closing the gender gap that was present among the untreated compliers. We also find a marginally significant boost in the probability that girls take either the ACT or SAT (+3.1 ppt, t=1.6), suggesting that girls who entered GEM become more interested in applying to selective college programs (a result confirmed below). In contrast we see an insignificant negative effect for boys, and no large differences by FRL status.

***d. Impacts of GEM Entry on College Outcomes***

Next we turn to college-level outcomes. Panels A and B of Figures 6, 7, and 8 show the reduced form relationships between $5^{th}$-grade reading scores and 3 main outcomes: the probability of finishing high school on time and entering college within 1 year (Figure 6); the probability of earning any college degree within 5 years after the "normal" high school completion year (Figure 7); and the probability of earning a bachelor's degree in STEM or business-related fields within 5 years (Figure 8). In each case, panel A shows the reduced forms by gender, while panel B shows the reduced forms by FRL status. We discuss panels C and D below.

Looking first at on-time college entry in Figure 6, we see a small upward jump at the 380 point threshold for girls, a small downward jump for boys, and no jumps for FRL or non-FRL students. Turning to Figure 7, we see some indication of a positive jump in the probability of obtaining a degree within 5 years for girls, but no such effect for boys. The reduced-form plots by FRL status also show little evidence of a jump at the GEM threshold. The reduced-form plots in Figure 8 for the probability of obtaining a STEM or business degree show a clear jump for girls, but no effect for boys, while the jumps for both FRL and non-FRL participants are comparable but modest in size.

Table 3 summarizes our estimates of the effect of GEM entry on college-level outcomes, using the same format as Table 2. The first stages for all the underlying models are the same as the ones reported in column 2 of Table 2. Column 1 presents estimates of the effect of GEM entry on the probability of graduating high school 7 years after $5^{th}$ grade (i.e., on-time). None of the estimated effects are large or statistically significant. Column 2 presents models for on-time college entry (the same outcome shown in Figure 6). The overall effect (in row 1) is very close to 0, but this average reflects a slightly positive effect for girls and a slightly negative effect for boys (-5 ppt, t=1.3). We note that the corresponding RD estimate from a model with controls is smaller in magnitude, and suspect that 0 may be a good estimate for both genders.

It is also useful to check whether the estimates are sensitive to choice of bandwidth. (Recall that we are using a symmetric bandwidth of 25 points in our main tables.) We investigate this issue in the two lower panels of Figure 6, where we show the 2SLS estimates as we vary the bandwidth from 10 points to 40 points, along with their 95% confidence intervals. For reference we also highlight the bandwidth choice selected by the CCT algorithm (using a symmetric uniform kernel and CCT's "regularization" procedure), which is 30 for girls but 23 for boys. The pattern of estimates in panel C show that the estimates for boys tend to become slightly smaller in magnitude (i.e., less negative) with longer bandwidths, whereas those for girls are uniformly small (but positive). A similar exercise for FRL and non-FRL participants (panel D) yields estimates that center around 0 for both groups for most of the range of choices.

Returning to Table 3, column 3 shows estimates for models of the probability of enrolling at a selective college or university with a year of normal high school graduation time.[20] Here the overall effect is slightly positive, reflecting a relatively large positive effect for girls (+8.0 ppt, t=2.1) and a smaller negative effect for boys (−6 ppt, t=1.4). Again, the estimates from models with controls are smaller in magnitude, though the estimate for girls remains marginally significant (t=2.0), suggesting there may be a positive impact. The effects of GEM on selective college enrollment for FRL participants and non-participants are both small and statistically insignificant.

Many U.S. students who enter college stay for only a year or less (see e.g., DesJardin et al., 2002), leading to student debt but little economic benefit. Column 4 presents estimates of the effect of GEM on the probability that students enroll in college and persist for at least a year. (For reference, the mean of this outcome for the untreated compliers 64%). We estimate relatively small and statistically insignificant effects on this outcome for our overall sample and for the gender groups, but somewhat larger, opposite-signed effects for non-FRL participants (+5 ppt, t=1.8) and FRL participants (−8 ppt, t=1.2). When we look at the probabilities of obtaining any degree in 5 years (column 5) or a bachelor's degree in 5 years (column 6), however, the effects for both groups become smaller in magnitude.

Indeed, the estimates in column 6 suggest that GEM entry has little or no effect on the probability that students obtain a bachelor's degree. The effects for males are very close to 0, while the estimates for FRL and non-FRL subgroups are both slightly positive but statistically insignificant. As shown in Figure 7C, the estimated effect of GEM on the probability that girls obtain a bachelor's degree within 5 years is highly robust to bandwidth choice, while the effect

---

[20] The selective designation is based on the school's Carnegie Classification in the year the student enrolled. We note that 80% of the untreated compliers who enter college on time enter a selective college.

for boys is more negative at low bandwidths and then tends to 0. A similar exercise comparing effects at different bandwidths for FRL and non-FRL participants (Figure 7D) suggests that the effects for FRL participants may be positive while the effect for non-FRL's is close to 0.

Table 3, columns 7-11 present estimated models for the probability of obtaining bachelor's degrees in various fields: any STEM field in column 7, STEM or business-related fields in column 8, health or education (two of the most common fields for females) in column 9, business-related fields in column 10, and health fields in column 11. For the overall sample (row 1) we estimate a small positive effect on STEM (+1.8 ppt, t=1.0), a somewhat larger effect on business-related fields (+2.6 ppt, t=1.5); a marginally significant positive effect on STEM or business-related fields (+4.5 ppt, t=1.9), and negative effects on health-related fields (−1.9 ppt, t=1.5) and health or education (−1.9 ppt, t=1.0).

All of the effects of GEM on field of degree are concentrated among girls. The estimates for girls show a relatively large boost in STEM degree completion (+5.0 ppt, t=2.4) and in business or economics (+4.0 ppt, t=1.7), and a combined boost of 9.1 ppt (t=2.5) in STEM or business-related fields. Some of this gain appears to be at the expense of degree completion in health fields (−3.1 ppt, t=1.7). In contrast, we estimate only small effects, all far below conventional significance levels, on the fields of degree of boys. The effects for FRL and non-FRL participants are broadly similar to the effects for the overall sample.

Figures 8C and 8D show the robustness of the estimated effects on degree completion in STEM or business related fields to choice of bandwidth. For girls, the estimated effects are very similar across bandwidths, pointing to an effect of around 9 ppt, while for boys the estimated effects are slightly more negative at low bandwidths but settle down very close to 0 for bandwidths above 25 points. The estimated effects on FRL and non-FRL students are also quite robust, and consistent with an impact of 4-5 ppt, roughly the average of the effects for girls and boys.

Finally, Table 3 columns 12-14 report estimates for three outcomes that allow more time for students to achieve post-secondary benchmarks. In column 12 we redefine on-time college entry as enrollment within 2 years of normal high school completion (i.e., allowing a "gap year"). The estimated effects of GEM on this outcome are very similar to those reported in column 2, which looks at enrollment within 1 year of normal high school completion. Columns 13 and 14 present models for obtaining a bachelor's degree within 6 years of normal high school completion, and obtaining a bachelor's degree in a STEM or business-related field within 6 years. The estimates from these models are very similar to those in columns 6 and 7, respectively, suggesting that conclusions based on the conventional 5-year window are quite robust.

The impacts of GEM on major choice are summarized in Figure 9, which shows the mean shares of the treated and untreated compliers who obtain bachelor's degrees in STEM, business/economics, health or education, and all other fields. Among the untreated compliers, girls earn only about one-half as many degrees in STEM and business/economics as boys. Exposure to the GEM program, however, closes that gap. GEM appears to have no effect on the shares of either gender who ultimately earn a bachelor's degree, so the rises in the shares of STEM and business/economics degrees for girls come at the expense of the other fields. By comparison, the field distributions of the treated and untreated complier boys are quite similar.

**V. Spillover Effects on Untreated Compliers?**

Our RD analysis relies on comparisons between students whose 5$^{th}$-grade math scores are just above or just below the 380 point GEM threshold. A concern with this design is that the GEM selection process could have some effect on students who missed the threshold (i.e., the untreated compliers). The existing literature suggests three potential channels for such effects: (1) a negative peer effect, arising from the shift of the higher-scoring students to the GEM classes (Sacerdote, 2011);[21] (2) a discouragement effect, arising from the stigma of being left out of the advanced math stream (Ellison and Swanson, 2011); (3) a "top of the class" effect (Murphy and Weinhardt, 2020) arising because the untreated compliers move from being around the 80$^{th}$ percentile of their classmates in math to being the highest scoring students in their classes. Peer and discouragement effects are a particular concern because they would lead to a positive bias in our RD estimate of the effect of GEM. Top of the class effects would work in the opposite direction.

The fact that the estimated effects of GEM on the longer run outcomes of boys are uniformly small could be interpreted as evidence against the presence of large peer effects, since

[21] We note that many recent studies find limited or no evidence of peer effects. For example, Card and Giuliano (2016) find no effects of tracking 4$^{th}$ and 5$^{th}$ grade students into gifted/higher achiever (GHA) classrooms and regular classrooms on the students who narrowly miss the threshold for the GHA class. Likewise Abdulkadiroglu et al. (2014) find no evidence of peer effects on students who attend elite exam-based high schools.

boys who narrowly miss the GEM threshold experience the same changes in peer quality as girls with the same scores. Similarly, one could argue that boys who miss the threshold experience the same potential discouragement effects as girls. To the extent that middle school girls are particularly sensitive to discouragement, however, this may be unpersuasive.[22]

To address these concerns, we use data from another large urban school district where we can follow students from 4th grade to the end of college. This district – which we refer to as "District 2" – has two main tracks in middle school math, with most of the students who score in the upper half of the statewide reading test in 5th grade participating in the upper track. It has no GEM program or third track for very high scoring students. Nevertheless, it has a policy of pushing nearly all higher-scoring students into early algebra, whereas the District has a far more selective policy. The difference is illustrated in Panel A of Figure 10, where we show the fractions of girls (upper panel) and boys (lower panel) who enroll in Algebra in middle school in the District and in District 2, conditional on 5th grade math scores. To ensure balance on pre-determined characteristics (e.g., demographics and 4th-grade test scores), we reweight the District 2 sample using propensity scores constructed from a logit model for the probability of being in the District.[23] For students with scores close to the GEM cutoff, around **80%** of students in District 2 enroll in Algebra in middle school – versus around 10% of students in the District with scores just below the cutoff, and 25-40% of students in the District with scores just above the cutoff.

As shown in Panel B, however, District 2 does not have a similar policy for early geometry. Among students with 5th grade math scores in the 380-400 range (i.e., 0-20 points above the GEM cutoff) less than 5% of girls or boys in District 2 enroll in geometry by 8th grade. In contrast, in the District this fraction is between 35 and 45% for girls, and between 20 and 40% for boys.

Given the differences in the middle school math sequence in District 2 – where the vast majority of students with 5th grade math scores on either side of the GEM threshold are taking

[22] Ellison and Swanson (2021) find evidence of larger discouragement effects for girls than boys from competing in math competitions and losing, and discuss earlier studies which often report this pattern. They interpret their findings in the context of a broader literature suggesting that many females prefer to avoid competition.

[23] The propensity score model is estimated using all students in the two districts who meet our main sample criteria, separately for boys and girls. The model includes dummies for 25-point bins of 5th-grade math scores; indicators for race, FRL status, English Language Learner status, a quadratic in age, 4th-grade math and reading scores interacted with race, and 5th-grade reading scores.

*the same* advanced track courses – and in the District, where these students are separated into GEM and advanced tracks based on having a math score above or below 380 points, it is interesting to ask how the longer run outcomes of students with different 5th grade scores compare in the two districts. In particular, it is interesting to ask whether there is any evidence that students who have 5th grade math scores just below the GEM cutoff do worse in the District than in the District 2 – as might happen if there are significant negative peer effects or discouragement effects caused by moving the higher-scoring students to the GEM track. It is also useful to confirm that girls in the District who score just above the 380 point threshold show gains relative to girls in District 2 with similar scores.

Evidence on these two questions is presented in Figure 11.[24] The upper panels of this figure show the predicted fractions of students who would earn a bachelor's degree in STEM or business-related fields within 5 years of their on-time high school graduation year, using student demographics, test scores in 4th and 5th grades, and school characteristics, and the coefficients from a prediction model fit to students in the District. The lower panel shows the actual fractions of students in the two districts completing degrees in STEM or business/economics.

To help interpret the graphs, we fit local linear models for the outcomes of students in District 2, assuming no discontinuity (or slope change) in the region $\pm 25$ points from the GEM cutoff. These are shown as the dashed red lines. As in our main RD analysis, we fit local linear models for students in the District, allowing potential discontinuities and/or slope changes. These are shown as the solid blue lines in Figure 11.

Looking first at the predicted shares of students with STEM or business-related degrees, we see that the predictions are quite similar for students in the District and in District 2, conditional on 5th-grade math scores. In particular, around the 380 point threshold students on either side are quite similar within each of the districts and between the two districts.

Turning to the actual shares of students with STEM or business-related degrees, we see that for girls, the means for girls with scores just below 380 in the District are quite similar to the means for girls in District 2. Moreover, the fitted lines to the left of the cutoff have nearly the

[24] In Appendix Figure A5, we present a comparison of measures of mathematics achievement in high-school, including enrollment in calculus and PSAT scores. The patterns are similar to those in Figure 11.

same slope and cross the 380 point threshold at virtually the same level.[25] To the right of the GEM threshold, however, there is an obvious boost for girls in the District relative to those in District 2. For boys the story is similar, though the fitted line to the left of the GEM threshold for boys in the District is slightly flatter than the benchmark line for boys in District 2. Given that the fitted line to the right of the threshold for boys in the District meets the line to the left at virtually the same height (i.e., that there is no discontinuity in majoring for boys in the District), and that the fitted line to the right for boys in the District lies virtually on top of the benchmark line for District 2, we believe there is no evidence of negative spillover effects from the GEM program to students of either gender who narrowly miss qualifying for the program.

## VI. Summary and Conclusions

There is much disagreement about how to structure the math curriculum in American middle and high schools. On one hand, the science and engineering programs at many selective schools ask for evidence of mastery of advanced math in high school, especially calculus. This leads the parents of higher-achieving students to demand access to algebra and geometry in middle school so their children have a pathway to calculus by the end of high school. On the other hand, previous efforts to push "algebra for all" in middle school are perceived as having failed, especially for weaker students. A possible resolution to this conundrum is the availability of targeted acceleration programs like the GEM program we study in this paper.

GEM is a three-year program that compresses the standard 6$^{th}$ and 7$^{th}$ grade curriculum into a single year, leading to Algebra I in 7$^{th}$ grade and Geometry in 8$^{th}$ grade. Admission is limited to students who score above the 80$^{th}$ percentile on the statewide 5$^{th}$-grade math test, and it has "off-ramps" after 6$^{th}$ and 7$^{th}$ grade that allow students to return to the conventional math track if they are struggling. In fact only about 40% of GEM entrants actually complete all three years, but 80% complete algebra by the end of 8$^{th}$ grade.

We use the admission threshold for GEM to conduct a regression discontinuity-based analysis of the effects of entering GEM on subsequent course taking and grades in middle and high school, and on post-secondary outcomes, including degree attainment and field of

---

[25] It may be surprising that students in the District with scores below the GEM threshold have the same probability of majoring in STEM or business as those in District 2, given that a much higher share of this group in District 2 enrolls in algebra in 8$^{th}$ grade. We interpret this as evidence that pushing a large share of students into early algebra has very limited effects. Notably, students below the threshold in District 2 are no more likely than those in the District to enroll in calculus in high school (see Appendix Figure A5).

specialization. We have two main findings. First, the program has little or no effect on the college outcomes of boys, but relatively large effects on outcomes for girls, leading to relatively large increases in the probability of majoring in math-intensive fields (STEM and business /economics). Second, the impacts are about the same for economically disadvantaged students (who were receiving free or reduced price meals in 5th grade) and more advantaged students. Thus, GEM helps narrow gender gaps without widening disparities across income groups.

The main impacts of GEM for girls are summarized in Figure 12. We show the means of each of 7 outcomes for the *untreated compliers* – the set of students who form the control group in a fuzzy RD design – as well as the causal impacts of entry to GEM on the *treated compliers* – the set of students who form the treatment group. Starting in middle school, GEM entry leads to a 22 ppt increase in the probability of enrolling in algebra by 8th grade, relative to a baseline of 60% for the untreated compliers. The rather high baseline illustrates the fact that most participants in GEM would have been in the advanced math sequence in the absence of GEM, which allows students who perform well in 7th grade to take algebra in 8th grade. Just over one-half of girls who enter GEM also enroll in geometry in 8th grade. By comparison, only 30% of boys make it all the way to the 3rd year of GEM.

In high school, participation in GEM leads to a 10 ppt increase in the fraction of girls who successfully complete calculus, and a 3.5 percentile increase in their PSAT math scores. Among students with math scores just below the GEM threshold, boys have higher PSAT math scores than girls. But for those with scores just above the threshold, GEM pushes girls' scores up to the level of boys.

At the college level, GEM leads to an insignificant (+4 ppt) gain in 5-year bachelor degree completion rates for girls but significant changes in their major choice. Most importantly, exposure to GEM leads to a 70% rise in the fractions of girls who graduate with bachelor's degrees in STEM and business or economics. Entry to GEM leads nearly 25% of girls to graduate in the relatively highly paid STEM and business-related fields – closing the gap with boys and potentially having large effects on the future careers of these young women.

A widespread concern with tracking programs like GEM is that they could harm lower-performing students who miss out on the advanced track. These concerns have led some school districts to eliminate math tracking in middle school, despite the resulting problems in meeting the "hidden curriculum" requirements for selective engineering and science programs. To address this concern, we use data from a second large school district in the same state, which has regular and advanced tracks in middle school math but not a third upper track like GEM for the highest performers in math. We find no evidence that students of either gender with scores just below the GEM threshold perform worse in the District than in this other school district.

Overall, we conclude that the introduction of a targeted accelerated math curriculum like GEM has the potential to help close the gender gap in the selection of STEM majors and higher-paid college majors in general, without hindering the academic performance of lower-performing students.

## References

Abadie, Alberto. 2002. “Bootstrap tests for distributional treatment effects in instrumental variable models.” *Journal of the American Statistical Association* 97 (457): 284–92.

Abdulkadiroglu, Atila, Joshua Angrist, and Parag Pathak. 2014. “The Elite Illusion: Achievement Effects at Boston and New York Exam Schools.” *Econometrica* 82 (1): 137–96.

Altonji, Joseph G., Peter Arcidiacano and A. Maurel. 2016. The determinants of field choice in college and graduate school: Determinants and wage effects.” In Eric A. Hanushek, Stephen Machin and Ludger Woessmann, editors. *Handbook of the Economics of Education* Volume 5. Elsevier: New York: 305-396.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. “Robust nonparametric confidence intervals for regression-discontinuity designs.” *Econometrica* 82 (6): 2295–2326.

Card, David and Laura Giuliano. 2016. “Can Tracking Raise the Test Scores of High-Ability Minority Students?” *American Economic Review* 106(10): 2783–2816.

Card, David, Eric Chyn and Laura Giuliano. 2024. “Can gifted education help higher-ability boys from disadvantaged backgrounds?” UC Santa Cruz Department of Economics Working Paper, September 2024.

Card, David and A. Abigail Payne. 2021. “High school choices and the gender gap in STEM.” *Economic Inquiry* 59(1): 9-28.

Clotfelter, Charles T., Helen F. Ladd and Jacob L. Vigdor. 2015. “The aftermath of accelerating algebra: Evidence from district policy initiatives.” *Journal of Human Resources* 50 (1): 159-188.

Cohodes, Sarah R. 2020. “The Long-Run Impacts of Specialized Programming for High-Achieving Students.” *American Economic Journal: Economic Policy* 12(1): 127-166.

Denning Jeffrey T., Richard Murphy and Felix Weinhardt. 2023. “Class Rank and Long-Run Outcomes.” *Review of Economics and Statistics* 105(6): 1426-1441.

DesJardins, Stephen L., Dennis A. Ahlburg, and Brian P. McCall. 2002. “A Temporal Investigation of Factors Related to Timely Degree Completion.” *Journal of Higher Education* 73 (5): 555-581.

Domina, Thurston, Andree McEachin, Andrew Penner and Emily Penner. 2015. “Aiming high and falling short: California’s eighth-grade algebra-for-all effort.” *Educational Evaluation and Policy Analysis* 37 (3): 275–295.

Dougherty, Shaun M., Joshua S. Goodman, Darryl V. Hill, Erica G. Litke and Lindsay C. Page. 2017. “Objective course placement and college readiness: Evidence from targeted middle school math acceleration.” Economics of Education Review 58: 141–161.

Ellison, Glenn, and Ashley Swanson. 2021. "Dynamics of the Gender Gap in High Math Achievement." *Journal of Human Resources* 58(5): 1679-1711.

Hahn, Jinyong, Petra Todd Wilbert van der Klaauw. 2001. "Identification and estimation of treatment effects with a regression-discontinuity design." *Econometrica* 69 (1): 201-209.

Huffaker, Christopher. 2023. "Cambridge schools are divided over middle school algebra." *Boston Globe* July 14, 2023.

Huffaker, Elizabeth, Sarah Novicoff and Thomas S. Dee. 2024. "Ahead of the game? Course-taking patterns under a math pathways reform." Stanford University Graduate School of Education Working Paper. June 2024.

Lee, David S. 2008. "Randomized experiments from non-random selection in U.S. House elections." *Journal of Econometrics* 142: 675–697.

Lee, David S., Justin McCrary, Marcelo J. Moreira and Jack Porter. 2022. "Valid t-ratio inference for IV." *American Economic Review*, 112 (10): 3260–90.

McCrary, Justin 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics,* 142: 698–714.

McNally, Sandra 2020. "Gender differences in tertiary education: What explains STEM participation?" IZA Policy Paper No. 165. October 2020.

National Mathematics Advisory Panel. 2008. Foundations for success: The final report of the National Mathematics Advisory Panel. U.S. Department of Education: Washington, DC 20008.

Peters, Scott J. and James A. Carter. 2023. "The where and why of accelerated middle-school math." *Gifted Child Quarterly* 67(3) 235–243.

Sacerdote, Bruce. 2011. "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" In Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, editors, *Handbook of the Economics of Education*. Volume 3. New York: Elsevier.

Speer, Jamin D. 2017. "The gender gap in college major: Revisiting the role of pre-college factors." *Labour Economics* 44: 69-88.

United State Department of Education National Center for Education Statistics. 2022. *Digest of Education Statistics.* Available at https://nces.ed.gov/programs/digest/. Accessed September 9 2024.

Yoder, Steven. 2024. "Eliminating advanced math tracks often prompts outrage. Some districts buck the trend." *The Hechinger Report* April 17, 2024.

**Figure 1. Math Pathways in Middle School**

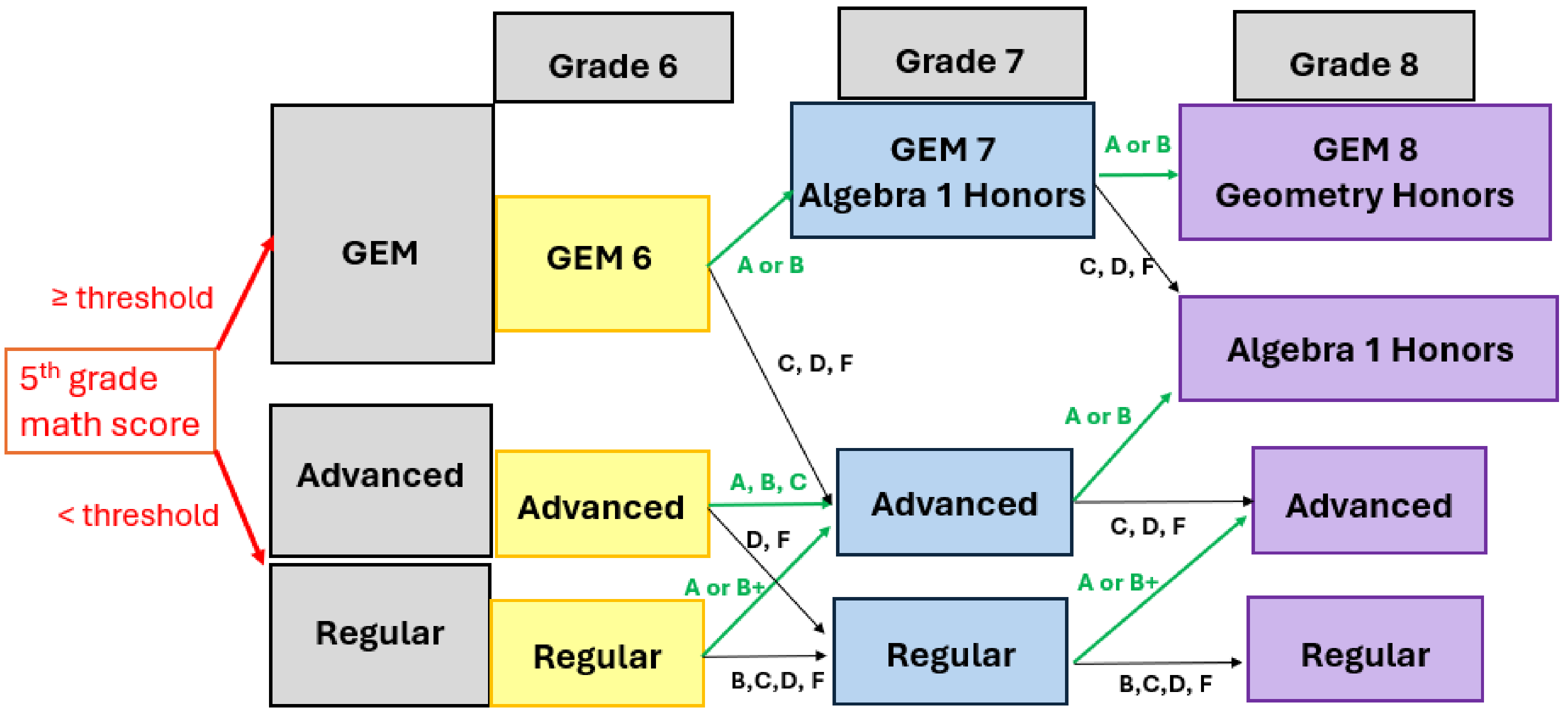
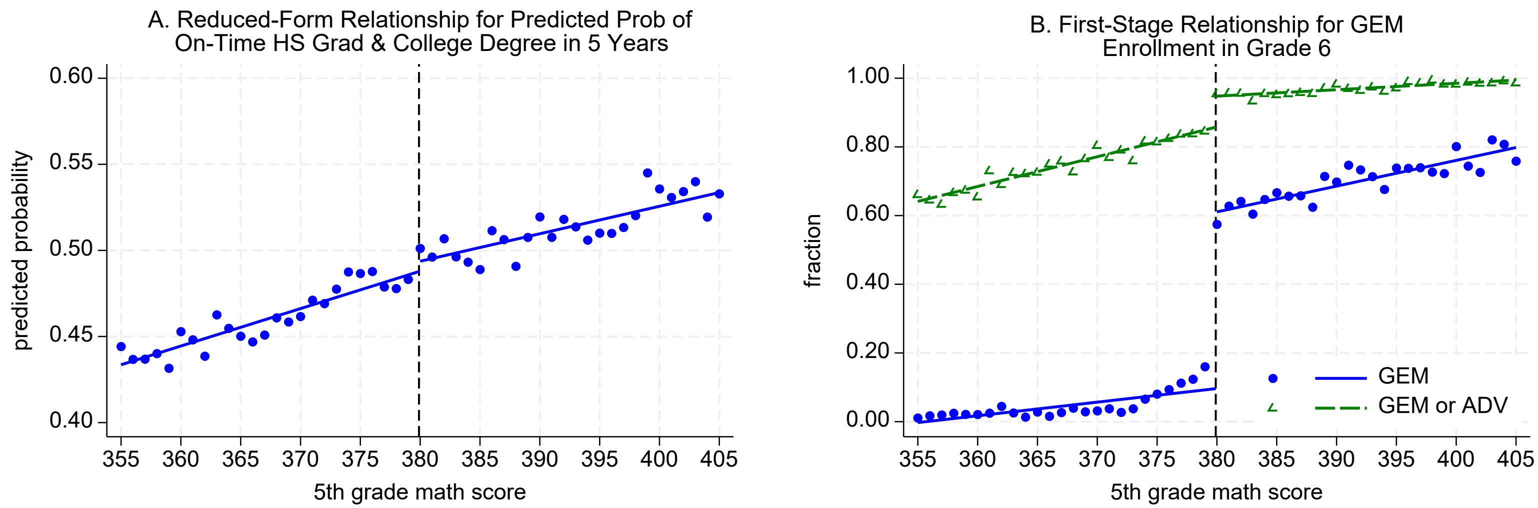
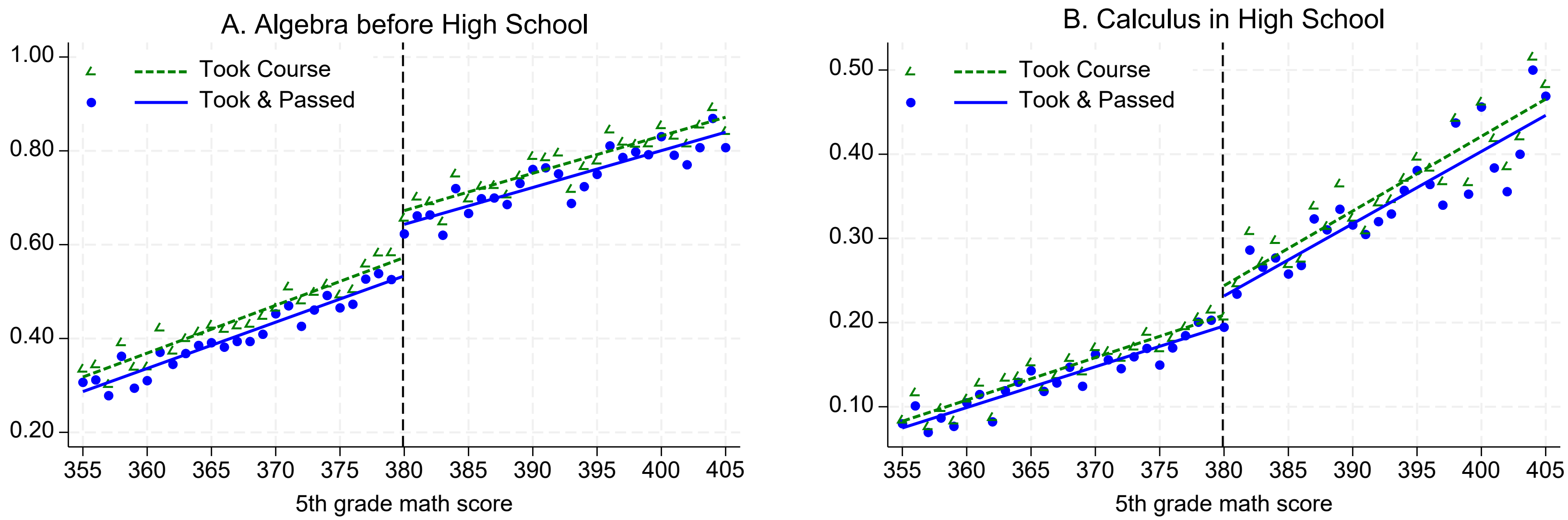


**Figure 2. Predicted Outcome and GEM Enrollment by Fifth-Grade Math Score, Pooled Sample**



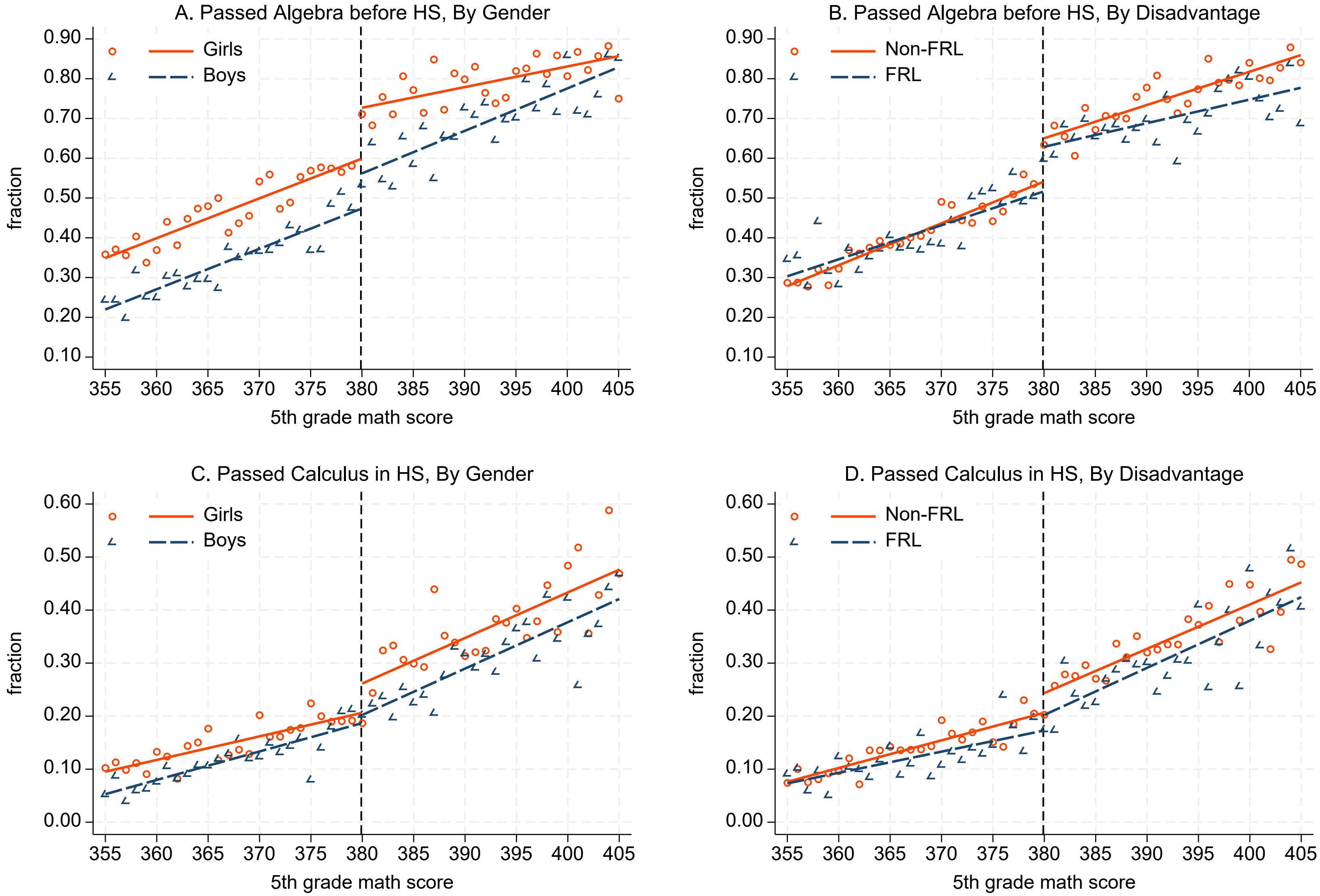

**Notes:** This figure plots sample means of the dependent variable at each value of fifth-grade math score (the running variable) along with fitted values from local linear RD models with a symmetric bandwidth of 25. In Panel A, the dependent variable is an index of pre-determined student characteristics constructed as the predicted probability of graduating from high school on time and earning a college degree within 5 years of high-school graduation. The prediction model is fit using students with fifth-grade math scores below the GEM eligibility threshold of 380 and includes second-order polynomials in math and reading test scores from fourth grade and in student age in fifth grade; indicators for student gender, race/ethnicity, FRL status, ELL status, and cohort; and average test scores and fraction FRL of the school where the student is enrolled in fifth grade. In panel B, the dependent variables are an indicator for participating in the GEM program in sixth grade (the treatment variable, in blue circles) and an indicator for participating in *either* GEM *or* advanced-track math (in green triangles).

**Figure 3. Reduced-Form Relationship for High School Outcomes, Pooled Sample**



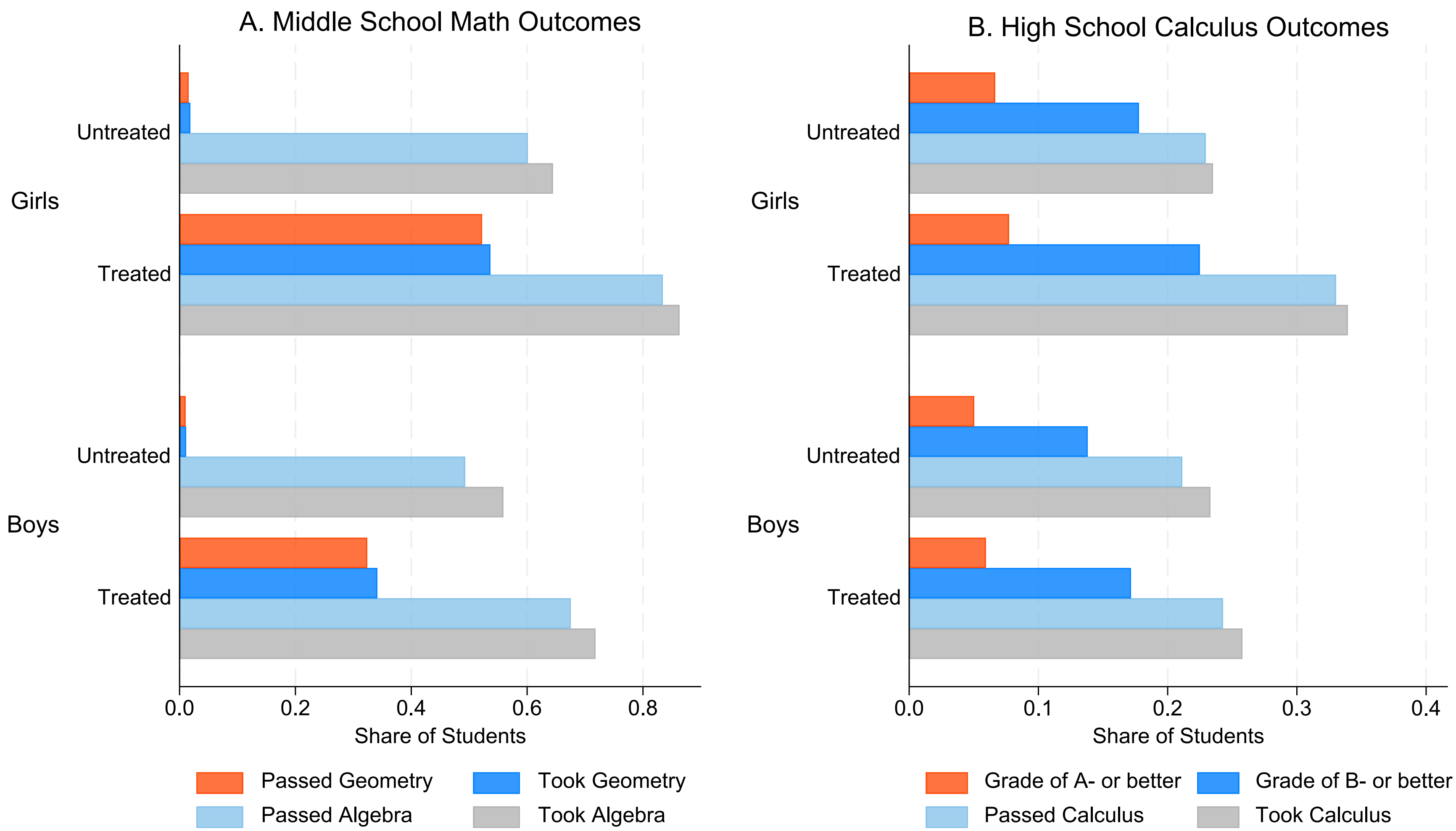

**Notes:** This figure plots sample means of the dependent variable at each value of fifth-grade math score (the running variable) along with fitted values from local linear RD models with a symmetric bandwidth of 25. In Panel A, the dependent variables are indicators for having enrolled algebra in middle school (grade 7 or 8) and having completed it with a passing grade (C- or better). In Panel B, the dependent variables are having enrolled in calculus in high school (grades 9-12) and having completed it with a passing grade.

**Figure 4. Reduced-Form Relationship for Successful Completion of Early Algebra / HS Calculus, by Gender and Disadvantage**



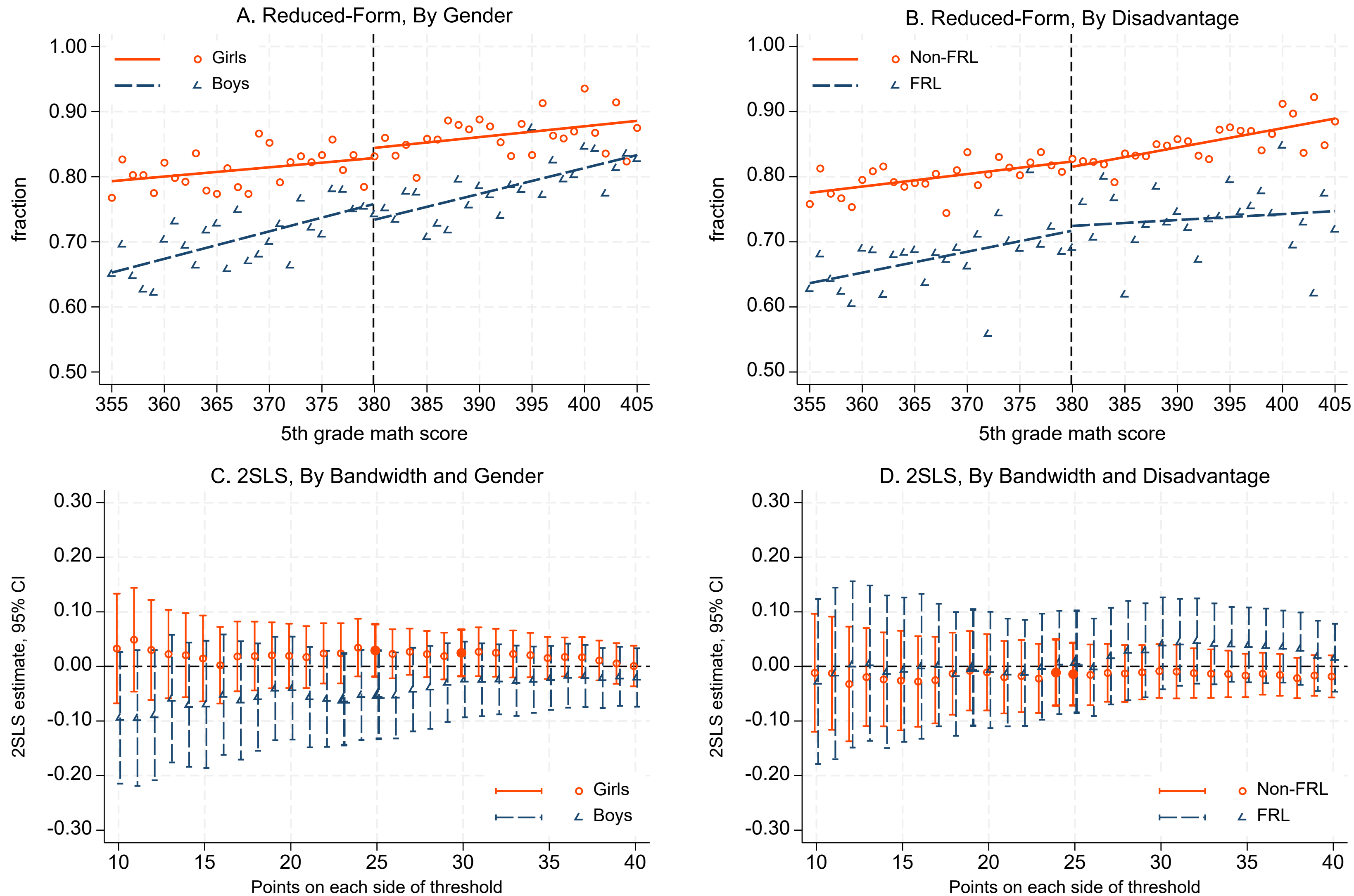

**Notes:** This figure plots sample means of the dependent variable at each value of fifth-grade math score (the running variable), along with fitted values from local linear RD models with a symmetric bandwidth of 25, separately by gender (Panels A and C) and by FRL status (Panels B and D). In Panels A and B, the dependent variable is an indicator for taking algebra in middle school (grade 7 or 8) and earning a passing grade (C- or better). In Panels C and D, the dependent variable is an indicator for taking calculus in high school and earning a passing grade (C- or better).

## Figure 5. Outcome Distributions for Treated and Untreated Compliers



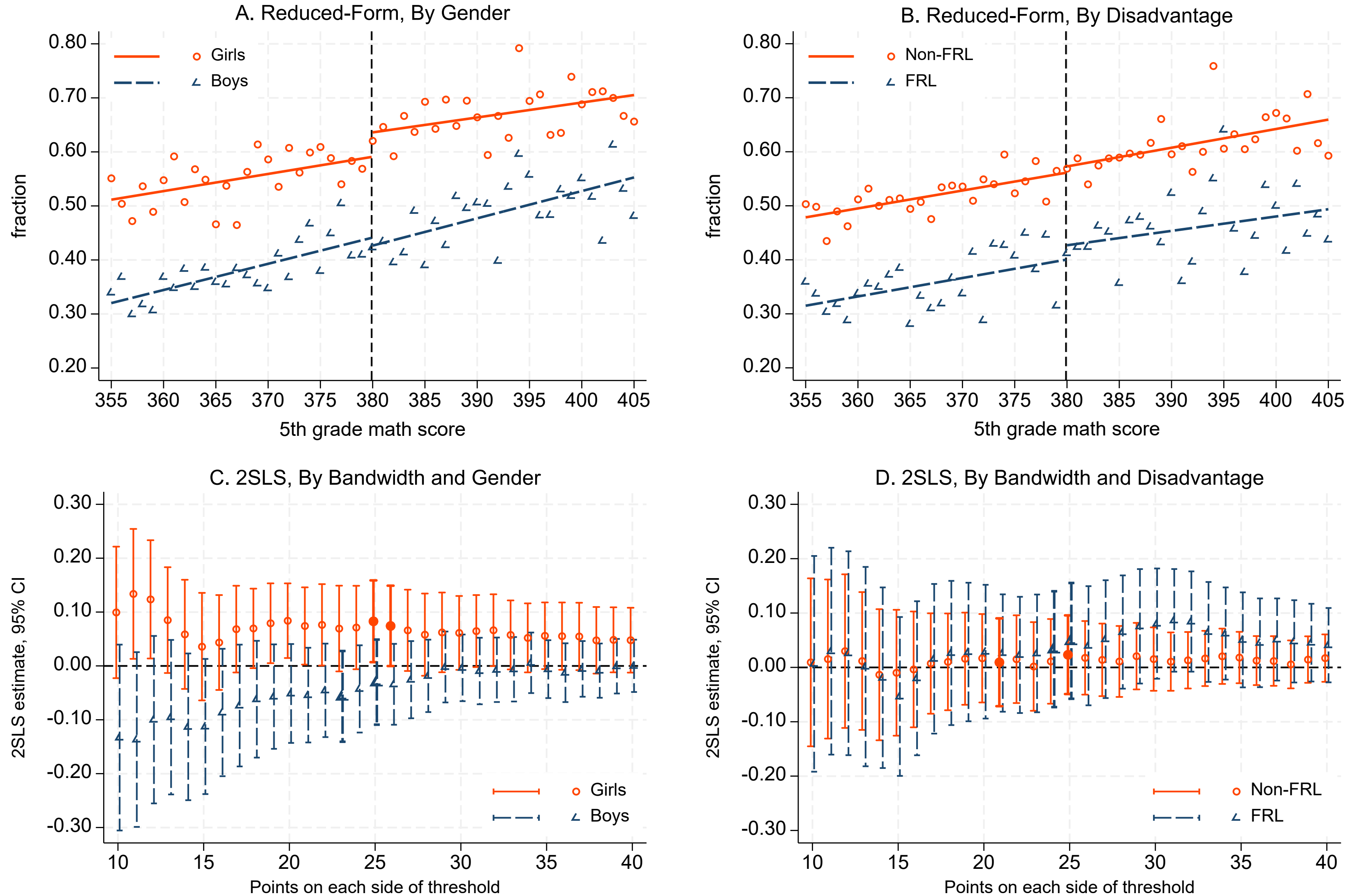

**Notes:** In panel A, taking algebra is defined as enrolling in algebra in either 7th grade or 8th grade. In both panels, passing Algebra (Geometry, Calculus) is defined as earning a C or better. For students who took Algebra twice, this variable is based on the grade the second time they took it.

**Figure 6. Graduate High School on Time and Enter College within 1 Year, by Gender and Disadvantage**



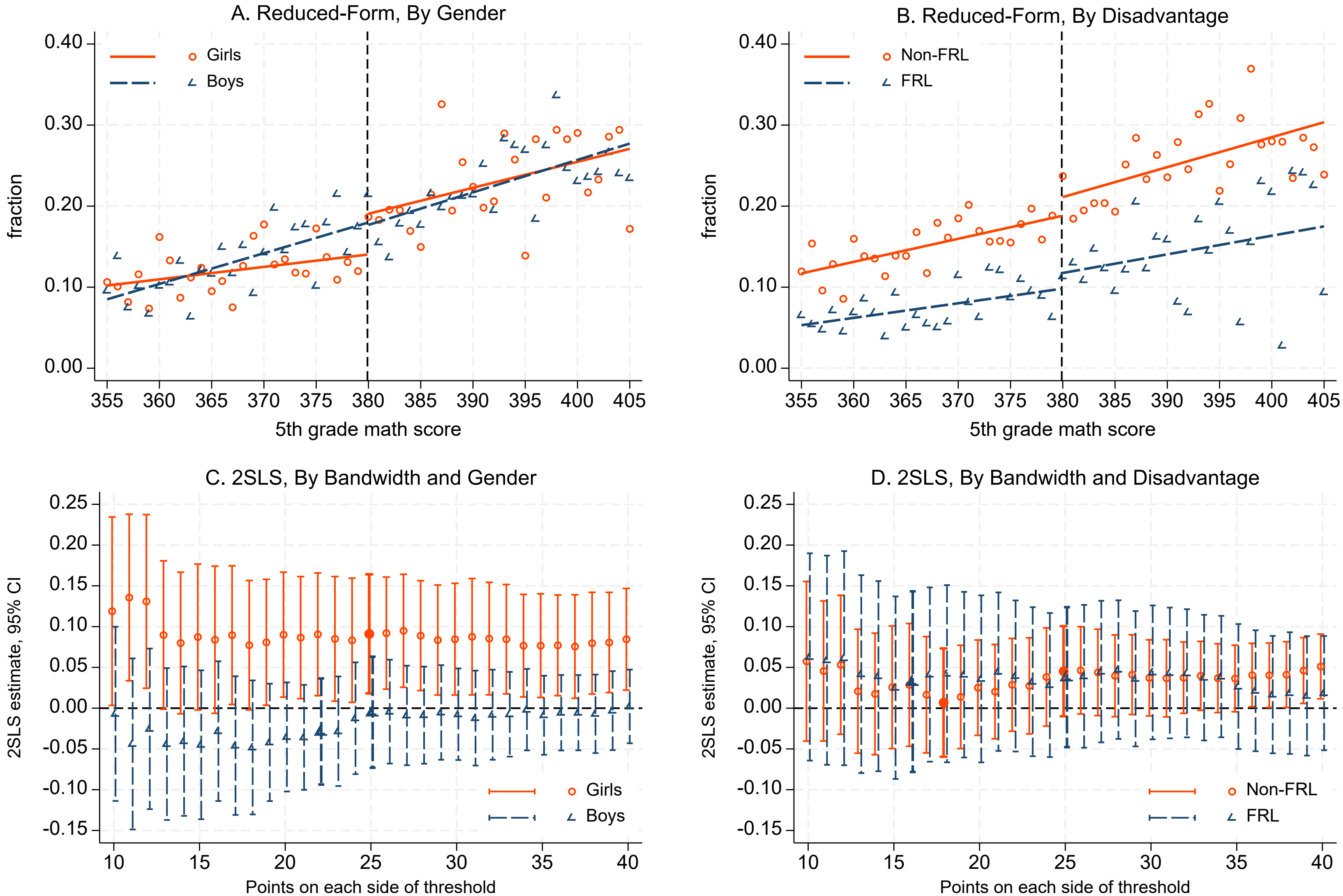

**Notes:** Panels A and B plot sample means of an indicator for graduating high school on time and enrolling in college within 1 year at each value of fifth-grade math score (the running variable), along with fitted values from local linear RD models with a symmetric bandwidth of 25, separately by gender (Panels A) and by FRL status (Panels B). Panels C and D plot 2SLS estimates for the treatment effect of GEM entry on the dependent variable for symmetric bandwidths between 10 and 40, with the CCT symmetric bandwidth and preferred bandwidth (+/-25) shown in bold.

**Figure 7. Graduate HS on Time and Earn Any College Degree within 5 Years, by Gender and Disadvantage**




**Notes:** Panels A and B plot sample means of an indicator for graduating high school on time and earning any college degree within 5 years at each value of fifth-grade math score (the running variable), along with fitted values from local linear RD models with a symmetric bandwidth of 25, separately by gender (Panels A) and by FRL status (Panels B). Panels C and D plot 2SLS estimates for the treatment effect of GEM entry on the dependent variable for symmetric bandwidths between 10 and 40, with the CCT symmetric bandwidth and preferred bandwidth (+/-25) shown in bold.

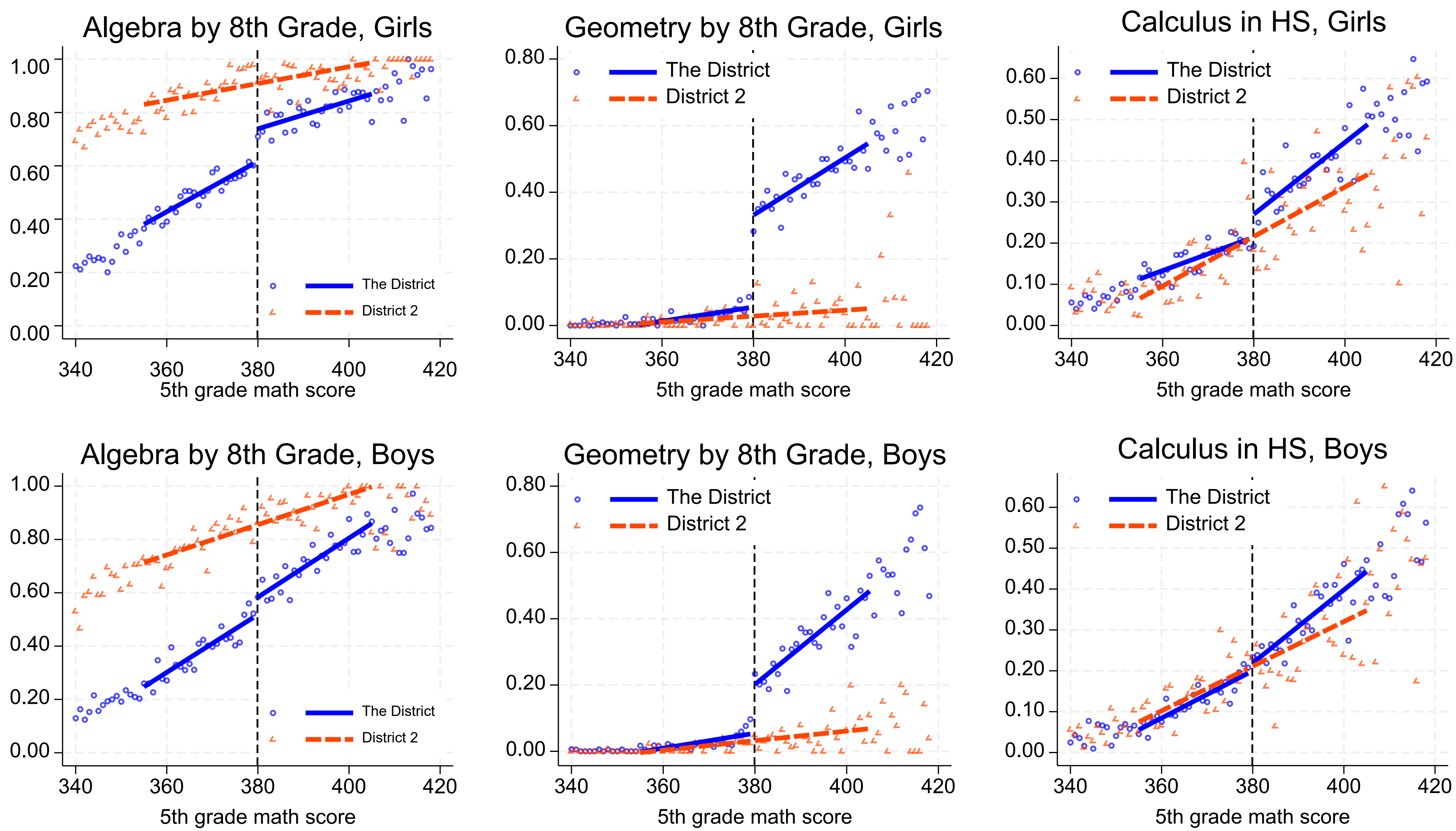
**Figure 8. Graduate HS on Time and Earn a Bachelors Degree in STEM/Bus/Econ, by Gender and Disadvantage**




**Notes:** Panels A and B plot sample means of an indicator for graduating high school on time and earning any college degree within 5 years at each value of fifth-grade math score (the running variable), along with fitted values from local linear RD models with a symmetric bandwidth of 25, separately by gender (Panels A) and by FRL status (Panels B). Panels C and D plot 2SLS estimates for the treatment effect of GEM entry on the dependent variable for symmetric bandwidths between 10 and 40, with the CCT symmetric bandwidth and preferred bandwidth (+/-25) shown in bo

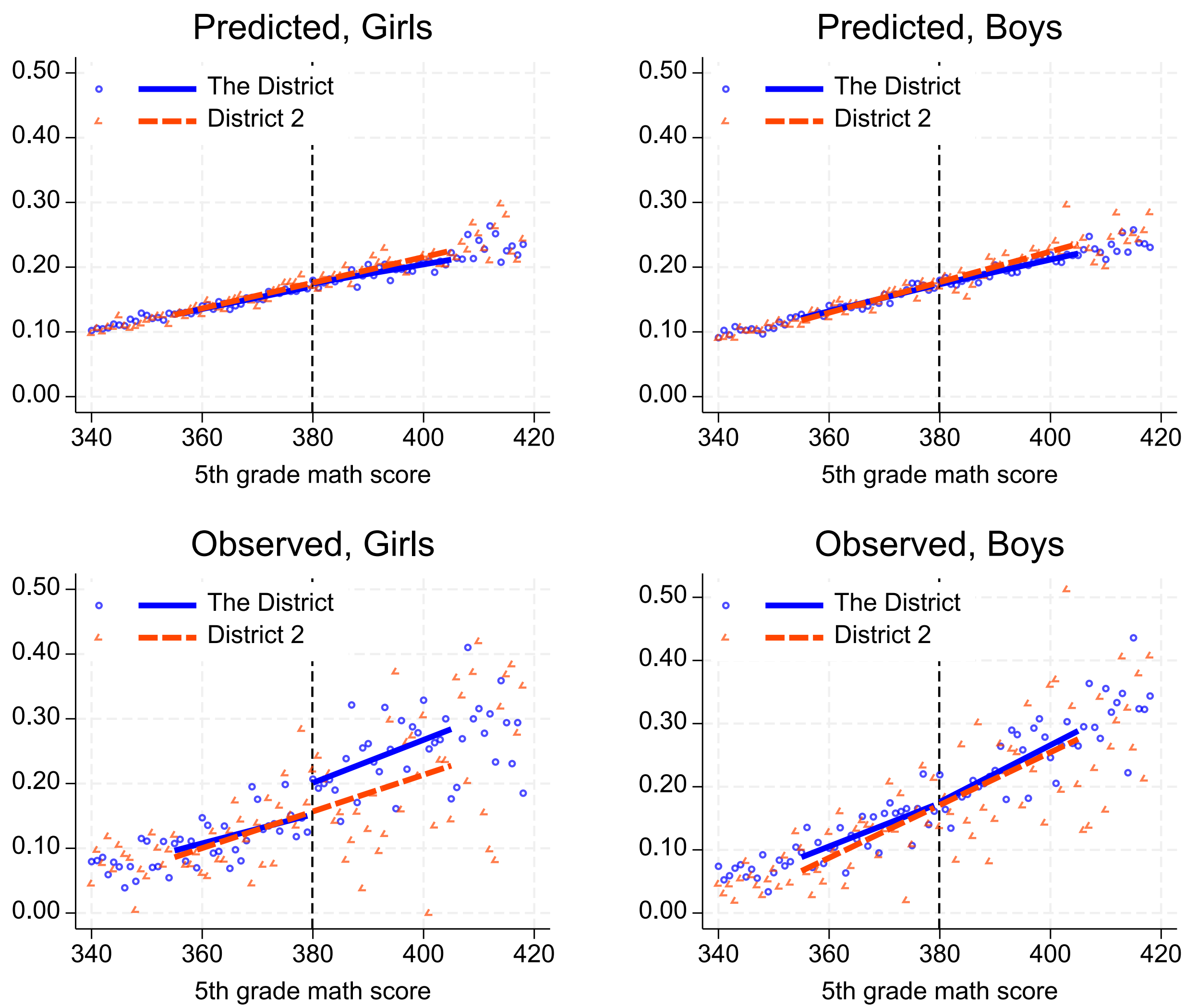
**Figure 9. Degree Earning Distributions for Treated and Untreated Compliers**

**Figure 10. Math Course Enrollments, Between-District Comparison**




**Notes**: This figure plots sample means of each variable at each value of fifth-grade math score, separately for the District and a comparison district (District 2) that did not offer an accelerated track like GEM for the highest performers but enrolled a larger share of lower-performers in algebra in 8th grade. Also shown are fitted values from local linear RD models with a symmetric bandwidth of 25 for the District (solid blue lines), and fitted values from linear models (with no discontinuity) using the same bandwidth for District 2 (orange dashed lines). Observations for District 2 are reweighted using propensity scores constructed from a logit model for the probability of being in BCPS , estimated using all students in BCPS and HCPS who meet our main sample criteria, separately for boys and girls. The model includes dummies for 25-point bins of g5math; dummies for race, ELL and FRL status, a quadratic in age, 4th-grade math and reading scores interacted with race, and 5th-grade reading scores.

# Figure 11. Bachelors Degree in STEM/Business/Economics in 5 years, Between-District Comparison




**Notes**: This figure plots sample means of each variable at each value of fifth-grade math score, separately for the District and a comparison district (District 2) that did not offer an accelerated track like GEM for the highest performers but enrolled a larger share of lower-performers in algebra in 8th grade. Also shown are fitted values from local linear RD models with a symmetric bandwidth of 25 for the District (solid blue lines), and fitted values from linear models (with no discontinuity) using the same bandwidth for District 2 (orange dashed lines). Observations for District 2 are reweighted using propensity scores constructed from a logit model for the probability of being in BCPS , estimated using all students in BCPS and HCPS who meet our main sample criteria, separately for boys and girls. The model includes dummies for 25-point bins of g5math; dummies for race, ELL and FRL status, a quadratic in age, 4th-grade math and reading scores interacted with race, and 5th-grade reading scores.

**Figure 12. Summary of Effects of GEM Entry on Girls**

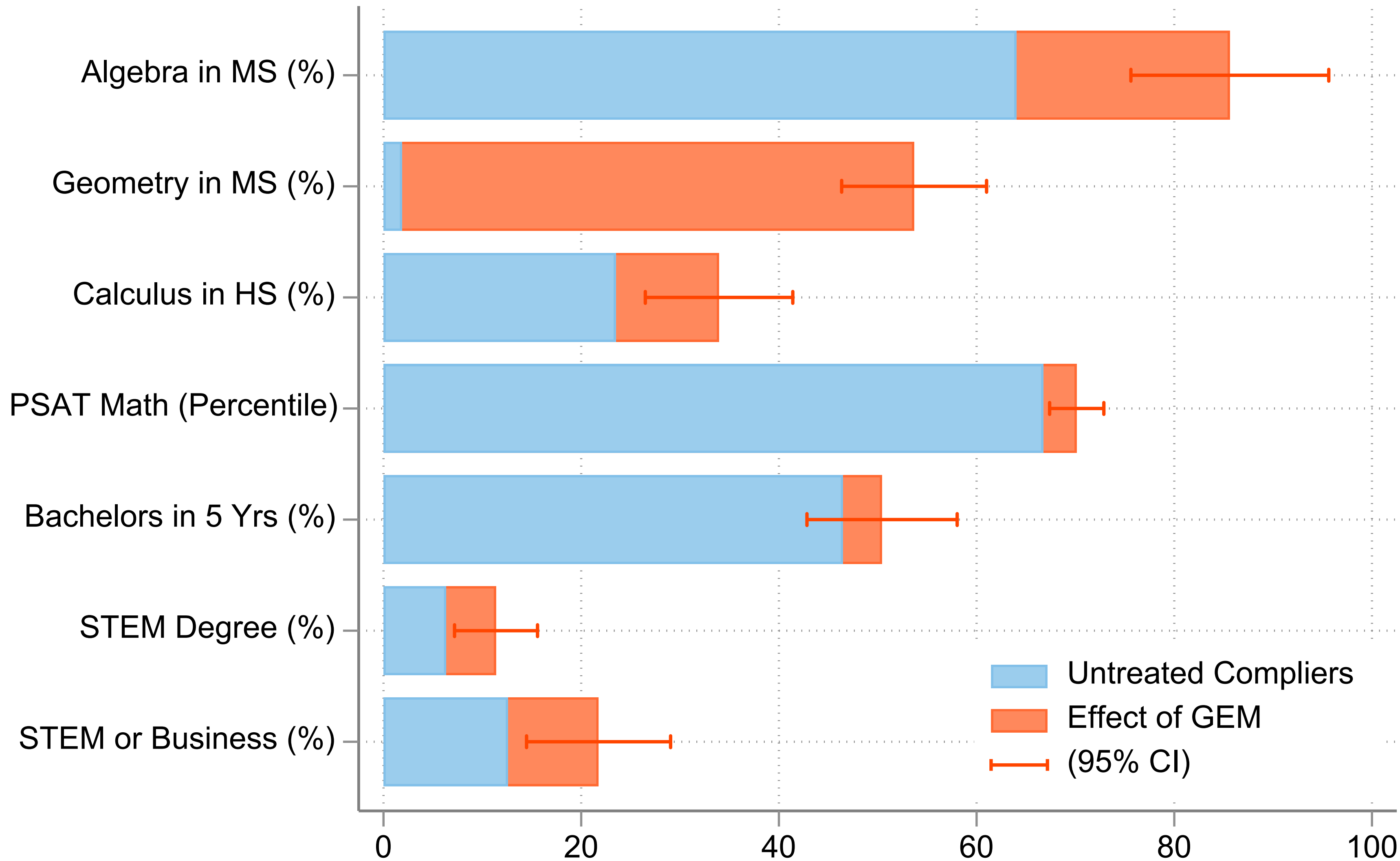


**Table 1. Estimated Means for Main Analysis Sample and for Compliers**

| A. Baseline Characteristics | Main Analysis Sample | Compliers | | | | |
|---|---|---|---|---|---|---|
| | | All | Girls | Boys | Non-FRL | FRL |
| **Demographic characteristics** | | | | | | |
| female | 0.49 | 0.49 | 1.00 | 0.00 | 0.47 | 0.54 |
| FRL | 0.31 | 0.31 | 0.33 | 0.28 | 0.00 | 1.00 |
| White | 0.44 | 0.41 | 0.39 | 0.43 | 0.52 | 0.17 |
| Black | 0.23 | 0.23 | 0.28 | 0.18 | 0.14 | 0.45 |
| Hispanic | 0.25 | 0.25 | 0.21 | 0.30 | 0.24 | 0.27 |
| **Baseline achievement** | | | | | | |
| grade 4 math score (standardized) | 0.65 | 0.78 | 0.75 | 0.80 | 0.81 | 0.71 |
| grade 4 reading score (standardized) | 0.53 | 0.66 | 0.71 | 0.61 | 0.70 | 0.58 |
| **School peer characteristics** | | | | | | |
| avg grade 4 math & reading scores | 0.13 | 0.12 | 0.11 | 0.14 | 0.18 | 0.00 |
| fraction FRL | 0.41 | 0.43 | 0.44 | 0.42 | 0.36 | 0.57 |
| fraction GEM eligible | 0.22 | 0.22 | 0.21 | 0.22 | 0.23 | 0.18 |
| **Predicted prob graduate high school on time & earn any college degree within 5 years** | 0.48 | 0.49 | 0.58 | 0.41 | 0.54 | 0.39 |
| **B. Mean intermediate outcomes or potential outcomes for compliers not in GEM** | | | | | | |
| **Math track in 6th grade** | | | | | | |
| GEM (treatment) | 0.28 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| advanced | 0.55 | 0.83 | 0.86 | 0.79 | 0.85 | 0.78 |
| regular | 0.18 | 0.17 | 0.14 | 0.21 | 0.15 | 0.22 |
| **Middle school course enrollments & grades** | | | | | | |
| algebra by 8th grade | 0.55 | 0.60 | 0.64 | 0.56 | 0.62 | 0.54 |
| completed algebra with B- or better by 8th grade | 0.39 | 0.44 | 0.51 | 0.37 | 0.46 | 0.39 |
| completed algebra with C- or better by 8th grade | 0.52 | 0.55 | 0.60 | 0.49 | 0.57 | 0.48 |
| geometry in 8th grade | 0.16 | 0.02 | 0.02 | 0.01 | 0.02 | 0.01 |
| 6th grade math GPA | 3.09 | 3.48 | 3.61 | 3.36 | 3.56 | 3.29 |
| 7th grade math GPA | 3.00 | 3.20 | 3.35 | 3.06 | 3.25 | 3.08 |
| 8th grade GPA | 2.90 | 2.99 | 3.10 | 2.88 | 3.04 | 2.87 |
| **High school course enrollments & grades** | | | | | | |
| took any advanced math (statistics, calculus, etc.) | 0.63 | 0.67 | 0.71 | 0.64 | 0.70 | 0.59 |
| took any calculus | 0.21 | 0.23 | 0.24 | 0.23 | 0.25 | 0.21 |
| completed any calculus with B- or better | 0.14 | 0.16 | 0.18 | 0.14 | 0.17 | 0.14 |
| completed any calculus with C- or better | 0.20 | 0.22 | 0.23 | 0.21 | 0.24 | 0.19 |
| took AP/IB calculus | 0.13 | 0.14 | 0.14 | 0.13 | 0.15 | 0.11 |
| took AP calculus exam + scored 3 or better | 0.07 | 0.06 | 0.06 | 0.07 | 0.07 | 0.05 |
| high school math GPA (all grades, all courses) | 3.21 | 3.26 | 3.38 | 3.15 | 3.30 | 3.17 |
| **College boards** | | | | | | |
| PSAT math percentile | 64.28 | 68.0 | 66.8 | 69.4 | 69.4 | 65.0 |
| PSAT reading percentile | 58.26 | 61.3 | 63.4 | 59.5 | 63.0 | 57.8 |
| Took SAT or ACT | 0.87 | 0.89 | 0.91 | 0.86 | 0.91 | 0.85 |

Note: table continues. See notes at end of second page.

**Table 1, Cont'd. Estimated Means for Main Analysis Sample and for Compliers**

| | Main Analysis Sample | Compliers | | | | |
|---|---|---|---|---|---|---|
| | | All | Girls | Boys | Non-FRL | FRL |
| <u>**C. Mean post-secondary outcomes or potential outcomes for compliers not in GEM**</u> | | | | | | |
| **College enrollment** | | | | | | |
| graduate high school on time | 0.95 | 0.97 | 0.96 | 0.98 | 0.97 | 0.95 |
| + enter any college within 1 year | 0.78 | 0.80 | 0.82 | 0.80 | 0.83 | 0.74 |
| + enter selective college within 1 year | 0.59 | 0.62 | 0.65 | 0.60 | 0.66 | 0.53 |
| + enter any college w/in 1 yr & persist >1 yr | 0.62 | 0.64 | 0.67 | 0.61 | 0.64 | 0.64 |
| **College degree earned** | | | | | | |
| + earn any college degree within 5 years | 0.50 | 0.51 | 0.56 | 0.46 | 0.55 | 0.41 |
| + earn bachelors (BA or BS) within 5 years | 0.35 | 0.38 | 0.46 | 0.30 | 0.42 | 0.29 |
| + earn BA/BS in STEM or Econ/Bus within 5 years | 0.16 | 0.16 | 0.13 | 0.21 | 0.19 | 0.10 |
| + earn BA/BS in STEM within 5 years | 0.08 | 0.08 | 0.06 | 0.10 | 0.10 | 0.05 |
| + earn BA/BS in Econ/Bus in 5 yrs | 0.08 | 0.08 | 0.06 | 0.10 | 0.10 | 0.05 |
| + earn BA/BS in Health/Educ in 5 yrs | 0.05 | 0.07 | 0.10 | 0.04 | 0.09 | 0.03 |
| + earn BA/BS in Health in 5 yrs | 0.04 | 0.06 | 0.09 | 0.04 | 0.08 | 0.03 |

Notes: The first column shows means of variables for all students in the main analysis sample (with math scores between 355 and 405). The remaining columns show estimated mean characteristics of all the compliers (Panel A), or estimated mean potential outcomes for compliers if they do not enter GEM (Panels B & C). See text for description of method of estimating mean characteristics of compliers and mean potential outcomes.

**Table 2. Estimated Treatment Effects on Math Course Enrollment and Performance**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Predicted College Degree in 5 Years | First Stage (GEM in Grade 6) | Course Enrollments | | | Took Algebra by 8th & Earned: | | Took Calculus in HS & Earned: | | Grade Point Avg in Math Courses | | | |
| | | | Algebra by 8th Grade | Geometry in 8th Grade | Calculus in HS | B- or better | C- or better | B- or better | C- or better | 6th Grade | 7th Grade | 8th Grade | All HS Math |
| **1. Pooled** | **0.006** | **0.515**** | **0.195**** | **0.429**** | **0.068*** | **0.145**** | **0.215**** | **0.042** | **0.069*** | **-0.696**** | **-0.369**** | **-0.051** | **0.006** |
| (N = 16,357) | (0.004) | (0.030) | (0.042) | (0.032) | (0.026) | (0.035) | (0.037) | (0.026) | (0.026) | (0.084) | (0.055) | (0.061) | (0.045) |
| with controls | | | 0.192** | 0.426** | 0.066* | 0.138** | 0.209** | 0.038 | 0.067* | -0.719** | -0.392** | -0.074 | -0.007 |
| | | | (0.041) | (0.032) | (0.026) | (0.034) | (0.036) | (0.025) | (0.026) | (0.083) | (0.053) | (0.062) | (0.044) |
| **2. Girls** | **0.007+** | **0.549**** | **0.215**** | **0.519**** | **0.104**** | **0.175**** | **0.232**** | **0.047** | **0.100*** | **-0.657**** | **-0.407**** | **-0.023** | **0.020** |
| (N =8,051) | (0.004) | (0.031) | (0.051) | (0.037) | (0.038) | (0.051) | (0.050) | (0.035) | (0.039) | (0.088) | (0.062) | (0.081) | (0.045) |
| with controls | | | 0.217** | 0.517** | 0.103** | 0.171** | 0.233** | 0.044 | 0.099* | -0.670** | -0.422** | -0.038 | 0.012 |
| | | | (0.048) | (0.037) | (0.038) | (0.049) | (0.048) | (0.035) | (0.039) | (0.087) | (0.060) | (0.080) | (0.044) |
| **3. Boys** | **-0.001** | **0.482**** | **0.164**** | **0.331**** | **0.025** | **0.103*** | **0.185**** | **0.033** | **0.031** | **-0.763**** | **-0.360**** | **-0.114** | **-0.030** |
| (N =8,306) | (0.004) | (0.034) | (0.046) | (0.040) | (0.038) | (0.040) | (0.040) | (0.036) | (0.037) | (0.102) | (0.082) | (0.078) | (0.087) |
| with controls | | | 0.164** | 0.328** | 0.026 | 0.102* | 0.183** | 0.032 | 0.032 | -0.769** | -0.364** | -0.115 | -0.034 |
| | | | (0.045) | (0.040) | (0.038) | (0.039) | (0.038) | (0.036) | (0.036) | (0.097) | (0.080) | (0.078) | (0.085) |
| **4. Non-FRL** | **0.004** | **0.521**** | **0.189**** | **0.452**** | **0.068*** | **0.163**** | **0.209**** | **0.059+** | **0.071*** | **-0.695**** | **-0.345**** | **-0.071** | **-0.000** |
| (N = 11,281) | (0.003) | (0.033) | (0.051) | (0.038) | (0.031) | (0.042) | (0.045) | (0.030) | (0.030) | (0.093) | (0.062) | (0.067) | (0.057) |
| with controls | | | 0.183** | 0.446** | 0.062* | 0.152** | 0.201** | 0.053+ | 0.065* | -0.717** | -0.370** | -0.097 | -0.018 |
| | | | (0.049) | (0.039) | (0.030) | (0.041) | (0.043) | (0.028) | (0.029) | (0.095) | (0.059) | (0.065) | (0.055) |
| **5. FRL** | **0.002** | **0.501**** | **0.214**** | **0.370**** | **0.064** | **0.097+** | **0.228**** | **-0.007** | **0.058** | **-0.718**** | **-0.439**** | **-0.031** | **-0.002** |
| (N = 5,076) | (0.006) | (0.036) | (0.053) | (0.041) | (0.042) | (0.054) | (0.048) | (0.041) | (0.044) | (0.108) | (0.113) | (0.099) | (0.081) |
| with controls | | | 0.222** | 0.373** | 0.073+ | 0.104+ | 0.237** | -0.002 | 0.066 | -0.713** | -0.432** | -0.013 | 0.002 |
| | | | (0.051) | (0.041) | (0.041) | (0.054) | (0.049) | (0.041) | (0.043) | (0.102) | (0.111) | (0.102) | (0.081) |

Notes: Table entries are 2SLS estimates for the effect of entry into GEM in sixth grade, from fuzzy RD models as specified in equation (3) in the text. (I.e., the entries are estimates of beta). The baseline models include four dummy variables for the student's cohort but no other controls. The models with controls include dummies for student gender, FRL and ELL status, and race; age; and quadratics in fourth-grade math and reading scores. Standard errors clustered on the middle school where the student was enrolled in 6th grade.

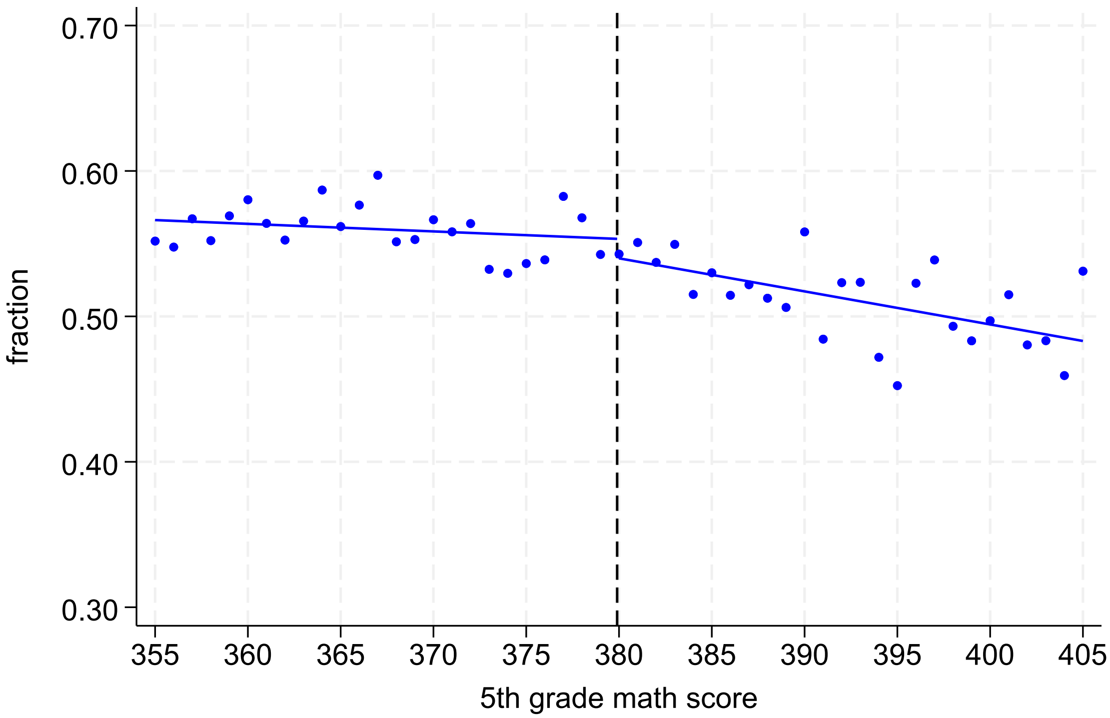
**Table 3. Estimated Treatment Effects on Post-Secondary Outcomes**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Bachelors degree within 5 years in (field): | | | | | Longer windows | | |
| | On-Time High School Grad | On-Time College Entry | Enroll in Selective College | Enroll & Persist 1+ Yrs, Any College | Any College Degree in 5 Years | Any Bachelors Degree in 5 Years | STEM | STEM or Bus/Econ | Health or Educ. | Bus/Econ | Health | College Entry within 2 Years | Any Bachelors Degree in 6 Years | STEM or Bus/Econ Degree in 6 Years |
| **1. Pooled** | **-0.005** | **-0.005** | **0.019** | **0.014** | **0.035** | **0.019** | **0.018** | **0.045+** | **-0.019** | **0.026** | **-0.021+** | **-0.003** | **0.015** | **0.044+** |
| (N = 16,357) | (0.010) | (0.025) | (0.031) | (0.027) | (0.030) | (0.030) | (0.018) | (0.024) | (0.013) | (0.017) | (0.011) | (0.021) | (0.029) | (0.025) |
| with controls | -0.007 | -0.012 | 0.007 | 0.006 | 0.022 | 0.005 | 0.017 | 0.042+ | -0.022 | 0.023 | -0.023* | -0.009 | 0.002 | 0.041+ |
| | (0.010) | (0.024) | (0.030) | (0.026) | (0.030) | (0.029) | (0.017) | (0.024) | (0.013) | (0.017) | (0.011) | (0.021) | (0.028) | (0.024) |
| **2. Girls** | **0.008** | **0.029** | **0.080*** | **0.041** | **0.083*** | **0.040** | **0.050*** | **0.091*** | **-0.025** | **0.040** | **-0.031+** | **0.027** | **0.069+** | **0.107*** |
| (N =8,051) | (0.011) | (0.025) | (0.038) | (0.032) | (0.039) | (0.039) | (0.021) | (0.037) | (0.021) | (0.024) | (0.018) | (0.022) | (0.036) | (0.041) |
| with controls | 0.006 | 0.025 | 0.066+ | 0.034 | 0.072+ | 0.025 | 0.050* | 0.087* | -0.027 | 0.037 | -0.031+ | 0.023 | 0.055 | 0.103* |
| | (0.011) | (0.024) | (0.033) | (0.032) | (0.038) | (0.037) | (0.020) | (0.035) | (0.021) | (0.023) | (0.018) | (0.021) | (0.034) | (0.039) |
| **3. Boys** | **-0.023** | **-0.052** | **-0.060** | **-0.026** | **-0.030** | **-0.018** | **-0.018** | **-0.005** | **-0.016** | **0.011** | **-0.014** | **-0.044** | **-0.058** | **-0.025** |
| (N =8,306) | (0.021) | (0.041) | (0.043) | (0.042) | (0.040) | (0.047) | (0.026) | (0.035) | (0.014) | (0.024) | (0.014) | (0.037) | (0.046) | (0.033) |
| with controls | -0.021 | -0.049 | -0.057 | -0.024 | -0.031 | -0.019 | -0.018 | -0.007 | -0.017 | 0.009 | -0.015 | -0.041 | -0.058 | -0.027 |
| | (0.020) | (0.041) | (0.040) | (0.042) | (0.041) | (0.046) | (0.026) | (0.034) | (0.014) | (0.024) | (0.014) | (0.037) | (0.045) | (0.033) |
| **4. Non-FRL** | **-0.005** | **-0.014** | **0.027** | **0.049+** | **0.023** | **0.023** | **0.019** | **0.045** | **-0.038*** | **0.025** | **-0.038*** | **-0.013** | **0.008** | **0.045** |
| (N = 11,281) | (0.012) | (0.029) | (0.034) | (0.028) | (0.037) | (0.031) | (0.019) | (0.028) | (0.017) | (0.022) | (0.015) | (0.024) | (0.032) | (0.028) |
| with controls | -0.006 | -0.020 | 0.017 | 0.041 | 0.011 | 0.008 | 0.017 | 0.043 | -0.043* | 0.023 | -0.042** | -0.018 | -0.005 | 0.043 |
| | (0.012) | (0.029) | (0.035) | (0.027) | (0.037) | (0.031) | (0.019) | (0.027) | (0.017) | (0.021) | (0.015) | (0.024) | (0.033) | (0.027) |
| **5. FRL** | **-0.007** | **0.008** | **-0.018** | **-0.082** | **0.049** | **-0.010** | **0.012** | **0.038** | **0.022** | **0.026** | **0.015** | **0.011** | **0.015** | **0.033** |
| (N = 5,076) | (0.022) | (0.048) | (0.058) | (0.065) | (0.054) | (0.065) | (0.036) | (0.044) | (0.018) | (0.025) | (0.019) | (0.045) | (0.056) | (0.045) |
| with controls | -0.003 | 0.016 | -0.010 | -0.075 | 0.055 | -0.003 | 0.015 | 0.041 | 0.025 | 0.026 | 0.018 | 0.018 | 0.025 | 0.037 |
| | (0.022) | (0.044) | (0.051) | (0.062) | (0.052) | (0.062) | (0.036) | (0.043) | (0.018) | (0.025) | (0.018) | (0.043) | (0.052) | (0.045) |

Notes: Table entries are 2SLS estimates for the effect of entry into GEM in sixth grade, from fuzzy RD models as specified in equation (3) in the text. (I.e., the entries are estimates of beta). The baseline models include four dummy variables for the student's cohort but no other controls. The models with controls include dummies for student gender, FRL and ELL status, and race; age; and quadratics in fourth-grade math and reading scores. Standard errors clustered on the middle school where the student was enrolled in 6th grade.

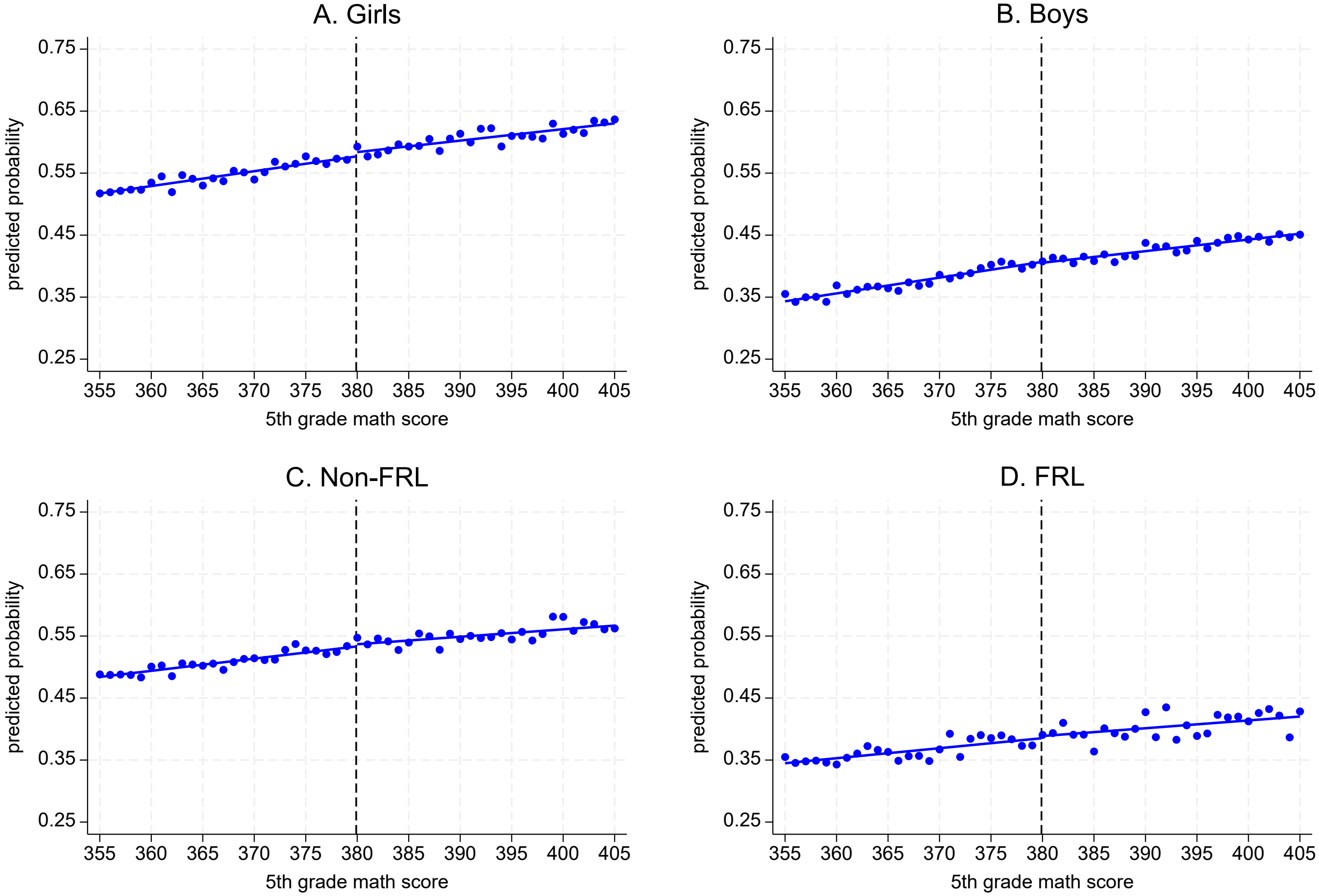
# Appendix A: Additional Figures and Tables

**Figure A1: Reduced-Form Relationship for Meeting Sample Criteria**



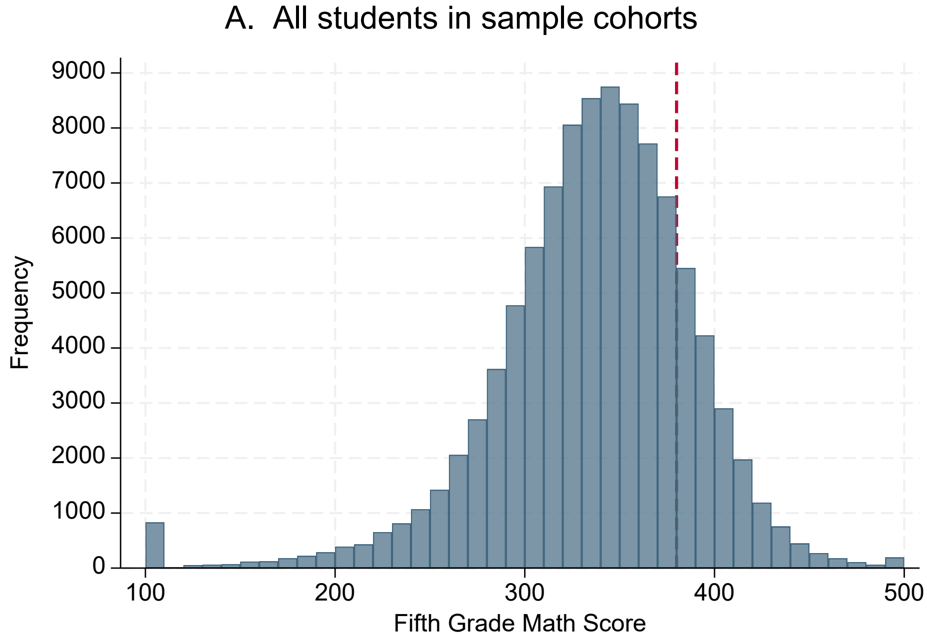

**Notes:** This figure plots means, for the full population of students enrolled in the District for fifth grade between 2004 and 2008, of an indicator for meeting the inclusion criteria for the main analysis sample (see Appendix Table A1 for details). It also shows fitted values from a local linear RD model with a symmetric bandwidth of 25. The estimated discontinuity is -0.013 with a clustered standard error (clustered on elementary school) of 0.011.

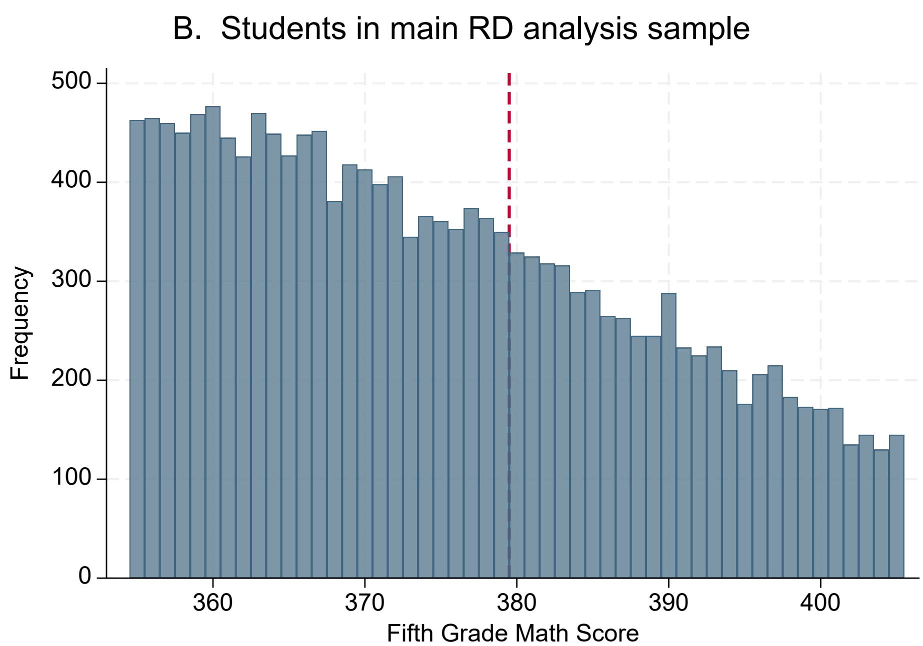
**Figure A2: Reduced-Form Relationships for Predicted On-Time Graduation and College Degree within 5 Years, by Gender and Disadvantage**




**Notes**: This figure plots sample means, along with fitted values from local linear RD models with symmetric bandwidths of +/- 25 points, of the predicted probability of graduating high school on time and completing any college degree (associates or bachelors) within 5 years. The prediction models include second-order polynomials in 4$^{th}$-grade math and reading test scores; student age; and indicators for student race/ethnicity, FRL status, ELL status, and cohort.

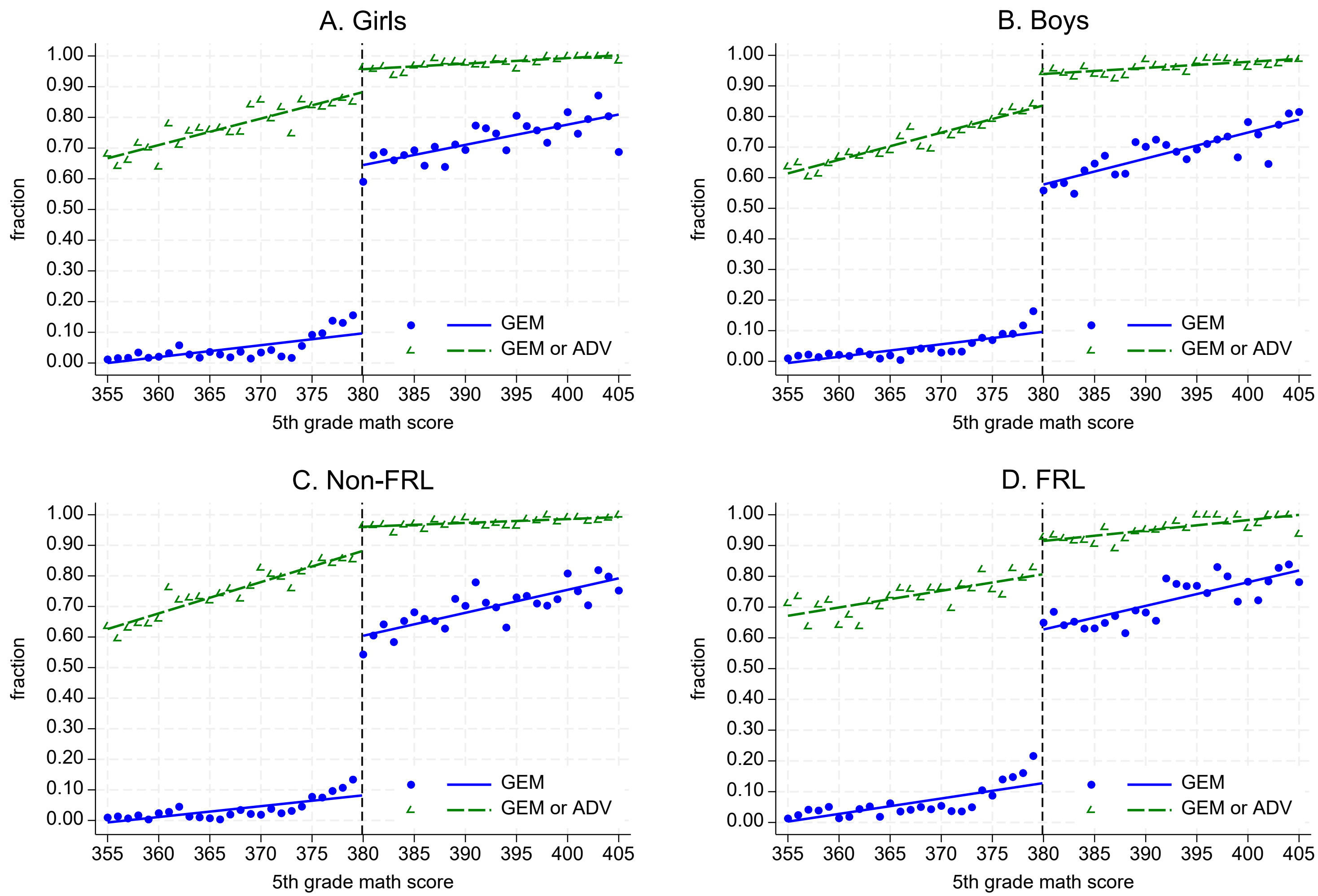
## Figure A3: Histogram of Running Variable




**Notes**: This figure shows the frequency distribution of fifth-grade math scores for all students in our sample cohorts (Panel A) and for students in the main RD analysis sample (Panel B). The sample consists of all students in the District who completed fifth grade in between 2003 and 2007, were not identified as gifted, had non-missing fourth-grade test scores, were enrolled in sixth grade in the following year (between 2004 and 2008), and were enrolled in (or graduated from) a District high school six years later (between 2010 and 2014). The RD analysis sample is further restricted to scores within +/- 25 points of the threshold for eligibility to participate in the GEM program in sixth grade. The dashed vertical lines in each panel indicate the threshold.

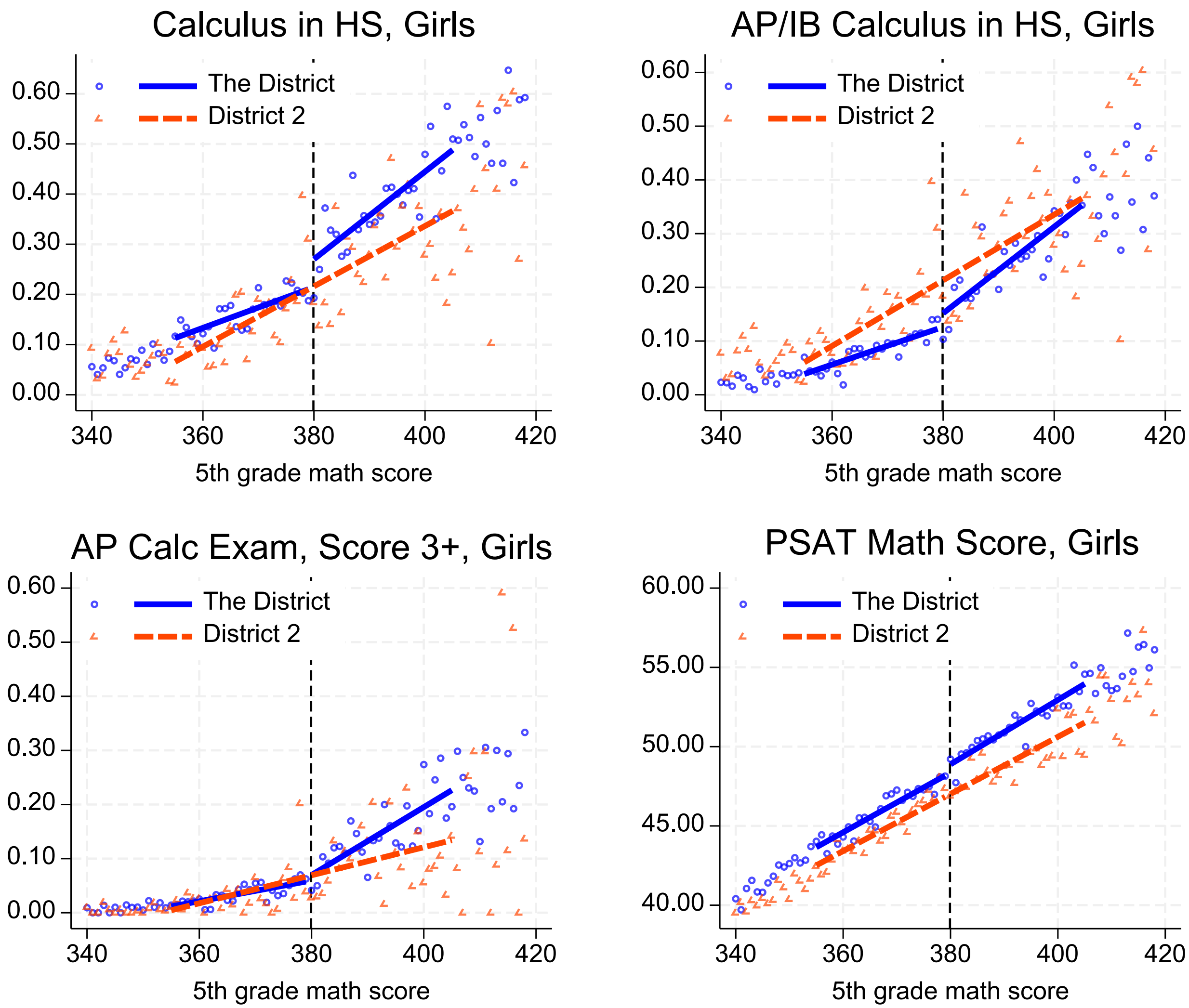
## Figure A4: First-Stage Relationships for GEM Enrollment, by Gender and Disadvantage




**Notes**: This figure plots sample means of an indicator for participating in the GEM program in sixth grade (the treatment variable, in blue circles) and an indicator for participating in *either* GEM *or* advanced-track math (in green triangles) at each value of fifth-grade math score (the running variable), along with fitted values from local linear RD models with a symmetric bandwidth of 25.

## Figure A5. High School Math Outcomes, Between-District Comparison

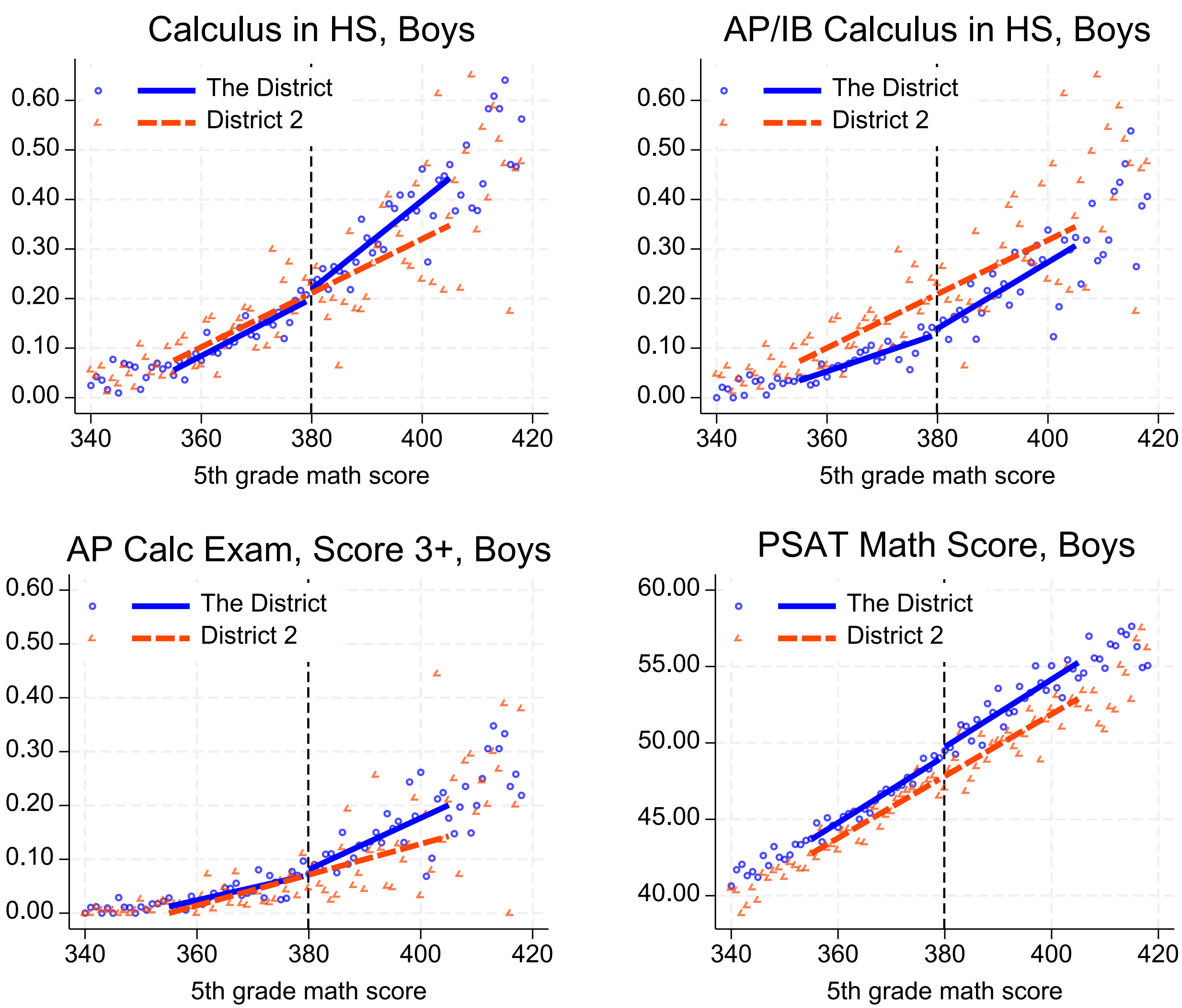
### A. Girls




**Notes**: This figure plots sample means of each variable at each value of fifth-grade math score, separately for the District and a comparison district (District 2) that did not offer an accelerated track like GEM for the highest performers but enrolled a larger share of lower-performers in algebra in 8th grade. Also shown are fitted values from local linear RD models with a symmetric bandwidth of 25 for the District (solid blue lines), and fitted values from linear models (with no discontinuity) using the same bandwidth for District 2 (orange dashed lines). Observations for District 2 are reweighted using propensity scores constructed from a logit model for the probability of being in BCPS , estimated using all students in BCPS and HCPS who meet our main sample criteria, separately for boys and girls. The model includes dummies for 25-point bins of g5math; dummies for race, ELL and FRL status, a quadratic in age, 4th-grade math and reading scores interacted with race, and 5th-grade reading scores.

**Figure A5. High School Math Outcomes, Between-District Comparison (Cont'd.)**

**B. Boys**




**Notes**: This figure plots sample means of each variable at each value of fifth-grade math score, separately for the District and a comparison district (District 2) that did not offer an accelerated track like GEM for the highest performers but enrolled a larger share of lower-performers in algebra in 8th grade. Also shown are fitted values from local linear RD models with a symmetric bandwidth of 25 for the District (solid blue lines), and fitted values from linear models (with no discontinuity) using the same bandwidth for District 2 (orange dashed lines). Observations for District 2 are reweighted using propensity scores constructed from a logit model for the probability of being in BCPS , estimated using all students in BCPS and HCPS who meet our main sample criteria, separately for boys and girls. The model includes dummies for 25-point bins of g5math; dummies for race, ELL and FRL status, a quadratic in age, 4th-grade math and reading scores interacted with race, and 5th-grade reading scores.

**Appendix Table A1: Estimated female wage gap with and without controls for field of degree**

| | Estimated female effect: | |
|---|---|---|
| | No controls for age (1) | Including age controls (2) |
| 1. No other controls | -0.328 | -0.328 |
| | (0.003) | (0.003) |
| 2. Controls for education (2 dummies) | -0.312 | -0.312 |
| and class of main job (7 dummies) | (0.003) | (0.003) |
| 3. Controls for education, job class and | -0.269 | -0.270 |
| STEM/business or economics ( 2 dummies) | (0.003) | (0.003) |
| 4. Controls for education, job class and field | -0.247 | -0.247 |
| of degree (172 dummies) | (0.003) | (0.003) |
| Share of gender effect in row 2 attributable to: | | |
| - STEM/business | 13.8% | 13.5% |
| - detailed field of degree | 20.8% | 20.8% |

Notes: table entries are estimated coefficients of indicator for female in models for log annual wage and salary earnings. Sample includes native born people age 30-35 in 2010-2018 American Community Survey (ACS) with BA degree or higher. Models in row 1 include no other controls (except age and age*female in column 2). Models in row 2 add controls for education (2 dummies) and class of main job (e.g., private, state and local goverment, etc.). Models in row 3 add 2 dummies for STEM fields and for fields in business or economics. Models in row 4 add dummies for the full set of 4-digit field of study variables in ACS. All models in column 2 also include a linear age term and female interacted with (age-32.5), so the coefficient for the main effect of female gender represents the wage gap at age 32.5 (the mid-point of the age range in the sample).

**Appendix Table A2. Sample Criteria & Sample Means**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | All in Grade 5 | Observed in Grades 4-6 | Traditional Middle School | Followed Through HS Grad | Not Gifted | Main RD Analysis Sample | | | | |
| | | | | | | Pooled | Girls | Boys | Non-FRL | FRL |
| **Demographic characteristics** | | | | | | | | | | |
| female | 0.49 | 0.49 | 0.49 | 0.51 | 0.52 | 0.49 | 1.00 | 0.00 | 0.49 | 0.50 |
| FRL | 0.43 | 0.43 | 0.44 | 0.41 | 0.42 | 0.31 | 0.32 | 0.30 | 0.00 | 1.00 |
| White | 0.35 | 0.35 | 0.36 | 0.36 | 0.34 | 0.44 | 0.43 | 0.46 | 0.56 | 0.17 |
| Black | 0.34 | 0.34 | 0.34 | 0.34 | 0.35 | 0.23 | 0.25 | 0.21 | 0.13 | 0.44 |
| Hispanic | 0.24 | 0.24 | 0.24 | 0.24 | 0.24 | 0.25 | 0.24 | 0.26 | 0.22 | 0.31 |
| **Baseline achievement** | | | | | | | | | | |
| grade 4 math score (standardized) | 0.08 | 0.08 | 0.09 | 0.18 | 0.09 | 0.65 | 0.63 | 0.68 | 0.70 | 0.54 |
| grade 4 reading score (standardized) | 0.08 | 0.08 | 0.08 | 0.17 | 0.09 | 0.53 | 0.62 | 0.45 | 0.60 | 0.38 |
| **School peer characteristics (6th grade)** | | | | | | | | | | |
| avg grade 4 math & reading scores | 0.07 | 0.07 | 0.07 | 0.09 | 0.08 | 0.13 | 0.13 | 0.14 | 0.20 | -0.02 |
| fraction FRL | 0.46 | 0.46 | 0.46 | 0.45 | 0.46 | 0.41 | 0.42 | 0.41 | 0.35 | 0.56 |
| fraction GEM eligible | 0.20 | 0.20 | 0.20 | 0.20 | 0.20 | 0.22 | 0.22 | 0.22 | 0.24 | 0.17 |
| **Math track in 6th grade** | | | | | | | | | | |
| GEM (treatment) | 0.16 | 0.16 | 0.16 | 0.18 | 0.14 | 0.28 | 0.28 | 0.27 | 0.29 | 0.26 |
| advanced math | 0.27 | 0.27 | 0.28 | 0.31 | 0.31 | 0.55 | 0.56 | 0.53 | 0.55 | 0.54 |
| regular math | 0.57 | 0.57 | 0.55 | 0.51 | 0.55 | 0.18 | 0.16 | 0.19 | 0.17 | 0.20 |
| **Number of observations** | 98,709 | 79,494 | 73,803 | 51,226 | 47,439 | 16,357 | 8,051 | 8,306 | 11,281 | 5,076 |

The sample in column 1 includes five cohorts of students who entered fifth grade in 2003-2007. Column 2 is restricted to students who were observed in 4th grade (and have 4th-grade test scores) and were also observed in 6th grade the following year (2004-2008). Column 3 further restricts the sample to those enrolled in a traditional (e.g., non-charter) middle school in 6th grade. Column 4 is limited to students who were enrolled in (or graduated from) a District high school six years later (between 2010 and 2014). Column 5 drops students who were identified as gifted (which usually occurs no later than 5th grade). Finally, columns 6-10 report means for the analysis samples for the RD models, which include students who meet the criteria in column 5 and also had 5th-grade math score within 25 points of the GEM eligiblity threshold (our preferred RD bandwidth).

**Appendix Table A3. More High School Outcomes**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | Took Any Advanced Math Course | Took AP/IB Calculus Course | Took Calc AP Exam + Got Score of >=3 | High School GPA (Overall) | PSAT Math | PSAT Verbal | Took SAT/ACT |
| **1. Pooled** | **0.077**** | **0.047*** | **0.024** | **0.030** | **2.453*** | **1.350** | **-0.004** |
| (N = 16,357) | (0.023) | (0.021) | (0.016) | (0.027) | (1.009) | (1.443) | (0.020) |
| **2. Girls** | **0.093*** | **0.055** | **0.021** | **0.079*** | **3.477*** | **3.090** | **0.031** |
| (N =8,051) | (0.036) | (0.034) | (0.023) | (0.033) | (1.434) | (1.954) | (0.019) |
| **3. Boys** | **0.051** | **0.036** | **0.027** | **-0.044** | **1.381** | **-1.044** | **-0.050** |
| (N =8,306) | (0.033) | (0.022) | (0.020) | (0.055) | (1.612) | (2.030) | (0.037) |
| **4. Non-FRL** | **0.055*** | **0.050+** | **0.032*** | **-0.005** | **3.465**** | **2.118** | **-0.017** |
| (N = 11,281) | (0.025) | (0.026) | (0.016) | (0.031) | (1.239) | (1.383) | (0.020) |
| **5. FRL** | **0.131**** | **0.037** | **0.001** | **0.091** | **-0.886** | **-1.326** | **0.016** |
| (N = 5,076) | (0.047) | (0.036) | (0.034) | (0.057) | (1.944) | (3.203) | (0.048) |

Notes: Table entries are 2SLS estimates for the effect of entry into GEM in sixth grade, from fuzzy RD models as specified in equation (3) in the text. (I.e., the entries are estimates of beta). The models include four dummy variables for the student's cohort but no other controls. Standard errors clustered on the middle school where the student was enrolled in 6th grade.