



ROCKWOOL Foundation Berlin

Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES

261/26

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

David Card, Laura Giuliano

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

Authors

David Card, Laura Giuliano

Reference

JEL Codes: I21

Keywords: mathematics instruction, college major, noncognitive skills

Recommended Citation: David Card, Laura Giuliano (2026): Targeted Acceleration in Middle School Math: Impacts on College Entry, Degree Completion, and STEM. RFBerlin Discussion Paper No. 261/26

Access

Papers can be downloaded free of charge from the RFBerlin website: <https://www.rfberlin.com/discussion-papers>

Discussion Papers of RFBerlin are indexed on RePEc: <https://ideas.repec.org/s/crm/wpaper.html>

Disclaimer

Opinions and views expressed in this paper are those of the author(s) and not those of RFBerlin. Research disseminated in this discussion paper series may include views on policy, but RFBerlin takes no institutional policy positions. RFBerlin is an independent research institute.

RFBerlin Discussion Papers often represent preliminary or incomplete work and have not been peer-reviewed. Citation and use of research disseminated in this series should take into account the provisional nature of the work. Discussion papers are shared to encourage feedback and foster academic discussion.

All materials were provided by the authors, who are responsible for proper attribution and rights clearance. While every effort has been made to ensure proper attribution and accuracy, should any issues arise regarding authorship, citation, or rights, please contact RFBerlin to request a correction.

These materials may not be used for the development or training of artificial intelligence systems.

Imprint

RFBerlin
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 151 143 444 67
E-mail: info@rfberlin.com
Web: www.rfberlin.com



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

David Card and Laura Giuliano¹

September 2026

Abstract

We study an accelerated three-year middle school math curriculum in a large urban district that lets students complete Algebra I and Geometry before high school. Admission is based on 5th-grade math scores, with a cutoff near the 80th percentile. Continued eligibility requires at least a B grade in the previous year; program completion is highly correlated with an index of non-cognitive skills. Following students for six years after high school, we find treatment effects that vary widely by non-cognitive skill and gender. Marginally eligible boys with high non-cognitive skills are 33 percentage points (ppt.) more likely to complete calculus by 11th grade and 12 ppt. more likely to complete a STEM or business major within six years. Boys with lower non-cognitive skills take more advanced high school math but show no significant college effects. Like their male peers, girls with high non-cognitive skills are more likely to complete calculus by 11th grade (+17 ppt.). Their rate of STEM/business majoring is unaffected, but they shift from social sciences to core STEM fields. Girls in the lower two non-cognitive terciles are also strongly affected, with more advanced high school math, higher entry to selective four-year colleges, and more STEM/business degrees (+12 ppt.).

¹ 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. We also received helpful feedback from Sarah Cohodes, Thomas Dee, and participants at the 2025 AEA Annual Meeting, the Spring 2025 NBER Education Meeting, the Fall 2024 NBER Economics of Talent Meeting, the 2025 EALE/SOLE Meeting, and from seminar participants at Stanford University (CEPR), UC Berkeley (ILRE), UC Davis (Economics), and University of Maryland (AREC). 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. In addition, Giuliano gratefully acknowledges support from the Polynera Fund.

The question of how to structure the math curriculum in U.S. public schools is widely debated. Many school districts offer accelerated tracks with algebra and even geometry in middle school, allowing students to complete higher-level math classes in high school (see e.g., Wolfe et al., 2023), though access to these courses varies widely even within districts (Peters and Carter, 2023). In the 2000’s several states attempted to push all students into algebra in 8th grade (one year ahead of the traditional start). Within just a few years, however, this effort was abandoned.² More recently some districts have tried to *delay* algebra until 9th grade for even high-achieving students, citing concerns over the impacts on low-achievers (see e.g., Huffaker et al. 2025 and Huffaker, 2023). Nevertheless, the demand for accelerated math in middle school is strong (Yoder, 2024), and enrollment in algebra and geometry in 7th and 8th grades is 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) offers pre-algebra in 6th grade, Algebra I in 7th grade, and geometry in 8th grade, accelerating students two years beyond the traditional curriculum, and one year beyond the advanced track offered in the District and other Florida districts. Eligibility for GEM 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 GEM entry are identified in a regression discontinuity (RD) analysis – have relatively strong math skills.⁴ Indeed, nearly all the students who barely miss the GEM threshold enter the advanced math track in 6th grade.

We use information on middle and high school course-taking and grades, combined with data from the National Student Clearinghouse (NSC) on college enrollment and degree completion, to study the effects of entry to the GEM program in 6th grade. Focusing on students whose 5th-grade scores are within a narrow bandwidth ($\pm 0.6\sigma$ ’s) of the entry threshold, we

² 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 Data collected by the U.S. Department of Education (2022) show that in the 2020-21 academic year, 4.4% of 7th grade students and 24% of 8th grade public school students were enrolled in Algebra I. Nationwide about 4% of 8th grade students were enrolled in Geometry, with rates of 13% in Virginia, 10% in Florida, and 9% in Utah.

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.

follow 23,000 students for 13 years – 6 years after they would normally graduate from high school. Our relatively large sample and long follow-up period enable us to conduct an informative RD analysis of the effects of GEM on high school math attainment, college completion, and major choice.

A key feature of GEM is that students must earn at least a B grade to continue in the program and receive high school credit for their work. Only about one-third of marginally eligible entrants successfully complete all three years. Our first main finding is that completion of GEM is highly correlated with an index of non-cognitive skills, formed by taking average high school course grades in subjects other than math and regression-adjusting for 5th-grade test scores in math and reading.⁵ Indeed, the GEM completion rate is close to zero for marginally eligible students with the lowest levels of non-cognitive skill (most of whom are boys), and around 60% for those with the highest levels (most of whom are girls). In light of these differences, we conduct much of our subsequent analysis stratifying students by gender and within-gender tercile of measured non-cognitive skill.

At the high school level, we find a heterogeneous pattern of effects for marginally eligible GEM entrants. For these students as a whole, entry to GEM leads to a rise in the number of advanced math classes (including AP's) taken in high school (+0.4 courses, $t \approx 4$), a rise in the probability of completing calculus by 11th grade (+13 ppts, $t \approx 5$), a slightly smaller rise in the probability of completing calculus by 12th grade (+10 ppts, $t \approx 4$), and a rise in PSAT math scores of roughly 0.1σ ($t \approx 3$). The impact on early completion of calculus is particularly salient for students hoping to be admitted to selective science and engineering programs, since they can include their calculus grade in their application.⁶

As might be expected given the variation in GEM completion rates, these impacts are unequally distributed by gender and non-cognitive (NC) skill level. For girls in the top tercile of NC skills we find a large positive effect on the probability of completing calculus by 11th grade (+18 percentage points, ppts.), but no significant effect on the number of advanced math courses

⁵ The idea of using residualized course grades as a measure of non-cognitive ability was suggested by Jacob (2002) and is consistent with the conclusions of Borghans et al. (2016) that grades are more affected by non-cognitive skills than are achievement test scores. A concern with the use of high school grades is that they may be affected by GEM participation. We show below, however, that there is no discernable discontinuity in average grades in non-math subjects at the GEM entry cutoff. We also find no discontinuity in any quantile of the non-cognitive index, including the tercile cutoffs that we use to stratify our analysis sample (see Appendix C).

⁶ Elite science and engineering programs often look for evidence of successful completion of calculus – see e.g., UC Berkeley College of Engineering (2024). Relatively few students in the District, however, apply to such programs.

or the probability of completing calculus by the end of high school, arguably reflecting the high rates for untreated compliers in this subgroup.⁷ For girls with lower NC skills, we find that GEM participation leads to a 0.4-0.5 course rise in the number of advanced math courses in high school and relatively large increases in calculus enrollment by 12th grade (+14 pts). For boys in the top tercile of NC skills we find large positive effects on the probability of completing calculus by 11th grade (+33 pts.) and on the probability of having any calculus by end of high school (+17 pts), combined with a sizeable gain (+0.3 σ 's) in math PSAT scores. For boys in the lower terciles there is also an increase in the number of advanced math classes (+0.3 courses), and a rise in the probability of completing calculus by 11th grade for the middle tercile group (+10 pts), but no significant effects on the share with calculus by the end of high school.

At the college level, NSC data allow us to measure the type of college entered after high school, degree completion (including AA and BA/BS degrees), and field of the degree for students who complete a bachelor's degree. Building on the taxonomy developed by the NSF (2023), we identify two main sets of fields that require a more intensive math background: STEM (Science, Technology, Engineering, and Mathematics) and business.⁸ Across students as a whole, entry to GEM leads to no significant impact on college entry or degree completion, but results in a +5.5 ppt. increase ($t \approx 2.4$) in the share completing a degree in STEM or business, a 25% rise relative to the baseline rate of the untreated compliers.

Again, these impacts vary widely by gender and NC tercile. For girls in the top NC tercile, there is no effect on the share graduating with a bachelor's degree in STEM or business (which is already around 45% for the untreated compliers). There is, however, a shift in majors *within* the STEM fields, with a marginally significant 7 ppt. rise in degrees in hard/life sciences offset by a decline in social sciences. For girls in the middle tercile of NC ability, participation in GEM leads to a significant 12 ppt. rise in the share who graduate with a STEM/business degree, attributable to a rise in the share who attain any bachelor's degree and a fall in the share who earn a degree outside of STEM or business. Marginally eligible girls in the bottom tercile of NC

⁷ For simplicity we refer to the mean potential outcome of compliers in the absence of GEM (i.e., the mean of $Y(0)$ in potential outcomes notation) as the mean outcome for the “untreated compliers”.

⁸ See Appendix B. Following NSF, we include social sciences in STEM but exclude health-related majors such as nursing. We analyzed the math intensity of different majors using O*NET ratings of the math requirements for the occupations that each major feeds into, derived from the NCES/BLS crosswalk of college major codes and occupation codes. As shown in Figure B.1, business majors lead to occupations in which the level and intensity of math use are comparable to those associated with majors in biology, computer science, and psychology.

skill show the largest responses to GEM participation, with a 14 ppt. rise in the share who enter a selective 4-year college after high school ($t \approx 2.4$), a 13 ppt. rise in the share who earn a bachelor’s degree after 6 years ($t \approx 2.2$), and a 14 ppt. rise in the share who graduate in STEM or business ($t \approx 2.8$). This pattern suggests that marginally eligible girls who lack gender-stereotypical non-cognitive skills benefit from GEM by shifting toward more “math-centered” courses in high school, leading to better college options and more STEM and business degrees.

For boys in the top tercile of NC skill, GEM entry leads to a 12 ppt. ($t \approx 2.4$) rise in the probability of majoring in STEM or business. Given the 43% rate for untreated compliers, this means that well over one-half of the treated compliers end up with a bachelor’s degree in STEM or business 6 years after high school. In contrast, GEM entry has no significant effects on the college outcomes for boys in the two lower tiers of NC ability. If anything, entry to GEM for marginally eligible boys in the lowest tier of NC skills is associated with slightly lower rates of entering a selective college (-6 ppt. effect, $t \approx 1.4$) and of completing a STEM or business degree (-4 ppt. effect, $t \approx 1.9$). The low exposure of low-NC boys to the GEM curriculum, however, and the sensitivity of the point estimates leads us to conclude that they are unlikely to reflect causal effects.⁹

A concern for our RD design is that the presence of GEM could affect students who narrowly missed the threshold (i.e., the untreated compliers). The existing literature points to three potential channels for such spillovers: (1) negative peer effects, arising from the shift of the higher-scoring students to the GEM classes (Lefgren, 2004; Paloyo, 2020); (2) discouragement effects, arising from the stigma of being left out of the advanced math stream (Buser and Yuan, 2019; Ellison and Swanson, 2023); (3) “top of the class” effects (Denning et al., 2023) arising because in the presence of GEM the untreated compliers become among the highest scoring students in their math 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.

To address this concern, we use data from another large urban school district in the same state – “District B” – which has no GEM program but strongly pushes all students with above-

⁹ As we discuss further below, most of these boys left GEM after 6th grade. In addition, in the between-district comparisons discussed below we have to drop one cohort of students and we find that in the resulting sample the point estimates of the effect of GEM entry for low-NC boys fall to near zero.

average math scores into an advanced track with Algebra I in 8th grade. Consistent with that policy, rates of completion of Algebra I by 8th grade are much higher in District B than in the District but trend smoothly through the GEM cutoff.

Relative to students in District B, we find that students in the District with scores below the GEM cutoff have lower rates of completing calculus by the end of high school, with gaps that become more negative just below the threshold for some groups. There are similar downward trends – albeit much smaller and statistically insignificant – in the between-district gaps in STEM/business majoring. While the pattern of gaps is suggestive of discouragement or peer effects, three key facts point to an alternative explanation. First, the gaps follow linear trends or are flat rather than exhibiting a dip below the threshold as might be expected if there were negative spillovers. Second, there are no such gaps in *early* calculus enrollment (by 11th grade). Third, the gaps in STEM/business majoring for high-NC girls – the group that is a focus of earlier work on discouragement effects – actually narrow continuously around the GEM threshold. Rather than spillovers, we suspect that the between-district gaps among students who are ineligible for GEM reflect the differences between District B’s broad-based acceleration policy and the more targeted policy of the District. Our cross-district comparisons underscore the point that the outcomes of the untreated compliers in an RD design depend on the underlying policy environment and could be different in a different setting.

Our empirical findings contribute to three main literatures. The first is a relatively small set of studies that have tried to identify the causal effects of accelerated math 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 harmful for weaker students – an effect that led to the suspension of the initiatives. Subsequent research by Dougherty et al. (2017) showed that a more targeted approach – leaving out the lowest-scoring quintile of students – has a positive effect on math course-taking of girls, including enrollment in pre-calculus in 11th grade, with little effect on boys. (They do not follow students to college.) Cohodes (2020) studies an accelerated curriculum 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.¹⁰ Most recently, Huffaker et al. (2025) study the effects of a decision by San Francisco Unified School District to eliminate access to Algebra I or geometry in middle school. They find that the policy led to a reduction in the fraction of students taking AP calculus by 12th grade, driven by declines among Asian and Pacific Islander students, but they do not track students beyond high school.

A second, related literature addresses within-school tracking more generally (see Lefgren 2004 for an early study in economics and Betts, 2011 for a review covering much of the older literature). A rigorously designed experimental study of first graders in Kenya (Duflo et al., 2011) suggested that tracking leads to test score gains for both higher and lower-ability students – though it does not address longer run effects. Card and Giuliano (2016) used an RD approach to study the impacts of access to separate “gifted and high achiever” (GHA) classrooms in 4th and 5th grades in the same district we study here. They complemented their RD analysis with a between-school analysis to address spillover effects on ineligible students and concluded that access to GHA classes raised the test scores of eligible non-gifted students, with no spillover effects on those who narrowly missed out. A recent study using data from Texas (Antonovics et al., 2023) shows that within-school tracking of middle and high school students is widespread in the state, and it appears to benefit higher-scoring students without harming lower-scoring students. Horvath (2025) studies classroom assignment in elementary schools in North Carolina and concludes that many schools assign higher and lower ability students to different classrooms even in early grades. Outside the U.S., Humphries et al. (2025) study the impacts of following a STEM track in high school on college outcomes and earnings for Swedish students. On the theoretical side, Ellison and Pathak (2025) address the optimal design of a tracking system when students differ along a single (observable) dimension of heterogeneity. An implication of their model is that in an optimized system there should be no discontinuity in outcomes at the threshold for track assignment.

¹⁰ An older study by Altonji (1995) 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. Özek (2021) studies the impacts of a “remediation track” offered to students who score below a threshold in their 7th-grade math test. He finds that entry to the track raises 8th-grade test scores of marginally eligible students but has negative impacts on participation in more advanced math classes in high school, and no effects on college-related outcomes.

Finally, we also contribute to the literature on college major choice (see Altonji et al., 2016). Much of that literature is concerned with obtaining credible estimates of the returns to different majors (e.g., Kirkebøen et al, 2016; Campos et al. 2026); another important strand focuses on the gender differences in major choice (see McNally 2020 for a recent review). Women are less likely than men to graduate in science, technology, engineering and math (STEM) fields, and in business-related fields.¹¹ While there is a positive 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* in STEM majoring.¹² Our results suggest that exposure to an accelerated math curriculum in middle school affects the major choices of relatively high-scoring students of both genders who have the non-cognitive skills to benefit from the curriculum.

I. Background

a. Conceptual framework

How does entry to an accelerated math program affect college outcomes? We assume that the primary mechanism is through the accumulation of math skills. Students who successfully complete courses like pre-calculus with strong grades can include that information in their college applications and are more likely to be admitted to selective colleges (where STEM majors are more readily available), and to programs like engineering that screen for math preparation. In addition, we believe that students with stronger math skills and the experience of successfully completing harder math classes in high school are more likely to succeed in college courses like calculus and statistics that mediate entry to fields like computer science, business and other constrained majors.

Following the literature on skills production (e.g., Heckman et al., 2006) we assume that math skills at the end of high school (Y^H) depend on a combination of cognitive ability and the initial level of skills (both proxied by the 5th grade test score in math T^m) and non-cognitive ability (N). The effect of GEM entry (indicated by $G \in \{0,1\}$) can then be modelled as a shift in the production function:

¹¹ Some studies – e.g., Cimpian and King (2024) -- focus on a narrower subset of STEM subjects (physics, engineering and computer science), which are the fields with the lowest shares of female students.

¹² 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.

$$Y^H = f(T^m, N) + g(T^m, N) \cdot G + \epsilon \quad (1)$$

where ϵ incorporates all other factors. As discussed below, entry to GEM depends on T^m , with a discontinuity at the threshold score (380 points), which we can normalize to 0. In an interval $\pm R$ around the threshold, equation (1) implies that

$$E[Y^H | T^m, G, N; -R \leq T^m \leq R] \approx \ell(T^m; N) + \beta(N) \cdot G \quad (2)$$

where $\ell(T^m; N)$ is a continuous piece-wise linear control function that can change slope at $T^m = 0$. The impact of GEM for students with a given range of non-cognitive ability, $\beta(N)$, can then be estimated using a fuzzy RD approach (Lee and Lemieux 2010), in which there is a first-stage model relating GEM entry to an indicator for having a test score above the GEM threshold, and a structural equation based on (2) relating the outcome to GEM participation, which is estimated by 2SLS.

GEM is a 3-year program with “off-ramps” for students who fail to achieve at least a B grade in the previous year. Arguably, the impacts of GEM on the accumulation of math skills by the end of high school will depend on the amount of time a student spends in the program. This suggests that we might expect bigger effects of *GEM entry* on the high school outcomes of students who are more likely to complete GEM. This is only part of the story, however, because different students have different rates of skill accumulation even in the absence of GEM. For example, as we show below, nearly one-half of girls with relatively high non-cognitive skills who narrowly miss the GEM threshold enroll in calculus by 12th grade. We wouldn’t necessarily expect large causal effects of GEM entry on calculus participation for these girls, given their high enrollment rates in the absence of GEM.

At the college level, a production-of-skills framework suggests that we could use a framework based on (1) and (2) to model outcomes like “successful completion of a bachelor’s degree in a STEM or related field.” Of interest here is whether there are particular components of Y^H – such as completion of calculus – that are most relevant for certain college outcomes. Suppose that the impact of GEM on a discrete college outcome of interest Y^C is mediated through a particular discrete high school outcome Y^H . Then, looking across subgroups, we might expect that the treatment effect of GEM entry on Y^C for a given subgroup would be proportional to the impact on Y^H .

Implicit in the specification of (1), and in the implementation of an RD model based on (2), is that the presence of GEM has no effect on students who are ineligible for the program.¹³ A concern about tracking programs in general is that there may be classroom-level peer effects that depend on the average ability or achievement of students in the class. In that case, the assignment of higher-achieving students to a different classroom could have a negative effect on lower-achieving students (see e.g., Lefgren 2004), causing an upward bias in the estimated effect of GEM entry.¹⁴ Related mechanisms are “discouragement effects” that cause students who miss the GEM threshold to put less effort into math classes (Ellison and Swanson, 2023), and “top of the class” effects (Denning et al., 2023) that lead to greater engagement and achievement for students who narrowly miss qualifying for GEM threshold.

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 B” – has two main tracks in middle school math, with most of the students who score in the upper half of the statewide math test in 5th grade participating in the upper track. It has no GEM program or third track for very high scoring students. We fit a series of local linear reduced-form RD models for the two districts that include full interactions of the control function and the GEM eligibility indicator with an identifier for the District, while adjusting for differences in the observable characteristics of students measured prior to 5th grade. These models allow us to see how outcomes of students with the same 5th-grade math score differ between the two districts. Negative peer effects or discouragement effects among students in the District who miss the GEM threshold will be evidenced by a negative between-district gap in outcomes that is largest for students with scores just below the GEM threshold.

b. 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; just under one-half are eligible for the federal free- and reduced-price lunch (FRL) program. During the past two decades, the District

¹³ Formally, this is identical to the question of whether there are stable unit treatment effects (SUTVA) when trying to evaluate a particular intervention.

¹⁴ A deeper concern raised by Oakes (1985) is that the curriculum of the lower track will change once the upper track is introduced. We do not believe this is likely to have happened in our setting given the curriculum rules in Florida, but if the advanced track classes are “watered down” when GEM is offered, it could have a negative impact on the math skill development of students who miss the GEM cutoff.

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.

Mainstream middle schools in the District offer three primary tracks of mathematics instruction (taught in separate classrooms): regular, advanced and GEM. There is also a fourth (remedial) 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 6-8.¹⁵ Students in the advanced track who perform well in 7th-grade math, however, can accelerate by taking Algebra I in 8th grade. The third track, known as Great Explorations in Mathematics (or GEM), compresses the standard 3-year middle school math curriculum into 6th grade, and offers Algebra I in 7th grade and Geometry in 8th grade. GEM is offered in all the District's mainstream middle schools, including those located in lower-income neighborhoods.

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).¹⁶ Entry to GEM also requires a teacher's recommendation and parental consent. Students must earn at least a B grade to stay in the program in 7th and 8th grades; otherwise, they return to the advanced track (as shown in Figure 1). This off-ramp feature addresses concerns that some students who meet the 5th-grade threshold may struggle with the faster pace of GEM. Even those who only spend a year in GEM, however, have an extra pathway to completing Algebra 1 by the end of middle school relative to those who miss the cutoff (see Figure 1).

Appendix Table A1 shows how the pathways in Figure 1 map to course-taking patterns through the end of high school. The table reports the distribution of math course enrollments in each grade from 7th through 12th for four groups of students defined by their 6th-grade math track and, for GEM entrants, by whether they successfully completed the 3-year sequence. Panel A shows that students in the regular track follow the conventional sequence closely: 90% take Algebra 1 in 9th grade, 73% take Geometry in 10th grade, and 45% take Algebra II in 11th grade,

¹⁵ 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 correspond to "accelerated" courses in current CPALMS terminology.

¹⁶ 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.

with fewer than 3% enrolling in any calculus class before the end of high school. Of those who start out in the advanced track (Panel B), just under half take Algebra 1 in 8th grade and Geometry in 9th, but only ~13% reach a calculus class by 12th grade. For the students who enter GEM but leave the program before completing it (Panel C), the modal math course in 9th grade is Geometry, with 63% enrolled (somewhat more than in the advanced track group), followed by Algebra II in 10th (52%) and pre-calculus in 11th (25%). Overall rates of calculus enrollment for this group are roughly double those of the advanced-track group (e.g., 13.7% vs. 7.6% enroll in AP Calculus). Finally, Panel D shows the enrollment patterns of successful GEM completers – nearly all of whom take Algebra II in 9th grade. This group typically goes on to take pre-calculus in 10th grade (65%) and some type of calculus in 11th grade, with 12th-grade enrollments spread across AP Calculus BC, AP Statistics, and dual enrollment math courses. Overall, rates of calculus enrollment are quite high among GEM completers, with 53% taking AP Calculus AB and 30% taking AP Calculus BC by the end of high school.

c. Regression Discontinuity Model

Building on equation (2) we use a “fuzzy RD” approach (Lee and Lemieux, 2010) to analyze the effect of entering GEM in 6th grade on outcomes of marginally eligible students. Specifically, we estimate local linear first-stage models for entering GEM in 6th grade using data for students with math scores in a symmetric bandwidth (BW) around the 380-point threshold, i.e., those with scores in the interval $[380 - BW, 380 + BW]$. Likewise, we estimate local linear reduced-form models for jumps in our outcomes at the 380-point threshold. Our IV estimator of the treatment effect of GEM entry on the mean outcome of the compliers is the ratio of the reduced form and first stage discontinuities at the entry threshold, which we implement using two-stage least squares. Our baseline models have no controls other than cohort fixed effects (4 dummies) to allow for trends over time in GEM participation and in college outcomes, but we show below that combinations of baseline characteristics that predict our main outcome trend smoothly at the 380-point threshold.

For our main analysis, we focus on a bandwidth of 30 points, which is equivalent to $\pm 0.6\sigma$'s (standard deviation units) of the statewide math test. The first-stage and reduced-form relationships are all relatively close to linear, so (as we show below) our estimates of the discontinuities in the first-stage and reduced-form models are relatively stable across bandwidths. We also present estimates using alternative bandwidths, including the one selected

by the algorithm of Calonico et al. (2014; hereafter, CCT): in most cases the CCT choice is close to 30.

III. Data and Analysis Sample

a. Data Sources

We obtained anonymized student-level data from the District’s centralized record system for all children who were in 5th grade between the 2003 and 2007 academic years. These records contain information on gender, age, race/ethnicity, and FRL status. They also include state-wide achievement test scores in reading and math for grades 3-8; course enrollments for grades 6-12, 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), which are available for all students who completed high school in the District.

Starting from this sample of 5th graders, we derive our main analysis sample in a series of steps summarized in Appendix Table A2. We start in column 1 with 106,170 children who were in 5th grade in the District in the academic years 2003-2007. About 43% participated in the FRL program; their racial composition is 34% White non-Hispanic (“White”), 36% Black non-Hispanic (“Black”), and 23% Hispanic. We then eliminate the roughly 20% of 5th grade students who were not also in grades 4 and 6 in the District (column 2). This has little effect on the characteristics of the sample. Next, we eliminate the roughly 30% of students who were present from 4th to 6th grades but left the District before 12th grade (column 3). This leads to a rise in the 4th-grade test scores of the remaining group, and a slight increase in the share of girls (51% vs. 49%). About 17% participated in GEM and 29% participated in the advanced math track math class in 6th grade, leaving about 51% in the regular math track and 3% in remedial math.

A concern for our RD design is that students who enter GEM may be more or less likely to stay in the District and be (potentially) matched to the NSC 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 against 5th-grade math scores. For the population as a whole (Panel A), the estimated discontinuity at the GEM threshold is very small (-0.001) and statistically insignificant ($t \approx -0.1$). The estimates are similarly small for both boys and girls.

Column 1 of Table 1 shows the characteristics and school outcomes of the roughly 56,000 students who meet these sample criteria. Panel A presents demographic and 4th-grade test

scores, as in Appendix Table A2. Panel B shows outcomes in middle and high school. In 8th grade, about a third of these students were enrolled in Algebra I (well above the national average, but about equal to the average for Florida as a whole). In high school, about a quarter took pre-calculus and 15% took calculus (of any kind, including AP). The data in Panel C show that just over 80% of District 5th graders who were present in grades 4-6 and remained in the District for at least six more years graduated from high school “on time” (7 years after 5th grade); just over 60% enrolled in any college in the year after normal high school completion; 41% earned a college degree over the next 6 years; 28% earned a bachelor’s degree in this window; and 16% (or about 57% of all those who earned a bachelor’s degree) earned a degree in STEM or business fields.

For our *RD analysis sample* shown in column 2 of Table 1, we focus on students who were in the sample described in column 1 and had a 5th-grade math score within 30 points of the GEM threshold of 380 points. This subset of nearly 23,000 students has a 30% FRL rate, significantly below the district-wide rate of 42% in column 1. Relative to the overall District population, the analysis sample also over-represents White students and under-represents Black students, reflecting racial differences in test scores. Not surprisingly, students with scores near the GEM cutoff had higher math scores and higher reading scores in 4th grade than the average student in the district-wide sample. About 29% ultimately entered GEM in 6th grade. Most of the others were in the advanced track math class in 6th grade; only 22% were in the regular math stream.

The remaining six columns of Table 1 show the mean characteristics of students in our RD analysis sample separately by gender and tercile of non-cognitive skills. We return to a discussion of these subsamples in the next section.

IV. RD Analysis of GEM Entry

a. Validity of RD Design and First Stage

We begin our empirical analysis by checking the potential validity of an RD design around the GEM cutoff. Panel A of Figure 2 presents a standard check proposed by Lee (2008), which uses the predetermined characteristics of students to predict our main outcome variable (obtaining a bachelor’s degree in STEM or business within 6 years of high school graduation),

then plots the average predicted probabilities by 5th-grade math score.¹⁷ (We use 1-point bins in this and all subsequent plots, reflecting the granularity of the underlying FCAT scores). The figure shows that the predicted probabilities rise steadily with 5th-grade math scores, with no jump at the GEM threshold. We also constructed similar figures by gender and tercile of non-cognitive skill. These are reported in Appendix Figure A2 and are all relatively smooth at the 380-point threshold.

As a second check, we looked for any discontinuity in the density of test scores around the GEM threshold, as suggested by McCrary (2008). 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.

In panel B of Figure 2 we show the relationship between 5th-grade math scores and GEM entry. For reference, we also plot the relationship between test scores and the probability that a student enters GEM *or* advanced math in 6th grade. Starting with GEM (shown in blue), we note three features. 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 up to 6 points below the threshold.¹⁸ Second, among the students with scores near the cutoff, there are both “always takers” (as evidenced by the ~10% GEM entry rate for students with scores just below the 380 cutoff) and “never takers” (as evidenced by the roughly 40% 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 50 ppts. in magnitude – implying that our RD design has a powerful first stage ($t \approx 18.2$ for the pooled sample).

One explanation for the rather high share of never-takers is that the rules for GEM specify that the student must have a math score above the GEM threshold *and* score 300 or more on the 5th-grade reading test. Over 90% of students who barely satisfy the math threshold have reading scores of 300 or more, however, so the latter requirement excludes relatively few students. More significantly, many students who score above both test thresholds are not assigned to GEM, reflecting two types of non-compliance. On one hand, it appears that students

¹⁷ Specifically, we use student demographic characteristics (gender, age, race/ethnicity, FRL status, English Language Learner status), gifted status, and 4th-grade math and reading scores to predict the probability of a STEM or business degree for students with math scores below the GEM threshold. Then we use predictions from this model for all students in our analysis sample.

¹⁸ 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.

with 5th grade math scores just a little above the GEM threshold *and relatively weak 4th grade math scores* are less likely to enter GEM – presumably reflecting teachers’ and/or parents’ concerns over their true math ability. On the other hand, non-compliance rates also vary by school and are typically higher in schools with lower FRL rates and higher shares of students in the upper quintile of 5th-grade math scores. We suspect that some schools may limit GEM enrollment to ensure that class sizes are roughly equal for all three math tracks, effectively raising the threshold for GEM entry.

In contrast to the large jump in GEM enrollment at the 380-point threshold, the green line in panel B of Figure 2 shows there is a much smaller jump – only about 3 ppts. – in the probability that students enter GEM *or* the advanced math track in 6th grade. This implies that most of the GEM compliers ($\approx 94\%$) come from the advanced track.

b. Persistence in GEM

As noted above, eligibility for the second and third years of GEM requires at least a B grade in the previous year. Panels C and D of Figure 2 plot the probabilities of taking Algebra I in 7th **grade** (i.e., remaining on the GEM track for 7th grade), and of taking and successfully completing algebra by 8th **grade** – which is possible for students who remain in GEM at least through 7th grade, as well as for students in the advanced track who perform relatively well in their 7th-grade math class. Looking first at the share who remain in GEM (the blue line), we see a roughly 25 ppt. jump at the threshold for GEM entry, with about 40% of students who scored just above 380 points entering Algebra I in 7th grade. Since the rise in GEM entry at the cutoff is about 60 ppts., we infer that only about two-thirds of marginally eligible GEM entrants successfully completed the first year of the program and entered Algebra I by 7th grade.

Looking next at the share of students who took algebra by 8th **grade** (yellow line in panel C) we see around a 70% rate for students who were marginally eligible for GEM, and a 60% rate for those who were marginally ineligible, implying that eligibility for GEM only raises the share taking Algebra I in middle school by 10 ppts. The green line in panel C shows the share of students who *successfully completed* Algebra I (earning a grade of B or better) by 8th grade. This is lower than the share who enrolled in the course. Interestingly, the ratio of the probability of successful completion to the share enrolled is around 68% for students just to the left or right of the GEM threshold, so the success rate in the course is not much affected by GEM.

Finally, panel D shows the fractions of students who enrolled in or earned at least a B grade in geometry in 8th grade. Both rates are very low to the left of the threshold but rise substantially at the threshold, with around 30% of marginally eligible students taking geometry, and 21% completing the course with a B or better grade. Comparing the latter to the 60-ppt. jump in GEM entry (Panel B), we infer that only about a third of marginally eligible entrants managed to successfully complete all three years of the program. Even among students with math scores near the upper limit of our analysis sample (400-410 points) who have a GEM entry rate of 80%, the success rate in the program is only around 63%.

Why do so many students who entered GEM fail to successfully complete all three years? Part of the answer is that the 5th-grade statewide test is a noisy measure of their math ability. But another important factor identified in the existing literature is non-cognitive skill (see e.g. Jacob, 2002; Heckman et al., 2006; Becker et al. 2010, Card et al., 2026). Good course grades, as required by GEM, require a combination of math skills and the ability to complete class assignments and exams. Even students with high math ability may lack these skills, making it unlikely they can complete GEM.

Unfortunately, our available data do not include direct measures of non-cognitive skills, or even of course grades until 6th grade. Building on Jacob (2002), we therefore use mean grades in high school classes other than math, adjusted for cognitive skills as measured by 4th-grade tests in reading and math, as an index of non-cognitive skills.¹⁹ An obvious concern with this measure is that it is potentially affected by GEM participation. In Appendix C we analyze this possibility by looking at how the mean value of this noncognitive index, and its quantiles, vary with 5th-grade math test scores. Reassuringly, we find that these are all quite flat across the range math test scores near the GEM cutoff.

The importance of noncognitive (NC) skills in the successful completion of GEM is shown in Figure 3. Here, we take students with marginally eligible 5th-grade math test scores (between 380 and 385 points) and divide them into quintiles of NC skills. The left panel shows the fraction of students in each NC quintile who enter GEM in 6th grade: this ranges between 60 and 68% and is nearly constant across the quintiles. The right panel shows the rate of successful completion of GEM, which in contrast is very highly correlated with NC skills. For the bottom

¹⁹ We exclude non-academic courses such as physical education. We have also experimented with excluding science class grades from the index. This has almost no effect on the results.

quintile group, the probability of successful completion of GEM is under 10% while the rate for the top quintile is over 40%.

As shown in Appendix Figure C1, the distribution of measured NC skills is shifted to the right (by about 0.5σ 's) for girls relative to boys. Thus, most of the marginally eligible students in the bottom quintile of NC skills are boys (73%) and most of those in the top quintile are girls (65%). Even within quintile, however, rates of successful completion of GEM are about 6 ppts. higher for girls than boys. We suspect that at least part of this gap reflects gender differences in how well high school grades measure non-cognitive skills in middle school. In light of these differences, and also well-known differences in college major choices between men and women, in the remainder of the paper we examine heterogeneity by gender and NC skill, classifying girls and boys into gender-specific terciles of our simple index of noncognitive skills.

Columns 3-8 of Table 1 show the demographic characteristics and average schooling outcomes for the resulting six subgroups of our main analysis sample. The data in panel A reveal that FRL participants and Black students are less likely to be classified in the higher NC terciles, while Hispanic students are more evenly distributed across the terciles. The 4th-grade math test scores of the 6 subgroups are relatively similar (from a high of $+0.76\sigma$'s above the statewide mean for high-NC boys to a low of $+0.60\sigma$'s above the mean for low-NC girls), while the reading scores are a little more widely dispersed (from a high of $+0.66\sigma$'s for high-NC girls to a low of $+0.44\sigma$'s for low-NC boys).

Consistent with the data in panel A of Figure 3, GEM entry rates are quite similar for the 6 groups, but rates of completing Algebra I by the end of 8th grade or of successfully completing GEM are quite different. About 62% of high-NC girls complete algebra by 8th grade with a grade of B or better; 21% of these girls successfully complete GEM. By comparison, only 17% of low-NC boys complete algebra with a B or better, and only 6% successfully complete GEM.

The differences across the six subgroups persist into high school and beyond. About 40% of high-NC girls and boys complete calculus by the end of high school while only 5% of low-NC boys do so. Likewise, 20% of high-NC boys and 17% of high-NC girls complete AP Calculus AB with a score of 3 or better (enough to qualify for college credit at many colleges), while only 1% of low-NC boys or girls do so. Rates of degree completion and majoring choices also vary widely across NC terciles and between males and females. Within 6 years of high school completion, rates of completing a bachelor's degree range from 73% for high-NC girls to 7% for

low-NC boys; rates of completing a bachelor's degree in STEM or business range from 44% for high-NC boys to 5% for low-NC boys.

c. Impacts of GEM Entry on Math Courses in Middle School

With this background, we turn to an analysis of the impacts of GEM entry on subsequent student outcomes. We begin by summarizing the impacts of GEM entry on the timing and successful completion of Algebra 1 and Geometry, using a fuzzy RD approach. For students as a whole, the first stage and reduced forms are illustrated in Figure 2. Row 1 of Table 2 presents the corresponding first-stage estimate (column 1) and 2SLS estimates of the effects of GEM entry (columns 2-6). Given the first-stage estimate of 0.508, the treatment effect estimates are essentially twice as big as the reduced-form jumps show in panels C and D of Figure 2. For reference, the number in square brackets in each row shows the mean outcome for the untreated compliers. As noted in the discussion of Figure 2C and 2D, these rates range from very close to 0 for geometry to around 65% for enrollment in algebra by 8th grade.

Turning to the effects on our six subgroups, Appendix Figure A4 presents first-stage plots by gender and tercile of NC index. All six 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 six subgroups, there is a much smaller rise (<5 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.

Rows 2-7 of Table 2 show that indeed the estimated first-stage impacts of passing the 380-point threshold are very similar across groups. However, the impacts of GEM entry vary widely, with a *negative* correlation between the treatment effects and the underlying means of the outcomes for the untreated compliers in a particular subgroup. For example, column 4 presents impacts on the probability of earning at least a B grade in Algebra I by 8th grade. The treatment effects are substantially larger for groups with lower means in the absence of GEM. For low-NC girls, with a mean for untreated compliers of 31.5%, GEM has an effect of +20 pts. For high-NC girls, with a mean for untreated compliers of 77.1%, GEM has only a + 5 ppt. effect. The implication is that access to the accelerated GEM curriculum raises the average rate of successful completion of algebra before high school while equalizing differences across

higher and lower NC groups. For the high- and middle-NC tiers it also has a larger effect on boys than girls, equalizing the gender gap in early completion of algebra.

One issue for students entering a more advanced curriculum is that teachers tend to reward relative rather than absolute performance in assigning grades. A student who is near the top of the class in the advanced math track may be closer to the middle in the GEM track, leading to lower grades. Columns 1-3 of Table 3 show the effects of GEM entry on math GPAs in 6th through 8th grades. For the overall sample and all subgroups there are large negative impacts of entering GEM on 6th-grade math GPAs, on the order of 0.7 of a letter grade. There are also negative, but smaller, effects on 7th-grade math GPAs, but close to a zero impact on 8th-grade GPAs, with even slightly positive effects for low-NC students (most of whom are no longer in GEM). The fact that the grade effect is close to zero by 8th grade suggests that there is no lasting negative impact of entering the GEM track for most students.

d. Impacts of GEM Entry on Math Courses in High School

Next, we look at impacts on high school math. Table 4 summarizes the impacts of GEM entry on 8 outcomes: enrollment in Algebra II by 9th grade (which is only possible for students who have already completed Algebra I and Geometry, i.e., those who successfully completed GEM); the total number of advanced math courses that a student has enrolled in by 12th grade; enrollment in pre-calculus; enrollment in any calculus; early enrollment in calculus by 11th grade; enrollment in any AP calculus class; enrollment in AP calculus BC (the second AP class in calculus); and enrollment in AP statistics. Panel A of Figure A5 plots the shares of students taking calculus by 11th grade (blue circles) and by 12th grade (red triangles). There is a clear upward discontinuity in both rates at the GEM entry threshold. Looking in row 1 of Table 4, we see about a +10 ppt. increase in any calculus associated with GEM entry, and a larger 13 ppt. gain for early calculus. Both estimates are highly significant (t 's around 4). Panel C of Figure A5 shows the sensitivity of the estimated impacts to bandwidth choices. Given the near linearity of the reduced form plots, the estimates are quite stable for bandwidths between 10 and 40 points; the CCT-selected bandwidths are 25 and 27, respectively, just slightly below our default bandwidth of 30. The other estimates in row 1 suggest that the average responses to GEM entry include a rise in the number of advanced math classes taken by 12th grade (+0.4 courses), a 7 ppt. impact on the probability of taking pre-calculus (from a baseline rate for the untreated compliers of 48%), and modest rises in the probabilities of taking AP classes in calculus and statistics.

As would be expected given the findings in Table 3, the impacts on high school math outcomes vary widely across our 6 key subgroups. These estimates are shown in rows 2-7 of Table 4; the reduced-form relationships for calculus enrollment by subgroup are shown in Panels A and B of Figure 4. A particularly interesting comparison is between high-NC girls and high-NC boys. The untreated complier means for girls in this category are relatively high, with an average of 2.7 advanced math classes in high school, an 80% rate of taking pre-calculus, a 48% rate of taking any calculus, and a 23% rate of early calculus (i.e., by 11th grade). The corresponding means for the boys are systematically lower – contrary to the fear of many analysts that girls with good math skills tend to select out of advanced math at a higher rate than similar boys (see e.g., Kim et al., 2018). The treatment effects of GEM entry are generally larger for the **boys** (and sometimes much larger, as in the +33-ppt. impact on early calculus versus a 17-ppt. impact for the girls). Thus, for marginally eligible students with high NC skills, entry to GEM narrows – or even reverses – the pre-existing gender gaps in favor of girls.

For students in the mid-tier of NC skills, the gender gaps for the untreated compliers are reversed, as are gaps in the treatment effects. For mid-NC students, the gender gaps among the untreated compliers tend to favor boys – e.g., 20.1% of untreated mid-NC boys enroll in calculus by 12th grade, versus 15.7% of girls. The treatment effects of GEM entry, however, are larger for the girls, in many cases reversing the gender gap to favor girls. For example, among treated complier girls, the share who enroll in calculus by 12th grade is 30.6% (=15.7+14.9), while among treated complier boys the share is 24.1% (=20.1+4.0).

Finally, among students in the lowest tier of NC skills, the untreated compliers of both sexes have relatively low rates of taking advanced math and AP courses: just under 1 class in total for the girls and less than ½ of a class for the boys. Entry to GEM leads to positive treatment effects that are larger for girls than boys – for example +11 versus +8 ppt. effects on the probability of taking pre-calculus, respectively, though rates of completion of calculus and AP math classes remain low for both genders.

Columns 4 and 5 of Table 3 look at the effects of entering GEM on high school GPAs in math and non-math courses. The impacts on math GPAs are slightly negative for students as a whole and for 5 of the 6 subgroups, suggesting that the push of students into more advanced math classes may have lowered their grades slightly, though even the largest effect – for low-NC

girls – is only -0.11 grade points (on a mean of 2.25), and is not significant ($t = 1.2$). The impacts on average grades in non-math subjects are all very close to zero.

Column 6 of Table 3 presents estimates of the effect of GEM entry on PSAT math scores (expressed as percentiles in the national distribution). The average impact of GEM entry is positive (+3.0 percentiles, $t = 3.0$), with particularly large effects on high-NC boys (+6.2 percentiles $t = 3.2$) and low-NC boys (+7.3 percentiles $t = 2.3$), and a marginally significant effect on middle-NC girls (+3.5 percentiles $t = 1.9$). These gains suggest that the exposure to more advanced math classes associated with GEM entry may help students perform better on the PSAT exam. As shown in column 7, we also find a marginally significant effect of GEM entry on the likelihood that students achieve a score of 3 or higher on the AP Calculus AB exam. Interestingly, the effects are largest for the 2 subgroups with the biggest gains in math PSAT scores, supporting the view that GEM entry led to gains in their math achievement.

e. Impacts of GEM Entry on College Outcomes

Next, we turn to college-level outcomes. Table 5 presents the estimated impacts of GEM entry on 6 outcomes: graduation from high school on time (7 years after completing 5th grade); graduation on time and entering college the next year; graduation on time and entering a selective college; earning any degree (associate or bachelor's degree) within 6 years of on-time high school graduation; and earning a bachelor's degree in a STEM field or business within 6 years. Again, for each group of students and each outcome we also show the mean for the untreated compliers: for example, pooling all students, 96.2% graduated high school on time.

Panel B of Appendix Figure 5 plots the reduced form for attaining any degree within 6 years, and attaining a STEM or business degree, while panel D shows the estimated treatment effects of GEM entry on these two outcomes, using the full sample of students and bandwidths of 10 to 40 points. The reduced form for attaining a bachelor's degree shows no evidence of a jump at the GEM entry threshold; the associated treatment effects are close to 0 for the full range of bandwidths. By comparison, the reduced form for attaining a STEM or business degree shows a jump of around 2-3 ppt; the associated treatment effects are close to 5 ppts. for the full range of bandwidths.

Returning to Table 5, row 1 summarizes the impacts of GEM entry on college outcomes for the pooled sample. The estimated impacts on high school graduation and college entry are very close to zero. The average impacts on entering a selective college, earning any degree, or

earning a bachelor's degree are all weakly positive, but not significantly different from zero (t 's in the neighborhood of 1). Consistent with the estimates in Appendix Figure 5, however, the impact on completing a STEM or business degree is positive and significant at conventional levels (estimate = 5.5 ppts.; $t \approx 2.4$), implying about a 25% increase in STEM/business degrees relative to the untreated compliers. To put this estimate in context, we found in Table 4 that entry to GEM led to a 9.5 ppt. increase in the probability of completing calculus by end of high school. So, the net impact on STEM or business is about one-half as big as the effect on any calculus.

Table 6 presents a more detailed analysis of degree outcomes, distinguishing between 3 main categories of STEM: math, engineering, physical sciences and computer science (column 2); life sciences (column 3); and social sciences (column 5). We also show impacts on degrees in business (column 4), health (column 6), and other non-STEM fields (column 7). Using the estimated treatment effects of GEM entry from row 1, we infer that the 5.5 ppt increase in STEM/ business degrees came from a combination of a 2.3 ppt. rise in bachelor's degree completion, a 1.2 ppt. reduction in health-related degrees, and a 1.9 ppt reduction in other non-STEM degrees. Within the STEM/business fields, there were gains in all 4 sub-categories shown in the table, with 36% ($=2.0/5.5$) coming from a rise in social science degrees; 31% ($=1.7/5.5$) coming from a rise in life science degrees, 22% ($=1.2/5.5$) coming from a rise in business degrees, and 11% ($=0.6/5.5$) coming from engineering, physical sciences and math.

As we have found at the middle school and high school levels, GEM entry has very different effects by gender and NC tercile. Panels C and D of Figure 4 plot the reduced forms for completing any bachelor's degree and completing a STEM/business degree for the 6 subgroups. The associated treatment effects, summarized in rows 2-7 of Table 5 show that GEM has very small effects on high-NC girls, arguably because even in the absence of GEM these girls have high rates of entering a selective college (82.2%), earning a bachelor's degree in 6 years (81.0%) and earning a STEM or business degree (45.3%).

There are also no significant effects of GEM entry on the probability of entering a selective college or earning a bachelor's degree for high-NC boys, but there is a +11.9 ppt. effect on completion of STEM/business degrees (visually evident in panel D of Figure 4) which pushes their rate from 42.8% in the absence of GEM (a little below the rate for high-NC girls) to 54.7% (a little above the rate for high-NC girls).

In contrast to the null effects for high-NC girls, GEM had larger effects on the college outcomes of mid-NC girls, and its largest effects on low-NC girls. Our estimates imply that GEM raised the share of low-NC girls entering selective colleges by 14.1 ppts. (a 60% increase relative to the untreated compliers); increased their rate of completion of a bachelor's degree by 13.1 ppts. (a 70% increase relative to the untreated compliers); and increased their rate of completion of degrees in STEM or business by 13.6 ppts. (a 270% increase relative to the untreated compliers). By comparison, GEM entry had no significant effects for mid- and low-NC boys on any of the outcomes in Table 5.

It is also interesting to look at the subgroup-specific impacts of GEM on the detailed categories of STEM (Table 6). For high-NC girls we see only a small, insignificant +1.8 ppt. effect on STEM/business as a whole, but within STEM we see a moderately large reduction in the share completing a degree in Social Sciences (-6.5 ppts., $t \approx 1.3$) offset by a rise in "pure" STEM fields (columns 2 and 3) that totals 7.1 ppts. ($t \approx 1.8$) – around 40% over the base rate. For high-NC boys the significant 11.9 ppt. rise in STEM/business degrees arises from a small increase in any degree completion and a nearly 9 ppt. diversion from health and other non-STEM fields. Within the STEM fields, the largest gain is for social science degrees, which rise by 8 ppts. ($t \approx 2.5$). High-NC boys also see an increase in pure STEM degrees that is similar in magnitude to the effect for high-NC girls (+7.2 ppts., $t \approx 1.6$).

For mid-NC girls, the GEM-associated rise in STEM/business majors is attributable about equally to a 6-ppt. rise in any degree completion and a 5.3-ppt. diversion from health and other non-STEM fields (though neither point estimate is significantly different from zero). For these girls there is a marginally significant rise in business degrees (+6.7 ppts., $t \approx 2.0$) and an imprecisely estimated rise in social sciences (+5.4 ppts., $t \approx 1.2$), but no change in life or physical sciences. For low-NC girls, GEM has no negative effect on completion of non-STEM degrees but instead leads to a net shift from non-degree completion to a mix of life sciences, social sciences, and business.

Finally, for low-NC boys there are no significant effects on any of the components of STEM, and indeed the estimated impacts of STEM on degree completion and STEM/business degree completion are both negative. Looking at the reduced-forms in panels C and D of Figure 4, and taking account of the fact that most of these boys spent only a year in GEM, we believe the negative point estimates are unlikely to reflect true causal effects of GEM entry. We return to

this issue below when we study differences in college outcomes between students in the District and those in District B, which has no GEM program.

f. Summarizing Heterogeneity by Non-Cognitive Skill and Gender

The results in Tables 2, 3, and 4 suggest that there is wide heterogeneity in the impacts of GEM entry by NC tercile and gender. We summarize this heterogeneity in Figure 5, where we plot the first stages for entering GEM and the means of 5 main outcomes for the treated and untreated compliers in each of the 6 subgroups against their NC skill index. The vertical distance between the means for the treated and untreated compliers is the treatment effect of GEM entry for that subgroup. To aid in the interpretation of the gaps, we show a confidence interval around the mean for the treated compliers with the width of the 95% confidence interval for the GEM treatment effect. Thus, if the interval excludes the mean for untreated compliers, the treatment effect of GEM entry is significantly different from zero at conventional levels.

Panel A of Figure 5 plots the probabilities of entering GEM, which as noted earlier are quite similar for the six subgroups. Panel B shows the probabilities of successfully completing GEM (i.e., completing geometry by 8th grade with at least a B grade). The means for the untreated compliers are essentially zero, while the means for the treated compliers range from just a few percent for low-NC boys to just over 60% for high-NC girls. The relationship between our NC skill index and the probability of completing GEM for marginally qualified entrants is nearly linear, confirming the importance of non-cognitive skills for GEM persistence.

Panel C presents the means for completing calculus by end of high school. For this outcome the means of the untreated compliers rise systematically with NC skills, though middle-NC girls and high-NC boys under-perform relative to the trend for the other subgroups. Interestingly, these two groups have relatively large positive treatment effects, particularly high-NC boys, for whom the mean of the treated compliers is close to 60%. In contrast, high-NC girls, who have the highest calculus enrollment rate in the absence of GEM, have only small boost from GEM. These patterns are similar for calculus by the end of 11th grade (shown in panel D), though GEM also has a positive effect on the share of high-NC girls who complete calculus early.

Panel E shows the mean rates of completing a bachelor's degree within 6 years of high school graduation. Degree completion is very strongly correlated with levels of NC skills, ranging from only a few percents for low-NC boys to just over 80% for high-NC girls. GEM

entry does not have a large effect on this outcome, apart from the marginally significant gain for low-NC girls. Finally, panel F shows rates of completion of a degree in STEM or business. For the untreated compliers there is again a very strong positive correlation with NC skills, ranging from close to zero for low-NC boys to 45% for high-NC girls. Entry to GEM boosts the rates of STEM/business majoring by low- and mid-NC girls and high-NC boys, with a pattern of treatment effects that closely parallels the effects on completing calculus in panel C. Indeed, the correlation across subgroups in the treatment effects is 0.70, and a simple regression implies that each 1 ppt. increase in calculus enrollment is associated with a 0.7 ppt rise in the probability of majoring in STEM or business.²⁰

The cross-group relationship between the GEM treatment effect on STEM/business majoring and the effect on enrolling in calculus by *12th grade* is stronger than the relationship with the treatment effect on enrolling in calculus by *11th grade* ($\rho=0.39$), suggesting that it is *any calculus* in high school, rather than *early calculus*, that mediates entry to math intensive majors. Interestingly, GEM's impacts on early calculus are more highly correlated with the treatment effects on obtaining a degree in the pure STEM fields ($\rho=0.59$). Notably, both high-NC girls and high-NC boys have large effects on early calculus and ≈ 7 ppt. rises in majoring in pure STEM. The connection is consistent with the fact that some of these fields enroll students directly out of high school and look for evidence of strong math skills in their admission decisions.

Overall, the patterns in Figure 5 point to two main conclusions. First, among students with 5th-grade math scores near the 80th percentile (the GEM cutoff), non-cognitive skills are highly correlated with persistence in GEM, calculus enrollment in high school, and the probability of graduating with a bachelor's degree in STEM/business. Second, the treatment effects of GEM entry tend to be higher for students in the middle range of NC skills, with negligible impacts on the group with the lowest NC skills (who typically leave GEM after the first year) and relatively small effects on the group with the highest NC skills (high-NC girls).

V. Spillover Effects on Untreated Compliers?

Our RD analysis relies on comparisons between students whose 5th-grade math scores are just above or just below the 380-point GEM threshold. As noted in Section I, the presence of

²⁰ Interestingly, there is no such relationship across groups between the treatment effects of GEM entry on probability of successfully completing Algebra I by 8th grade (i.e., earning a grade of B or better) and the probability of taking calculus by end of high school. The correlation across the 6 groups in these effects is -0.10 .

GEM could affect students just below the cutoff through peer effects, discouragement effects, or top-of-the-class effects, leading our RD estimators to overstate or understate the true effect of GEM on the treated compliers. In this section we present a between-district comparison that can potentially detect such spillover effects.

At the outset, it is worth noting that the reduced-form plots in Figure 4 do not suggest large spillovers on the untreated compliers. The scatter of points just to the left of the GEM threshold shows no evidence of concavity for any of the 6 group, as might be expected if students who narrowly missed GEM were discouraged in pursuing math, or convexity, as might be expected if they benefitted from the removal of their higher-scoring classmates. A particularly interesting subgroup is high-NC girls, for whom there is no discontinuity in enrolling in calculus by 12th grade, or in attaining a STEM or business degree. Given the findings of Buser and Yuan (2019) and Ellison and Swanson (2023) on math competitions, one might suspect that these girls would be particularly sensitive to discouragement effects. Similarly, significant peer effects would be expected to lead to lower achievement by high-NC girls who are ineligible for GEM.²¹ Yet, there appears to be no long-term disadvantage of narrowly missing the GEM threshold for high-NC girls.

Nevertheless, to pursue a deeper investigation, we use data from another large urban school district in Florida where we can follow students from 4th grade to the end of college. This district – which we refer to as “District B” – has only two main tracks in middle school math, with no GEM program or third track for very high scoring students. However, District B’s criteria for placing students in *advanced* math are *less* selective than the District’s. Among students in our sample cohorts whose math scores were within 30 points of the GEM threshold, District B assigned nearly all (96%) to the advanced track, while the District assigned only 80% to the advanced and GEM tracks combined. Similarly, District B pushed students into early algebra at much higher rates than the District.²² This difference is illustrated by the patterns in

²¹ While some studies find evidence of classroom peer effects, we note that two influential studies of elite public high schools, by Abdulkadiroglu et al. (2014) and Dobbie and Fryer (2014), find no discontinuity in outcomes between marginally eligible and ineligible students, despite large differences in peer test scores at the admission threshold. Likewise, in Card and Giuliano (2016) we found no evidence of peer effect arising from tracking in 4th and 5th grades.

²² Appendix Table D1 presents some comparisons of student characteristics in the two districts. Here, and in the rest of this section, we combine data from our main District analysis sample with data from District B that meet the same sample criteria. Data from District B are only available for cohorts in 5th grade after 2004: thus, we also restrict our District sample by dropping the first cohort of students.

the top row of Figure 6, where we show the fractions of students in each of the six subgroups who enroll in Algebra I by 8th grade in the District and in District B, conditional on 5th-grade math scores. To help interpret the graphs, we fit local linear models for the two districts using data ± 30 points from the GEM cutoff, allowing for intercept and slope shifts at the GEM cutoff in the District and for slope shifts in District B.²³

On average, over 85% of students in District B with scores near the GEM threshold enroll in Algebra in middle school – versus around 60% of students in the District with scores just below the cutoff, and 70% of students in the District with scores just above (see panel C of Figure 2). Across the six subgroups, the *levels* of Algebra enrollment in the two districts vary widely, but the pattern of differences between the two districts is similar, with higher enrollment in District B for students on either side of the cutoff and no discontinuity, versus a ≈ 10 ppt. jump in the District associated with entry to GEM (for all groups except high-NC girls). Among high-NC boys, for example, nearly 100% of those in District B with scores near the cutoff enroll in algebra by 8th grade, versus 70% of those in the District with scores below the cutoff and nearly 80% for those with scores above the cutoff.

The differences between the two districts are summarized in Figure 7. Here, we show estimated between-district gaps in different outcomes from models of the form:

$$Y = a_0 + a_1 T^m + a_2 T^m [T^m \geq 0] + (b_0 + b_1 T^m + b_2 [T^m \geq 0] + b_3 T^m [T^m \geq 0])D + \eta \quad (3)$$

where T^m is a student's 5th-grade math score (recentered at the GEM cutoff) and D is an indicator for students in the District (versus District B). The coefficients b_0, b_1, b_2, b_3 parameterize the gap in mean outcomes between the two districts at each test score T^m as a linear function of the test score, with a potential jump and change in slope at the GEM cutoff. We fit this model using a sample that combines our main analysis sample from the District and a parallel sample from District B, and we use the gender-specific NC terciles from the District to classify District B students into six parallel gender-NC subgroups.

To adjust for differences in observed covariates between students with similar 5th-grade math scores in the two districts, we reweight District B observations by $p/(1 - p)$, where p is a propensity score for being in the District, estimated separately for each gender-NC subgroup

²³ In models that allow for intercept shifts in District B, the shift is never significantly different from zero.

within 5-point bins of the 5th-grade math score.²⁴ Because all subgroups show little-to-no overlap for observations with $p > 0.9$ (see Appendix Figure D1), we also trim each sample to exclude observations in this range.²⁵ To assess how well this approach balances the covariates most relevant to our main outcome, we use the baseline characteristics used in the propensity score model to construct a predicted probability that students complete a bachelor's degree in a STEM or business field.²⁶ In Appendix Figure D2, we plot these predicted probabilities against 5th-grade math scores for students in the two districts after reweighting and trimming the data. The means are very similar in the two districts across all values of 5th-grade math scores, although we see slightly higher values among high-scoring students in District B for girls in the middle NC tercile.

While the evidence in Appendix Figure D2 is reassuring, it is important to keep in mind that the projection of student outcomes on a locally linear function of test scores would not necessarily have the same intercept or slope in the two districts, even if the math curriculums were the same and the observed student characteristics were perfectly balanced. Evidence on the intergenerational transmission of occupations (Beller and Hout, 2006; Greenberg et al., 2024) and college majors (Dahl et al., 2020) suggests that the probabilities of completing different types of college degrees depend on parental education, income, and occupation. We have only crude controls for family background based on race/ethnicity and FRL eligibility. If the distributions of parental education and occupations differ between districts conditional on these controls, we might expect differences in the relationship between the probability of completing a STEM degree and 5th grade math scores. Similarly, degree completion and major choices also depend on the programs offered by local colleges and universities (Acton et al., 2026), and jobs available at local employers (Weinstein, 2022). If these differ between the districts it could also lead to differences in the projection of our main outcomes on 5th-grade scores. As we show in Appendix Table D2, however, the average characteristics of adults and parents in the two

²⁴ For each sample, the propensity score is estimated using a logit model that includes quadratics in 4th-grade reading and math scores, student age, FRL status, English language learner status, gifted status, and cohort fixed effects.

²⁵ Further analysis reveals that the excluded observations (with $p > 0.9$) are primarily Black students who are English Language learners, which is consistent with relatively large differences between districts in the population shares of Blacks and immigrants (see Appendix Table D2).

²⁶ This index is analogous to the one plotted in Figure 2A (for the pooled District sample) and Appendix Figure A2 (for the District subsamples) except that here, we use only the four cohorts that are available in both districts. As before, the models are fit to observations in the District RD sample with scores below the GEM threshold, and the coefficient estimates are used to predict the outcome for the full sample of students in the District and District B.

districts are remarkably similar, apart from higher shares of immigrants and Black families in the District.²⁷ In particular, the distributions of education, household income, and industry of work, and the share of adults with degrees in STEM or business are very similar between the districts. Moreover, the two districts have quite similar local colleges and universities.

The top row of Figure 7 shows fitted values for the covariate-adjusted, between-district gap in enrollment in Algebra I by 8th grade. We note the large gaps for students with scores close the threshold, the discontinuities in the gap for all groups except high-NC girls, and the fact that middle school algebra enrollment rates in the District only catch up with those in District B for students with scores near the upper end of the range of our analysis sample. While District B's middle school curriculum channels most higher-scoring students into Algebra I in 8th grade, it does not include a pathway to 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 B enroll in Geometry by 8th grade. In contrast, in the District, this fraction is around 35% (see panel D of Figure 2).

Given the differences in the middle school math sequence between the District and District B, it is interesting to ask how the longer run outcomes of students with different 5th-grade math scores compare in the two districts. In particular, we focus on the question of whether students in the District with scores just below the GEM cutoff do relatively worse in the District than in the District B – as might happen if there are significant negative peer effects or discouragement effects caused by moving the higher-scoring students to the GEM track.

The lower panels in Figure 6 show the fractions of students in the two districts who enroll in calculus by 12th grade and by 11th grade, and the fractions who complete a STEM or business major 6 years after high school graduation; the lower panels in Figure 7 show the associated between-district gaps in these outcomes at each value of T^m . Looking first at calculus by the end of high school, we note that there is indeed evidence that students with scores below the GEM cutoff tend to have lower rates of calculus enrollment in the District relative to their counterparts in District B. The rates are most similar for the low-NC groups, and the gaps are largest for the high-NC groups – especially high-NC boys. For three groups (low-NC boys and girls and high-NC boys) there is also some evidence that the gap is more negative just below the cutoff. On the

²⁷ Consistent with this difference, and as noted above, Black children who are English language learners make up a large share of those excluded when we trim observations with $p > 0.9$ from our sample.

whole, however, we view the evidence for calculus enrollment as inconsistent with spillover effects on students who just miss GEM eligibility. First, instead of displaying a dip in the District relative to District B just below the GEM cutoff (as might be expected if there were negative spillovers), the enrollment gaps either trend linearly as T^m approaches the threshold from below or are flat over the range of scores below the cutoff.²⁸ Second, in light of the wide gaps in algebra enrollments seen in the first row of Figures 6 and 7, the gaps in calculus enrollment appear linked to District B's policy of broad access to early algebra rather than any spillover effects within the District. Similarly, a likely explanation for the widening of the gap below the cutoff (e.g., for high-NC boys) is that early algebra is more likely to lead to calculus in high school for students with higher math scores.

The third panel in Figures 6 and 7 compares rates of enrollment in calculus by 11th grade for students in the two districts. Here, the patterns are more similar across our six groups. For low- and middle-NC students, rates of early calculus in District B are very low. Even for high-NC students, early enrollment in calculus is relatively rare in District B, with rates below 5% for those with math scores near the GEM threshold. By comparison, students in the District have higher rates of early calculus, with jumps at the GEM threshold for all groups but low-NC boys. The between-district gaps in Figure 7 are all positive and increasing with math test scores before and after the GEM threshold. These patterns do not show any evidence of spillover effects on students who narrowly miss qualifying for GEM.

Finally, the bottom panels in Figures 6 and 7 show rates of completion of a STEM or business degree. Looking first at the unadjusted means in Figure 6, rates of STEM/business degree completion look remarkably similar in the two districts for students below the cutoff. The covariate-adjusted gaps in Figure 7 reveal some evidence of negative trends in the District relative to District B in the range of scores below the cutoff. Interestingly, these trends largely mirror the patterns in calculus enrollment discussed above. However, Figure 7 also confirms that among observably similar students in the two districts with scores in the 30-point range below the cutoff, differences in STEM/business degree earning are mostly small and insignificant – even for high-NC boys for which the calculus enrollment gap is relatively large.

²⁸ Arguably, there is some evidence of a dip in the figure for low-NC girls, where the mean gap is negative only in the 5-point bin below the cutoff; however, there is also evidence of a downward trend over the full range of scores below the threshold.

Our main conclusion from these between-district comparisons is that there do not appear to be large spillover effects on students in the District who were marginally ineligible for GEM. The outcomes for high-NC girls – a group that might have been expected to exhibit discouragement effects, given the literature on math competitions – trend smoothly through the GEM threshold. And the patterns of early calculus enrollment suggest that outcomes for marginally ineligible students are in line with patterns for the broader group of students with scores below the cutoff for GEM. We do see lower rates of enrollment in calculus by the end of high school for students in the District with scores below the GEM threshold relative to comparable students in District B. However, we believe that these gaps are attributable to the broader-based acceleration policy of District B, rather than to spillover effects within the District.

Another feature of the between-district comparisons is the pattern of outcomes for students with math scores above the GEM threshold. Looking at the bottom rows of Figures 6 and 7, we note that in three cases (low-NC girls, mid-NC girls, and high-NC boys), there is clear evidence of a jump at the GEM cutoff in the District relative to District B, reflecting the treatment effects noted in the discussion of Table 5.²⁹ In all three cases, the treatment effect results in a positive gap between the districts in the rate of STEM/business majoring for students with scores above the GEM threshold.

The pattern of between-district gaps at scores above the threshold is particularly interesting for high-NC girls. While there is no discontinuity in STEM/business majoring for high-NC girls at the cutoff, below the cutoff the rates for this group are high relative to other groups in the District. In Figure 7, we see that there is also a positive between-district gap in STEM/business majoring for high-NC girls that, while insignificant below the cutoff, increases monotonically with the math score, leading to a relatively large divergence further above the threshold. Interestingly, these patterns of STEM/business degree completion mirror some distinctive patterns in math enrollments for this group. In particular, a combination of relatively high rates below the threshold (compared to other groups in the District) and a lack of threshold-

²⁹ We also note that there is no negative discontinuity in STEM/business majoring for low-NC boys in the District, contrary to the small negative discontinuity seen in row 7 of Table 5 and in panel D of Figure 4. This difference arises from the exclusion of the 2004 cohort from the samples used in our cross-district comparisons.

crossing effects is also seen in early algebra and calculus by 12th grade, while a relatively steep positive trend above the threshold is also seen in early calculus.

Together, these patterns suggest that girls in the District with high noncognitive skills and math scores near the cutoff are neither helped nor hurt by their access to GEM: they enroll in calculus and earn degrees in STEM/business at similar rates on both sides of the GEM cutoff, and their rates of STEM/business degree earning are also similar to comparable girls in District B. At higher levels math ability, however, the STEM/business degree rates of these girls surpass those of their District B counterparts. The between-district gap in early calculus is also especially large, suggesting that the GEM policy supports math-oriented pursuits among girls with high non-cognitive skills and high math ability.³⁰

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 6th- and 7th-grade curriculum into a single year, leading to Algebra I in 7th grade and Geometry in 8th grade. Admission is limited to students who score above the 80th percentile on the statewide 5th-grade math test, and it has “off-ramps” after 6th and 7th grade that allow students to return to the conventional advanced math track if they are struggling. In fact, only a third of marginal GEM entrants actually complete all three years, but 70% complete algebra by the end of 8th 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 completion and majoring in

³⁰ Under the assumptions posed by Ellison and Pathak (2025) about the upper and lower tracks in a model with a single index of skills, these patterns are consistent with the District’s math placement policies being “optimal” for high-NC girls.

more math-intensive fields (STEM and business). We have three main findings. First, among students with 5th-grade math scores in a relatively narrow range around the GEM threshold, persistence in the GEM program, math course outcomes in high school, college degree attainment, and STEM/business majoring are all highly correlated with a simple index of non-cognitive skills based on (adjusted) course grades in high school classes other than math. Controlling for this index, they are also higher for girls than boys, leading us to investigate the variation in the treatment effects of GEM entry by gender and NC skill tercile. Second, entry to GEM leads to a rise in the average number of advanced math classes in high school, with a relatively large effect on the probability of completing calculus by 11th grade and a slightly smaller effect on the probability of completing calculus by 12th grade. These average impacts are concentrated among students in the middle-ranges of NC skills; we find little effect on any calculus-related outcome for the lowest NC boys, while for the highest NC girls, GEM affects only the timing and level of calculus enrollment among marginal entrants. Third, entry to GEM has no effect on the average rate of attaining a bachelor's degree within 6 years of high school graduation, but leads to a 5-6 ppt. increase in the share of students who major in STEM or business. Across groups, the gains in STEM/business majoring are highly correlated with the gains in calculus enrollment by the end of high school, suggesting a causal pathway from advanced math completion in high school to math intensive major choices in college.

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 suggested math preparation for more selective engineering and science programs. To address this concern, we use data from a second large school district in the same state (District B), which has regular and advanced tracks in middle school math but not a third upper track like GEM for the highest performers in math. In comparison to the District, the policies in District B lead to a much higher share of students with 5th grade scores around the GEM threshold who enroll in algebra in 8th grade. We also find interesting differences between the districts in the relation between 5th-grade scores and completion of calculus in high school, particularly for high NC boys, who may benefit from the more inclusive acceleration policy in District B. Taking the evidence as a whole, however, we conclude that GEM has no large spillover effects on marginally ineligible students, supporting the internal validity of our RD design.

Overall, we conclude that a targeted accelerated math curriculum like GEM has the potential to help raise participation in advanced math classes in high school and increase the share of students who major in more math-intensive fields in college. Based on the evidence from GEM, we suspect the gains from *any* accelerated curriculum are likely to vary substantially with the non-cognitive skills of students. That opens up the question of whether targeting should be based solely on test scores or should take account of course grades that also reflect non-cognitive skills. A related question is whether it would be useful to add a component to the advanced curriculum that addresses non-cognitive skill formation. A growing body of research suggests that non-cognitive skills are increasingly important in the labor market and are complementary with cognitive skills, particularly for high-cognitive-ability workers (e.g., Deming, 2017; Edin et al., 2022).

References

- 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.
- Acton, Riley, Kalena E. Cortes, Louis Miller and Camila Morales. 2026. “Distance to Degrees: How College Proximity Shapes Students’ Enrollment Choices and Attainment Across Race-Ethnicity and Socioeconomic Status.” *Economics of Education Review*, forthcoming.
- Altonji, Joseph G. 1995. “The Effects of High School Curriculum on Education and Labor Market Outcomes.” *Journal of Human Resources* 30(3): 409-438.
- Altonji, Joseph G., Peter Arcidiacono 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.
- Antonovics, Kate, Sandra E. Black, Julie Berry Cullen and Akiva Yonah Meiselman. 2023. “Patterns, Determinants, and Consequences of Ability Tracking: Evidence from Texas Public Schools.” NBER Working Paper 30370. August 2023.
- Becker, Gary S., William H. J. Hubbard, and Kevin Murphy. 2010. “Explaining the Worldwide Boom in Higher Education of Women.” *Journal of Human Capital* 4(3): 203-241.
- Beller, Emily and Michael Hout. 2006. “Intergenerational Social Mobility: the United States in Comparative Perspective.” *Future Child* 16(2): 19-36.
- Betts, Julian R. 2011. “The Economics of Tracking in Education.” In Hanushek, Eric A., Stephen Machin and Ludger Woessmann (eds.) *Handbook of the Economics of Education*, Volume 3, Amsterdam: North Holland, pp. 341-381.
- Borghans, Lex, Bart H. H. Golsteyn, James J. Heckman, and John Eric Humphries. 2016. “What Grades and Achievement Tests Measure.” *Proceedings of the National Academy of Science* 113(47): 13354-13359.
- Buser, Thomas and Huaiping Yuan. 2019. “Do Women Give Up Competing More Easily? Evidence from the Lab and the Dutch Math Olympiad.” *American Economic Journal: Applied Economics* 11(3): 225–252.
- Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. “Robust nonparametric confidence intervals for regression-discontinuity designs.” *Econometrica* 82 (6): 2295–2326.
- Campos, Christopher, Pablo Muñoz, Alonso Bucarey, and Dante Contreras. 2026. “College Major Choice, Payoffs, and Gender Gaps,” NBER Working Paper 4736, January 2026.
- 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. 2026. “Can Gifted Education Help Higher-ability Boys from Disadvantaged Backgrounds?” *Journal of Political Economy*, forthcoming.

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

Cimpian, Joseph R. and Jo R. King. 2024. “An Institution-level Analysis of Gender Gaps in STEM Over Time.” *Science* 386(6724): 853-856.

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.

Dahl, Gordon B., Dan-Olof Rooth, and Anders Stenberg. 2020. “Family Spillovers in Field of Study.” NBER Working Paper 27618, July 2020.

Deming, David. 2017. “The Growing Importance of Social Skills.” *Quarterly Journal of Economics*, 132(4): 1593–1640.

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

Dobbie, Will and Roland G. Fryer Jr. 2014. “The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools.” *American Economic Journal: Applied Economics* 6(3): 58–75.

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.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. “Peer Effects, Teacher Incentives and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya.” *American Economic Review* 101(3): 1739-1774.

Edin, Per-Anders, Peter Fredrickson, Martin Nybom, and Bjorn Ockert. 2022. “The Rising Return to Noncognitive Skill.” *American Economic Journal: Applied Economics* 14(2): 78-100.

Ellison, Glenn, and Ashley Swanson. 2023. “Dynamics of the Gender Gap in High Math Achievement.” *Journal of Human Resources* 58(5): 1679-1711.

Ellison, Glenn and Parag Pathak. 2025. “Optimal School System and Curriculum Design: Theory and Evidence.” NBER Working Paper 34091, August 2025.

Greenberg, Kyle, Matthew Gudgeon, Adam Isen, Corbin L. Miller, and Richard W. Patterson. 2024. “Intergenerational Transmission of Occupation: Lessons from the United States Army.” NBER Working Paper 33009, November 2024.

Heckman, James J., Jora Stixrud, and Sergio Urzúa. 2006. “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior.” *Journal of Labor Economics* 24(3): 411-482.

Horvath, Hedvig. 2025. “Classroom Assignment Practices and Implications for Teacher Value-Added Estimation.” School of Education, University College London Unpublished Working Paper.

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

Huffaker, Elizabeth, Sarah Novicoff and Thomas S. Dee. 2025. “Ahead of the game? Course-taking patterns under a math pathways reform.” *Educational Researcher* 54(2): 91-102.

Humphries, John Eric, Juanna Schreter Joesen and Gregory F. Veramedi. 2025. “Complementarities in High School and College Investments.” Cowles Foundation Discussion Paper 2446, June 2025.

Jacob, Brian A. 2002. “Where the Boys Aren’t: Non-cognitive Skills, Returns to School, and the Gender Gap in Higher Education.” *Economics of Education Review* 21: 589-598.

Kim, Ann Y., Gale M. Sinatra and Viviane Seyranian. 2018. “Developing a STEM Identity Among Young Women: A Social Identity Perspective.” *Review of Educational Research* 88(4): 589-625.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad. 2016. “Field of Study, Earnings, and Self Selection.” *Quarterly Journal of Economics* 131(3): 1057–1111.

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

Lee, David S. and Thomas Lemieux. 2010. “Regression Discontinuity Designs in Economics.” *Journal of Economic Literature* 48(2): 281-355.

Lefgren, Lars. 2004. “Educational Peer Effects and the Chicago Public Schools.” *Journal of Urban Economics* 56: 169-191.

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 Science Board, National Science Foundation. 2023. *Higher Education in Science and Engineering*. Science and Engineering Indicators 2024. NSB-2023-32. Alexandria, VA. <https://nces.nsf.gov/pubs/nsb202332/>

Oakes, Jeannie. 1985. *Keeping Track: How Schools Structure Inequality*. New Haven, CT: Yale University Press.

Owen, Stefanie. 2023. “College Major Choice and Beliefs About Relative Performance: An Experimental Intervention to Understand Gender Gaps in STEM.” *Economics of Education Review*. 97: 102479.

Özek, Umut. 2021. “The Effects of Middle School Remediation on Postsecondary Success: Regression Discontinuity Evidence from Florida.” *Journal of Public Economics* 203: 104518.

Paloyo, Alfredo R. 2020. “Peer Effects in Education: Recent Empirical Evidence.” In Dominic Brewer and Patrick J. McEwan, editors, *Economics of Education* (2nd edition). New York: Elsevier.

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

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

UC Berkeley College of Engineering. 2024. “Undergraduate Admissions.” Retrieved from <https://engineering.berkeley.edu/admissions/undergraduate-admissions/> May 2024.

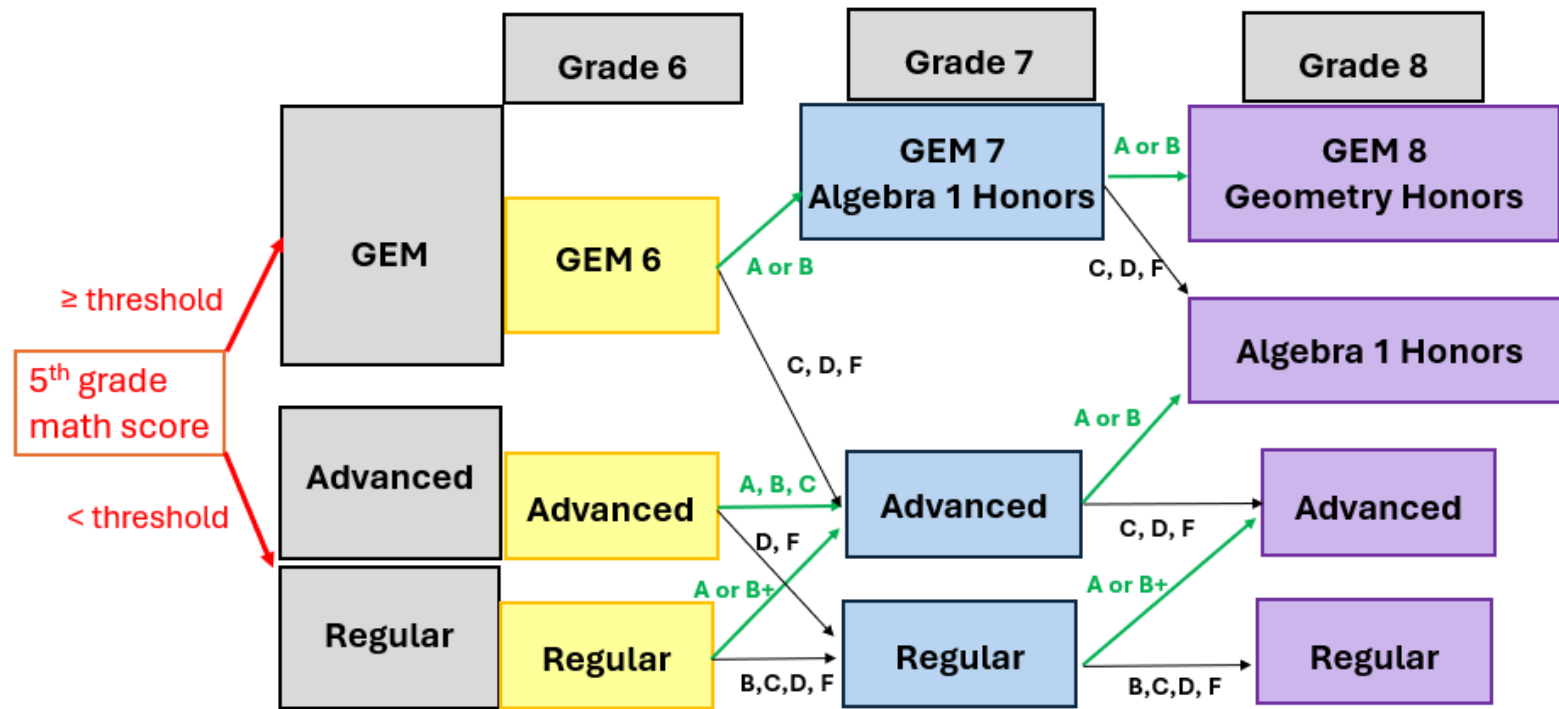
United States Department of Education National Center for Education Statistics. 2022. *Digest of Education Statistics*. Available at <https://nces.ed.gov/programs/digest/>. Accessed September 9, 2024.

Weinstein, Russell. 2022. “Local Labor Markets and Human Capital Investments.” *Journal of Human Resources* 57(5): 1498-1525.

Wolfe, Rebecca L., Elizabeth D. Steiner, and Jonathan Schweig. 2023. “Getting Students to (and Through) Advanced Math: Where Course Offerings and Content Are Not Adding Up.” Rand Research Report RR-A827-10, 2023.

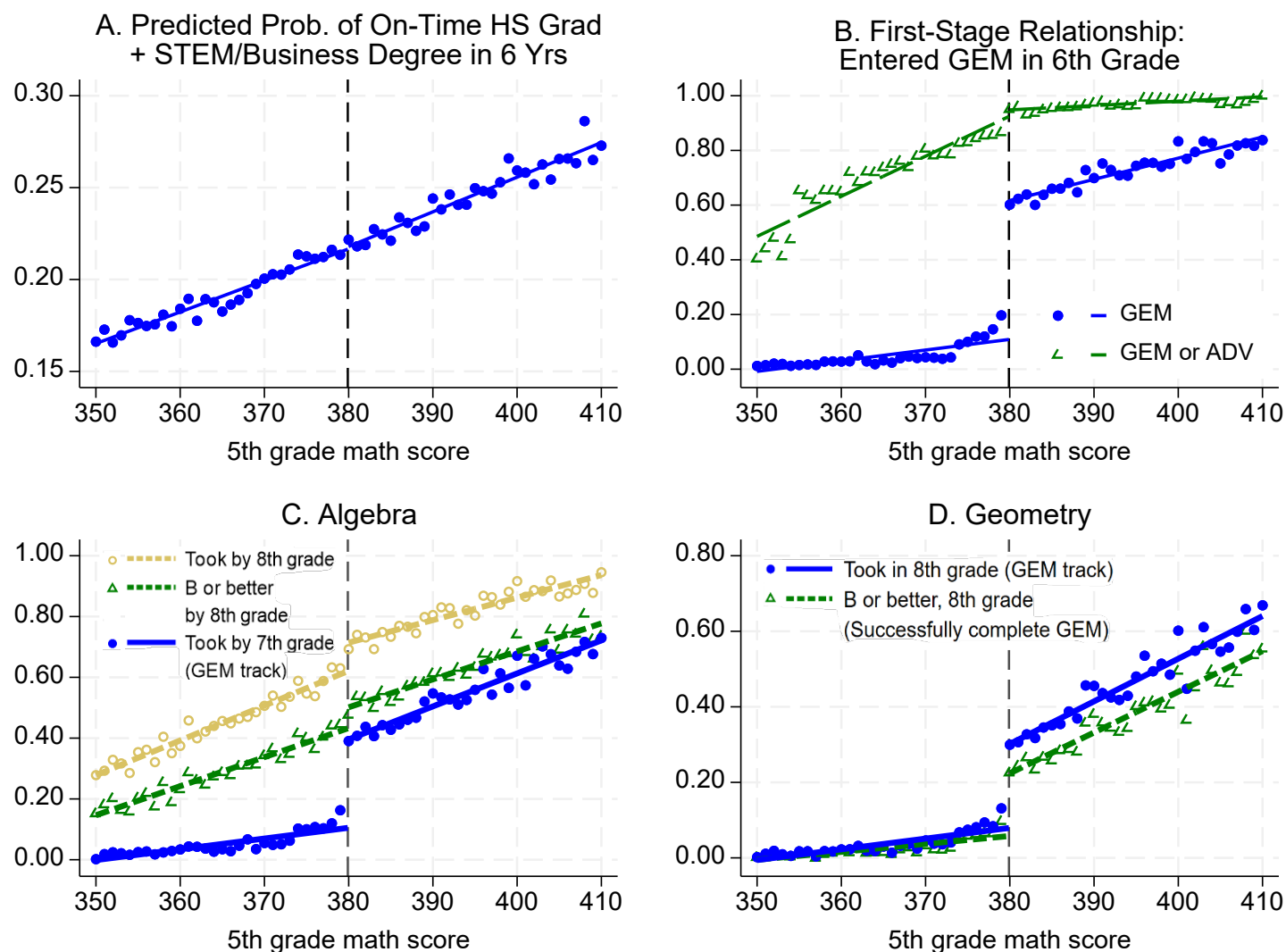
Yoder, Steven. 2024. “Eliminating Advanced Math Tracks Often Prompts Outrage. Some Districts Buck the Trend.” *The Hechinger Report* April 17, 2024.

Figure 1. Math pathways in middle school



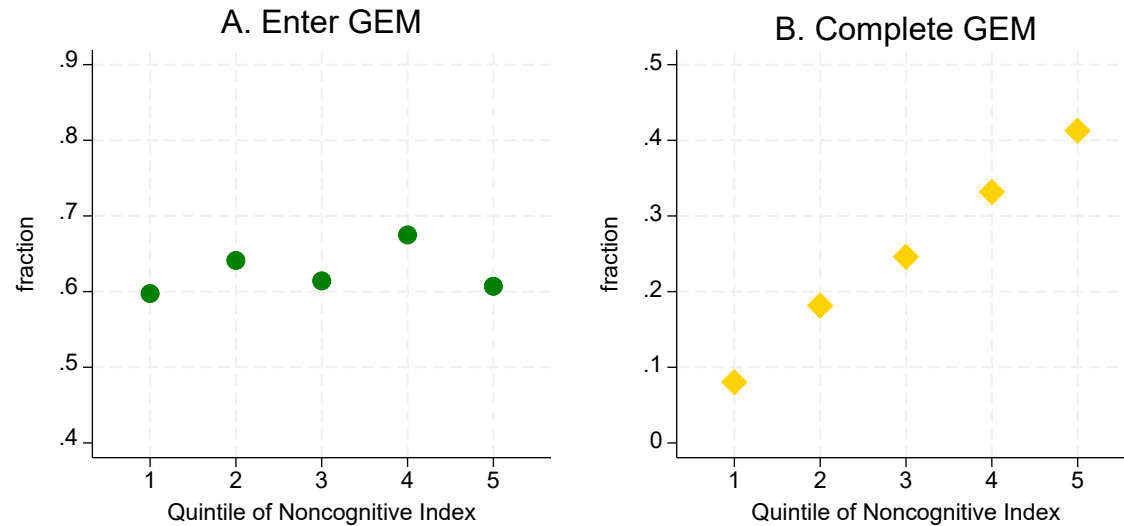
Notes: This figure illustrates various math pathways available to students in all District middle schools.

Figure 2. Predicted outcome, GEM entry, and GEM progression, by fifth-grade math score, pooled sample



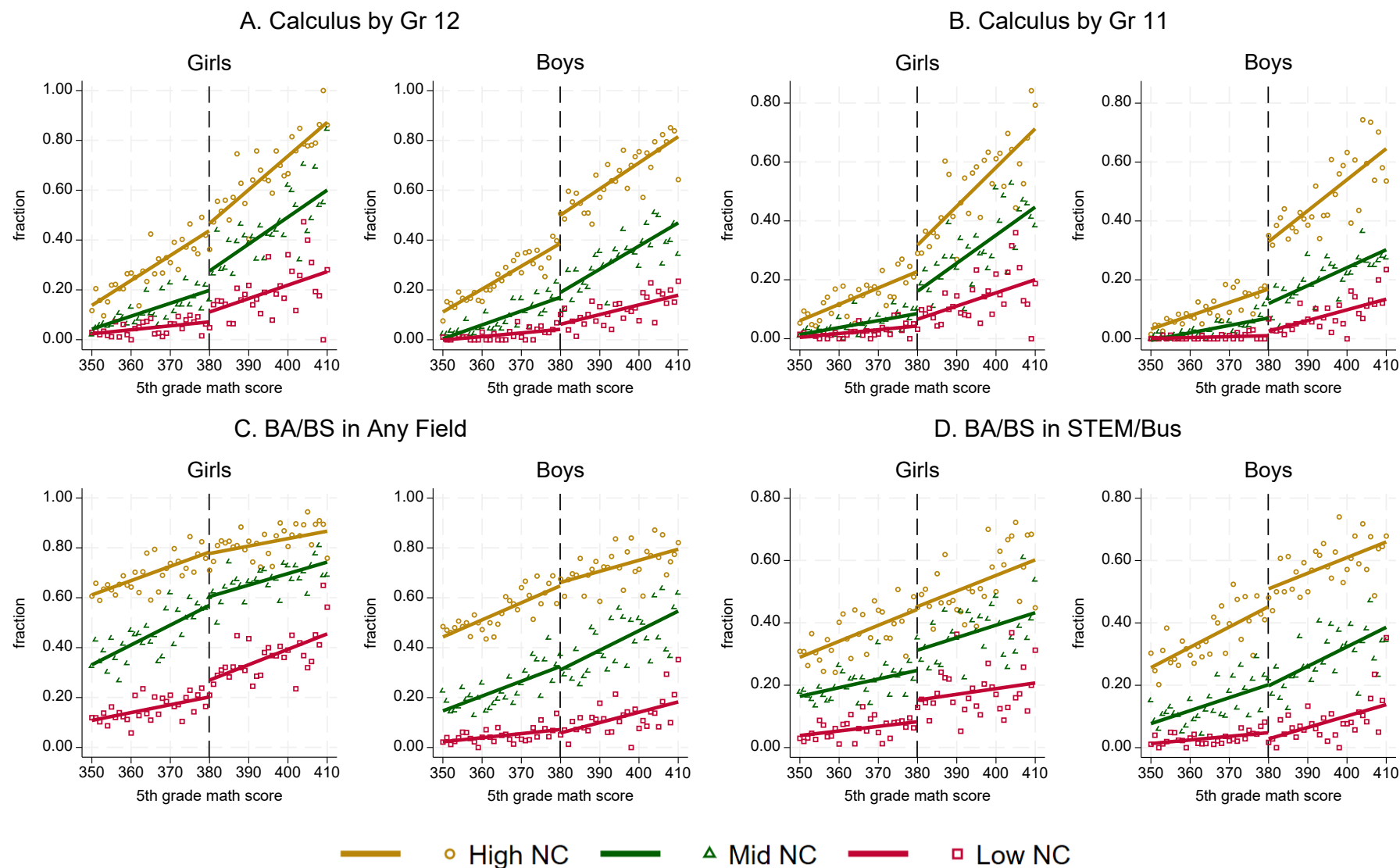
Notes: This figure plots sample means of the dependent variable at each value of 5th-grade math score along with fitted values from local linear RD models with a symmetric bandwidth of 30. The dependent variables (by panel) are: A) an index of pre-determined student characteristics, constructed as the predicted probability of graduating from high school on time and earning a bachelor's degree in a STEM or business field within 6 years; B) indicators for entering the GEM program in 6th grade (the treatment variable, in blue circles) and for entering *either* GEM or advanced math (in green triangles); C) indicators for enrolling in Algebra I in middle school (by 7th or 8th grade) and for completing it with a grade of B or better by the end of 8th grade; D) indicators for enrolling in Geometry in 8th grade and for completing it with a B grade or better (i.e., successfully completing the 3-year GEM sequence).

Figure 3. Shares of barely eligible students who enter GEM in 6th grade and who complete GEM successfully, by quintile of non-cognitive index



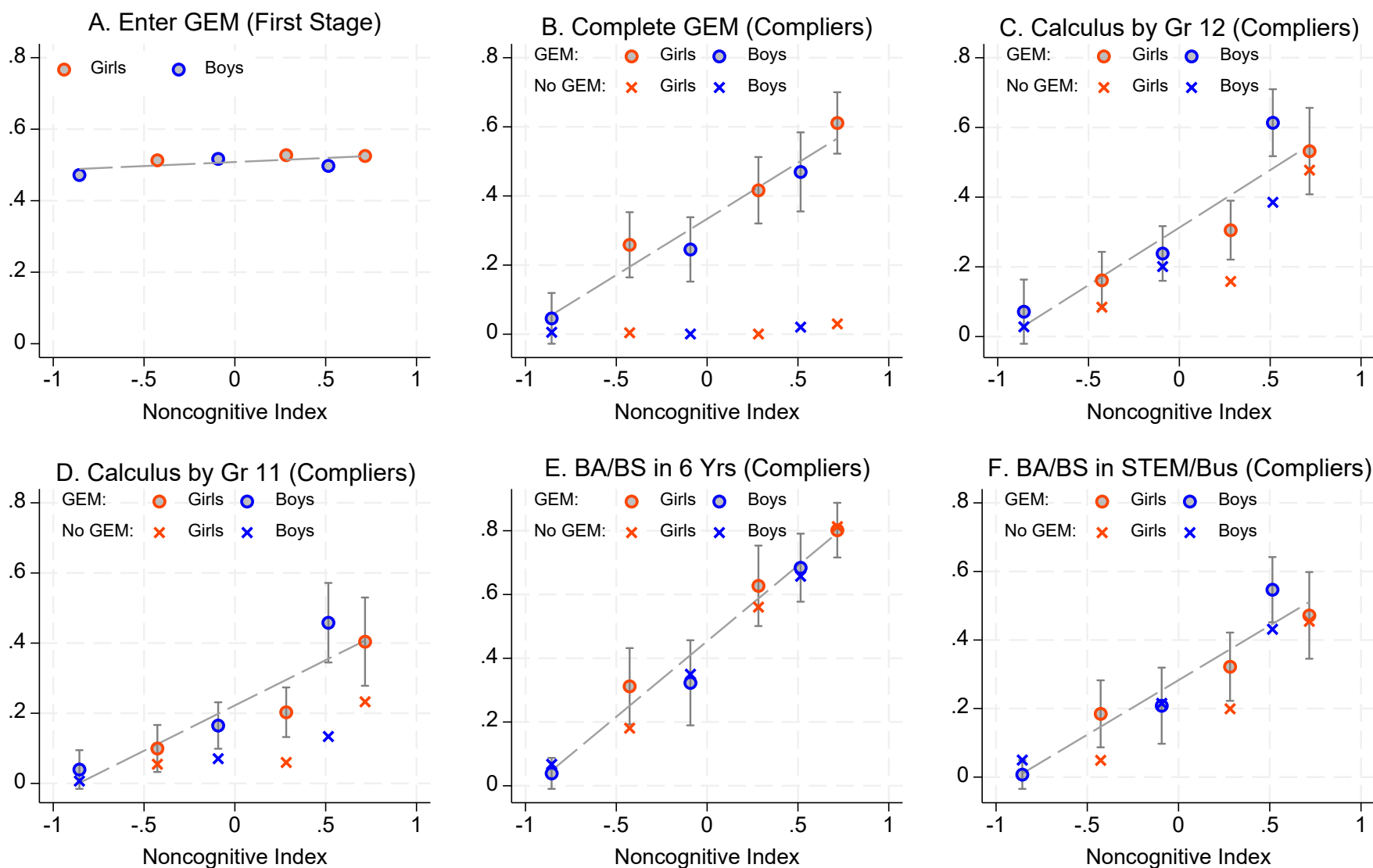
Notes: Panel A plots means of an indicator for entering GEM in 6th grade by quintile of a noncognitive index for the sample of observations with fifth-grade math scores between 380-385 (i.e., within 5 points above the GEM eligibility threshold). Panel B plots means of an indicator for successfully completing all three years of the GEM track by taking Geometry in 8th grade and earning a grade of B or better.

Figure 4. Reduced-form relationships for intermediate and post-secondary outcomes, by gender and noncognitive skill



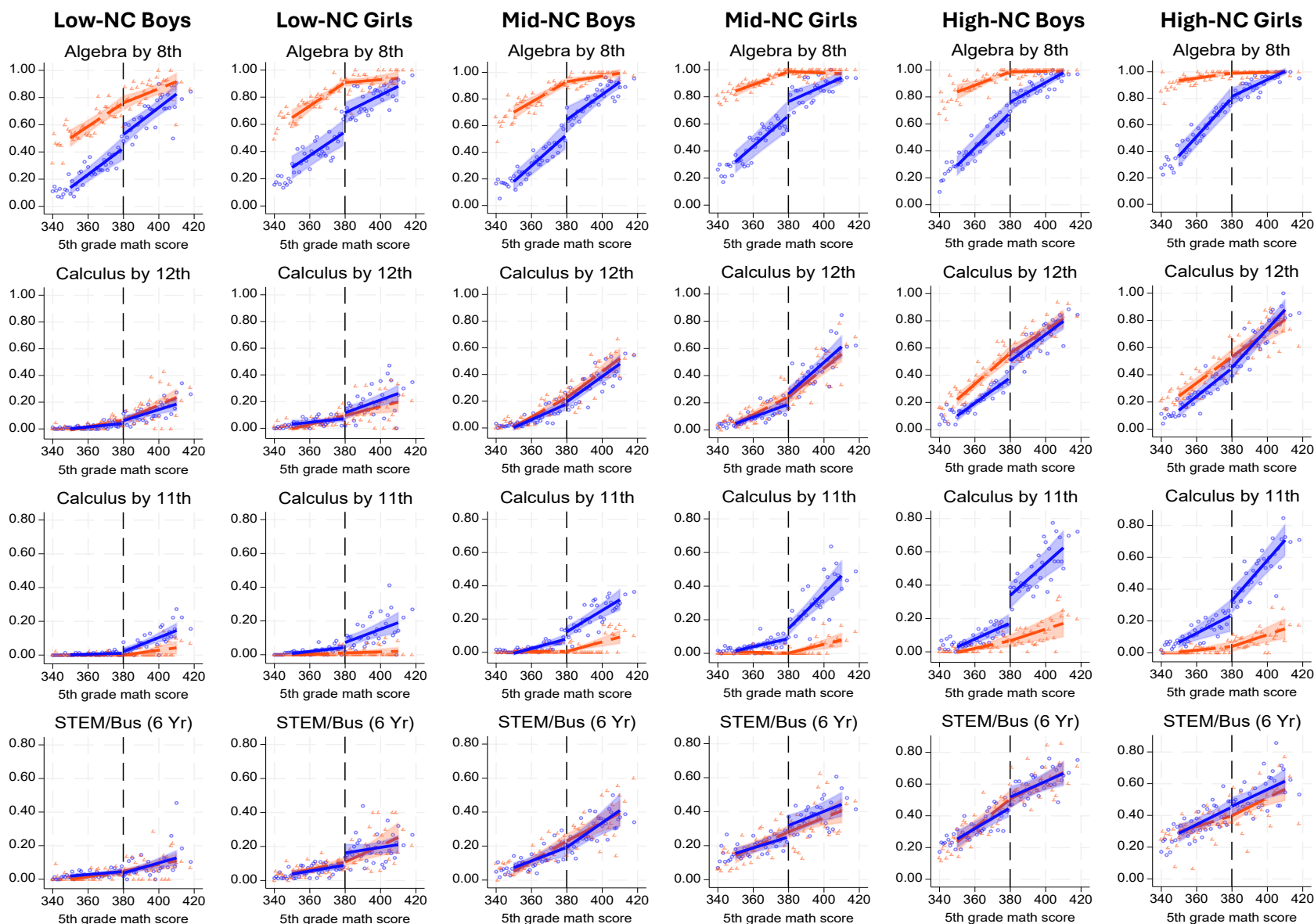
Notes: This figure plots sample means of the dependent variable at each value of 5th-grade math score, along with fitted values from local linear RD models with a symmetric bandwidth of 30, stratifying the sample by gender by gender-specific tercile of the noncognitive skill index. The dependent variables are indicators for: enrolling in any calculus by the end of high school (Panel A); enrolling in calculus by 11th grade (Panel B), earning a bachelor's degree in any field within 6 years of on-time high school graduation (Panel C), and earning a bachelor's degree in a STEM or business major within 6 years (Panel D).

Figure 5. First stage estimates and complier mean outcomes, treated vs. untreated, by gender and noncognitive skill



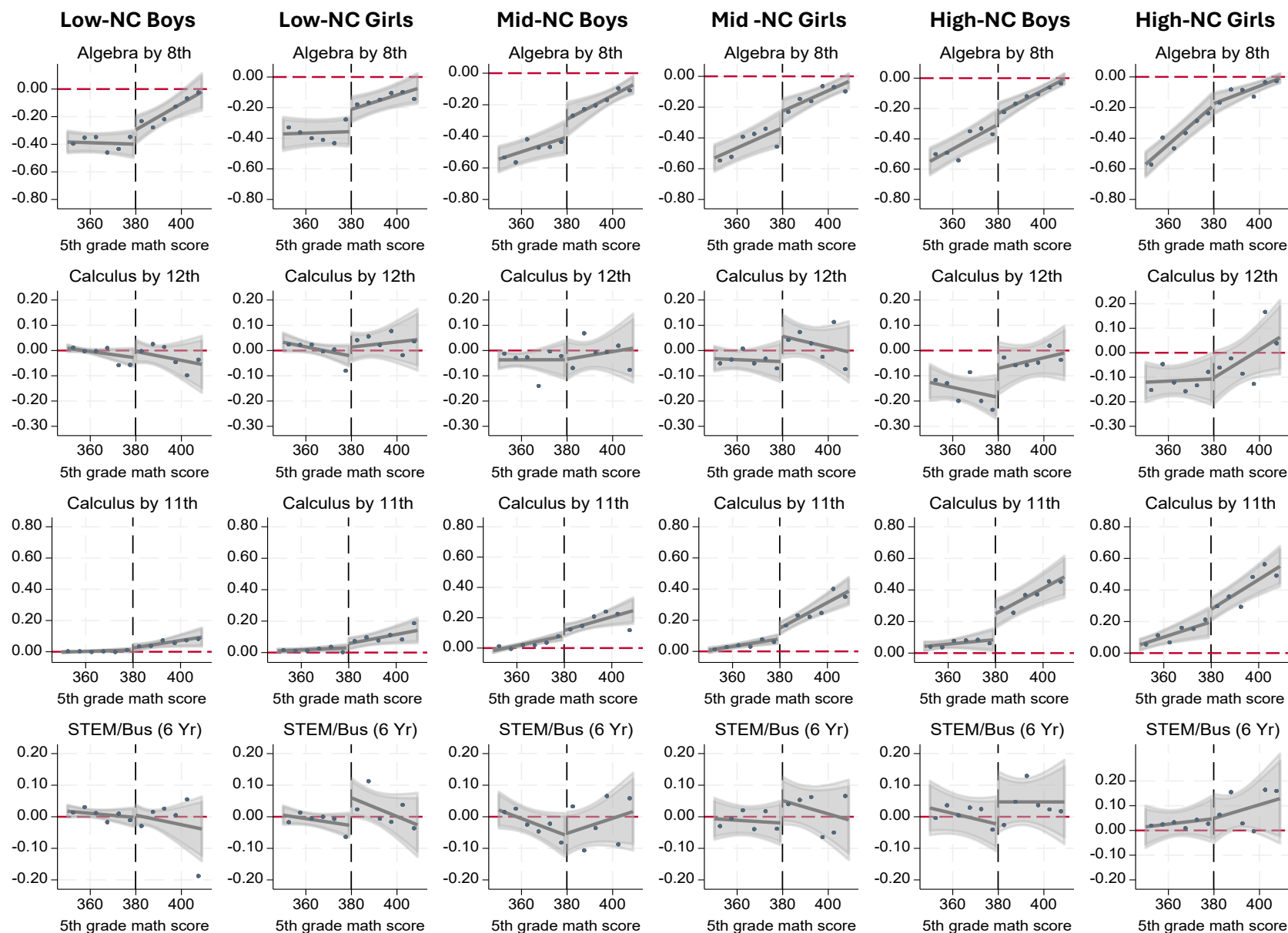
Notes: Panel A plots the first-stage RD estimates for entering GEM in 6th grade separately for six groups defined by gender and gender-specific terciles of noncognitive index against the complier mean of the noncognitive index. Panels B-F plot estimated mean outcomes for treated (GEM) and untreated (No GEM) compliers against the complier mean of the noncognitive index, along with recentered 95% confidence intervals for the 2SLS estimated treatment effects of GEM entry.

Figure 6. Between-district comparison of reduced-form relationships, by gender and noncognitive skill



Notes: Plotted points show sample means of each outcome at each value of 5th-grade math score for the District (blue circles) and District B (orange triangles). Solid blue and dashed orange lines are local linear fits (using a 30-point bandwidth) for the District and District B, respectively; both allow a slope change at the GEM cutoff (380); the District's fit also allows an intercept jump. District B students are classified into gender-NC subgroups using cut points based on the District's gender-specific terciles. Shaded areas are 95% confidence bands; standard errors are clustered on middle school.

Figure 7. Between-district, covariate-adjusted differences, by gender and noncognitive skill



Notes: Each panel plots the estimated outcome gap between the District and District B against the 5th-grade math score, for the indicated gender-NC subgroup. Solid lines are linear fits to data on each side of the cutoff (vertical dashed line), allowing the gap to jump and change slope; markers are mean differences within 5-point bins of the math score. District B observations are reweighted using inverse probability weights and p -scores > 0.9 are trimmed (see section V for details). Shaded areas are 90% and 95% confidence bands; standard errors are clustered on middle school.

Table 1. Estimated means for main analysis sample and subgroups

	District- wide sample	Main analysis sample	High Noncog Girls	High Noncog Boys	Mid Noncog Girls	Mid Noncog Boys	Low Noncog Girls	Low Noncog Boys
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<u>A. Baseline characteristics</u>								
Female	0.51	0.49	1.00	0.00	1.00	0.00	1.00	0.00
Free or reduced price lunch (FRL) recipient	0.42	0.30	0.24	0.21	0.28	0.28	0.40	0.39
White (non-Hispanic)	0.35	0.44	0.46	0.50	0.44	0.46	0.39	0.42
Black (non-Hispanic)	0.35	0.22	0.18	0.13	0.24	0.21	0.31	0.28
Hispanic	0.23	0.24	0.25	0.26	0.23	0.25	0.23	0.24
Gifted	0.07	0.10	0.10	0.12	0.10	0.10	0.10	0.09
Fourth grade (baseline) test scores								
Math score (standardized)	0.14	0.67	0.67	0.76	0.66	0.71	0.60	0.65
Reading score (standardized)	0.13	0.57	0.66	0.54	0.66	0.50	0.62	0.44
School peer characteristics								
Average fourth grade math & reading scores	0.06	0.12	0.17	0.18	0.13	0.13	0.05	0.06
Fraction FRL recipients	0.43	0.38	0.34	0.33	0.37	0.38	0.45	0.44
Fraction GEM eligible	0.19	0.21	0.22	0.22	0.21	0.21	0.19	0.19
<u>B. Mean intermediate outcomes</u>								
Math track in 6th grade								
GEM (treatment)	0.17	0.29	0.27	0.31	0.30	0.31	0.30	0.28
Advanced	0.30	0.52	0.55	0.50	0.52	0.50	0.52	0.50
Regular	0.52	0.22	0.20	0.20	0.20	0.22	0.22	0.28
Middle school course enrollments and grades								
Take algebra 1 by 8th grade	0.34	0.57	0.68	0.64	0.63	0.53	0.54	0.42
Complete algebra 1 with A or B by 8th grade	0.24	0.40	0.62	0.54	0.48	0.33	0.30	0.17
Take algebra I by 7th grade	0.14	0.23	0.24	0.27	0.24	0.23	0.21	0.17
Take geometry in 8th grade	0.12	0.19	0.22	0.24	0.20	0.18	0.16	0.11
Complete geometry with A or B by 8th grade	0.10	0.15	0.21	0.22	0.18	0.12	0.11	0.06
6th grade math GPA	2.79	3.10	3.55	3.35	3.32	2.98	2.91	2.51
7th grade math GPA	2.71	2.99	3.52	3.29	3.20	2.84	2.78	2.35
8th grade math GPA	2.63	2.88	3.41	3.21	3.07	2.74	2.60	2.25
High school course enrollments & grades								
Take algebra 2 in 9th grade	0.15	0.24	0.33	0.32	0.26	0.21	0.19	0.12
Number of advanced math courses	0.97	1.51	2.33	2.20	1.68	1.33	0.94	0.58
Take pre-calculus	0.26	0.41	0.62	0.58	0.46	0.38	0.27	0.18
Take any calculus	0.14	0.22	0.39	0.40	0.22	0.17	0.09	0.05
Take any calculus by 11th grade	0.09	0.14	0.25	0.24	0.13	0.10	0.06	0.03
Take any AP/IB calculus	0.10	0.16	0.29	0.31	0.14	0.12	0.05	0.03
Take AP calculus - BC	0.04	0.05	0.10	0.12	0.03	0.03	0.01	0.01
Take AP statistics	0.07	0.10	0.17	0.18	0.11	0.09	0.05	0.03
High school math GPA (all math courses)	2.20	2.49	3.20	3.00	2.67	2.34	2.04	1.69
High school GPA (non-math academic courses)	2.66	2.93	3.63	3.44	3.21	2.83	2.48	1.98
High school GPA (all academic courses)	2.56	2.83	3.54	3.35	3.10	2.72	2.39	1.91
College board exams								
Take AP calculus exam + scored 3 or better	0.06	0.08	0.17	0.20	0.05	0.05	0.01	0.01
PSAT math percentile	47.5	63.2	67.5	71.2	62.3	63.8	56.1	56.8

Note: Table continues; see notes on next page.

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

	District- wide sample	Main analysis sample	High Noncog Girls	High Noncog Boys	Mid Noncog Girls	Mid Noncog Boys	Low Noncog Girls	Low Noncog Boys
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
C. Mean post-secondary outcomes								
Transition to college								
Graduate high school on time	0.83	0.93	1.00	0.99	0.99	0.98	0.88	0.74
+ enroll in any college within 1 year	0.62	0.74	0.89	0.86	0.84	0.76	0.66	0.46
+ enroll in selective college within 1 year	0.29	0.46	0.73	0.67	0.57	0.41	0.26	0.11
College degree(s) earned within 6 years								
Any degree	0.41	0.54	0.82	0.72	0.67	0.47	0.38	0.18
Bachelor's (BA or BS) in any field	0.28	0.41	0.73	0.61	0.52	0.30	0.22	0.07
Bachelor's in STEM or Business field	0.16	0.24	0.41	0.44	0.26	0.19	0.10	0.05
Bachelor's degree with major in:								
math, engineering, physical sciences, comp. science	0.03	0.04	0.04	0.14	0.02	0.05	0.01	0.01
life sciences	0.03	0.05	0.12	0.07	0.06	0.03	0.02	0.00
business	0.05	0.08	0.11	0.15	0.09	0.08	0.03	0.02
social sciences	0.05	0.07	0.14	0.08	0.10	0.04	0.05	0.01
health (nonSTEM)	0.03	0.04	0.10	0.02	0.07	0.01	0.03	0.00
nonSTEM (all other)	0.09	0.13	0.23	0.15	0.20	0.10	0.10	0.02
Number of observations	56,003	22,754	3,732	3,852	3,732	3,853	3,732	3,853

Notes: The first column shows means of variables for all students in the district who meet the sample inclusion criteria (see Appendix Table A2, column 3). Column 2 shows means for the main analysis sample (math scores between 350 and 410). The remaining columns show means for six subgroups defined by gender and gender-specific tercile of a noncognitive index.

Table 2. Estimated treatment effects for marginal GEM entrants on middle school math course enrollments

	First stage (GEM entry)	Algebra I			Geometry	
		Took by 7th grade (GEM)	Took by 8th grade	Earned B or better by 8th gr.	Took by 8th grade (GEM)	Earned B or better by 8th gr.
		(1)	(2)	(3)	(4)	(5)
1. Pooled (N=22,754)	0.508 (0.028)	0.573 (0.033) [0.040]	0.184 (0.036) [0.647]	0.136 (0.029) [0.453]	0.438 (0.033) [0.019]	0.328 (0.026) [0.010]
2. Girls, High Noncog (N=3,732)	0.525 (0.040)	0.746 (0.035) [0.055]	0.037 (0.053) [0.869]	0.052 (0.061) [0.771]	0.696 (0.040) [0.034]	0.583 (0.044) [0.030]
3. Boys, High Noncog (N=3,853)	0.498 (0.032)	0.637 (0.055) [0.038]	0.168 (0.062) [0.738]	0.129 (0.068) [0.611]	0.531 (0.061) [0.020]	0.453 (0.057) [0.022]
4. Girls, Mid Noncog (N=3,732)	0.528 (0.037)	0.620 (0.047) [0.031]	0.186 (0.061) [0.707]	0.132 (0.062) [0.573]	0.522 (0.052) [0.003]	0.417 (0.048) [0.000]
5. Boys, Mid Noncog (N=3,853)	0.517 (0.044)	0.517 (0.052) [0.040]	0.229 (0.064) [0.589]	0.176 (0.060) [0.351]	0.381 (0.051) [0.012]	0.246 (0.047) [0.001]
6. Girls, Low Noncog (N=3,732)	0.513 (0.030)	0.538 (0.052) [0.038]	0.235 (0.058) [0.587]	0.200 (0.048) [0.315]	0.358 (0.052) [0.020]	0.255 (0.047) [0.004]
7. Boys, Low Noncog (N=3,853)	0.472 (0.043)	0.396 (0.055) [0.041]	0.248 (0.064) [0.393]	0.146 (0.053) [0.095]	0.159 (0.050) [0.024]	0.041 (0.037) [0.006]

Notes: Column (1) shows first-stage estimates for the probability of entering the GEM track in 6th grade. Entries in columns (2)-(6) are 2SLS estimates for the effect of entry into GEM in 6th grade for the marginally eligible compliers (i.e., they are estimates of β in equation 3). Standard errors (in parentheses) are clustered on the school where the student was enrolled in 6th grade. Square brackets contain estimated means for untreated compliers.

Table 3. Estimated treatment effects for marginal GEM entrants on grades and college board scores

	Course grades/GPA					College boards	
	6th grade math GPA	7th grade math GPA	8th grade math GPA	HS math GPA	HS non-math GPA	PSAT math pctile	AP calc exam score ≥3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. Pooled (N=22,754)	-0.672 (0.082) [3.416]	-0.382 (0.051) [3.176]	-0.056 (0.054) [2.961]	-0.054 (0.047) [2.575]	-0.012 (0.032) [3.009]	3.032 (1.010) [67.724]	0.029 (0.014) [0.065]
2. Girls, High Noncog (N=3,732)	-0.610 (0.091) [3.817]	-0.460 (0.075) [3.708]	-0.143 (0.087) [3.468]	-0.041 (0.072) [3.238]	0.008 (0.022) [3.706]	1.126 (1.881) [73.790]	0.062 (0.047) [0.146]
3. Boys, High Noncog (N=3,853)	-0.503 (0.114) [3.505]	-0.284 (0.100) [3.353]	-0.025 (0.107) [3.167]	0.022 (0.085) [3.050]	0.036 (0.038) [3.453]	6.221 (1.916) [74.248]	0.120 (0.046) [0.141]
4. Girls, Mid Noncog (N=3,732)	-0.572 (0.093) [3.557]	-0.393 (0.083) [3.338]	-0.120 (0.099) [3.168]	-0.047 (0.056) [2.731]	0.022 (0.022) [3.270]	3.513 (1.855) [66.548]	0.017 (0.030) [0.032]
5. Boys, Mid Noncog (N=3,853)	-0.715 (0.104) [3.314]	-0.338 (0.096) [3.021]	-0.175 (0.093) [2.915]	-0.067 (0.060) [2.387]	-0.010 (0.026) [2.901]	-0.631 (2.297) [72.094]	-0.025 (0.032) [0.060]
6. Girls, Low Noncog (N=3,732)	-0.774 (0.137) [3.315]	-0.494 (0.109) [3.073]	0.112 (0.108) [2.700]	-0.111 (0.095) [2.256]	-0.008 (0.064) [2.572]	1.852 (2.483) [60.769]	-0.016 (0.011) [0.017]
7. Boys, Low Noncog (N=3,853)	-0.825 (0.132) [2.996]	-0.307 (0.132) [2.583]	0.042 (0.119) [2.360]	-0.028 (0.080) [1.791]	-0.041 (0.072) [2.135]	7.322 (3.119) [56.637]	0.030 (0.013) [0.000]

Notes: Table entries are 2SLS estimates for the effect of entry into GEM in 6th grade for the marginally eligible compliers (i.e., they are estimates of β in equation 3). All models include 4 dummy variables for the student's cohort. Standard errors (in parentheses) are clustered on the school where the student was enrolled in 6th grade. Square brackets contain estimated means for untreated compliers.

Table 4. Estimated treatment effects for marginal GEM entrants on high school math course enrollments

	Algebra 2 by 9th grade	Number advanced math courses	Pre- calculus	Any calculus	Any calculus by 11th grade	Any AP calculus	AP calculus BC	AP statistics
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. Pooled (N=22,754)	0.333 (0.025) [0.127]	0.394 (0.094) [1.563]	0.070 (0.030) [0.478]	0.095 (0.025) [0.221]	0.134 (0.023) [0.092]	0.068 (0.019) [0.145]	0.029 (0.014) [0.035]	0.039 (0.016) [0.105]
2. Girls, High Noncog (N=3,732)	0.483 (0.059) [0.237]	0.189 (0.235) [2.667]	-0.087 (0.060) [0.801]	0.056 (0.061) [0.478]	0.173 (0.063) [0.232]	0.131 (0.060) [0.304]	0.084 (0.044) [0.068]	0.041 (0.043) [0.196]
3. Boys, High Noncog (N=3,853)	0.474 (0.054) [0.154]	0.712 (0.159) [2.182]	0.186 (0.058) [0.614]	0.236 (0.048) [0.380]	0.330 (0.055) [0.131]	0.178 (0.049) [0.263]	0.077 (0.048) [0.075]	0.085 (0.045) [0.151]
4. Girls, Mid Noncog (N=3,732)	0.385 (0.054) [0.113]	0.526 (0.146) [1.707]	0.105 (0.055) [0.520]	0.149 (0.043) [0.157]	0.144 (0.035) [0.059]	0.063 (0.042) [0.106]	0.008 (0.022) [0.019]	0.054 (0.041) [0.102]
5. Boys, Mid Noncog (N=3,853)	0.329 (0.043) [0.077]	0.353 (0.127) [1.420]	0.047 (0.058) [0.497]	0.040 (0.040) [0.201]	0.097 (0.034) [0.070]	0.027 (0.040) [0.133]	0.002 (0.023) [0.039]	0.009 (0.039) [0.102]
6. Girls, Low Noncog (N=3,732)	0.235 (0.050) [0.108]	0.413 (0.142) [0.925]	0.107 (0.055) [0.275]	0.076 (0.041) [0.085]	0.045 (0.034) [0.055]	0.008 (0.025) [0.049]	-0.002 (0.013) [0.005]	0.028 (0.030) [0.057]
7. Boys, Low Noncog (N=3,853)	0.112 (0.053) [0.075]	0.294 (0.156) [0.456]	0.080 (0.050) [0.163]	0.043 (0.046) [0.028]	0.032 (0.028) [0.007]	0.029 (0.027) [0.018]	0.015 (0.009) [0.004]	0.031 (0.024) [0.028]

Notes: Table entries are 2SLS estimates for the effect of entry into GEM in 6th grade for the marginally eligible compliers (i.e., they are estimates of β in equation 3). All models include 4 dummy variables for the student's cohort. Standard errors (in parentheses) are clustered on the school where the student was enrolled in 6th grade. Square brackets contain estimated means for untreated compliers.

Table 5. Estimated treatment effects for marginal GEM entrants on college entry and degree earning

	Graduate high school on time	Graduate HS & enter college on time	Enter selective college on time	Earn any degree in 6 years	Earn any bachelors degree in 6 years	Bachelors in STEM/ business in 6 years
	(1)	(2)	(3)	(4)	(5)	(6)
1. Pooled (N=22,754)	-0.009 (0.012) [0.962]	0.006 (0.021) [0.776]	0.018 (0.024) [0.514]	0.022 (0.020) [0.561]	0.023 (0.020) [0.437]	0.055 (0.023) [0.232]
2. Girls, High Noncog (N=3,732)	0.000 (0.007) [0.995]	0.036 (0.036) [0.895]	0.013 (0.046) [0.822]	0.006 (0.041) [0.851]	-0.009 (0.042) [0.810]	0.018 (0.063) [0.453]
3. Boys, High Noncog (N=3,853)	0.002 (0.008) [0.996]	0.008 (0.047) [0.871]	0.031 (0.047) [0.708]	-0.002 (0.052) [0.776]	0.031 (0.055) [0.652]	0.119 (0.049) [0.428]
4. Girls, Mid Noncog (N=3,732)	0.003 (0.010) [0.993]	0.066 (0.039) [0.819]	0.057 (0.055) [0.632]	0.050 (0.066) [0.691]	0.066 (0.063) [0.561]	0.124 (0.050) [0.199]
5. Boys, Mid Noncog (N=3,853)	0.001 (0.017) [0.988]	0.003 (0.047) [0.836]	-0.035 (0.062) [0.526]	0.011 (0.076) [0.517]	-0.028 (0.068) [0.352]	-0.008 (0.056) [0.216]
6. Girls, Low Noncog (N=3,732)	-0.025 (0.040) [0.914]	-0.016 (0.062) [0.675]	0.141 (0.058) [0.229]	0.093 (0.067) [0.342]	0.131 (0.060) [0.181]	0.136 (0.049) [0.049]
7. Boys, Low Noncog (N=3,853)	-0.019 (0.050) [0.878]	-0.040 (0.066) [0.548]	-0.061 (0.045) [0.156]	-0.004 (0.051) [0.182]	-0.027 (0.024) [0.065]	-0.041 (0.021) [0.048]

Notes: Table entries are 2SLS estimates for the effect of entry into GEM in 6th grade for the marginally eligible compliers (i.e., they are estimates of β in equation 3). All models include 4 dummy variables for the student's cohort. Standard errors (in parentheses) are clustered on the school where the student was enrolled in 6th grade. Square brackets contain estimated means for untreated compliers.

Table 6. Estimated treatment effects for marginal GEM entrants on earning field-specific bachelor's degree

	STEM/ business	Math, engr., phys.sci., comp.sci.	Life sciences	Business	Social sciences	Health	Non-STEM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. Pooled (N=22,754)	0.055 (0.023) [0.232]	0.006 (0.011) [0.045]	0.017 (0.009) [0.054]	0.012 (0.014) [0.077]	0.020 (0.017) [0.057]	-0.012 (0.011) [0.046]	-0.019 (0.014) [0.159]
2. Girls, High Noncog (N=3,732)	0.018 (0.063) [0.453]	0.026 (0.026) [0.029]	0.043 (0.035) [0.140]	0.005 (0.048) [0.105]	-0.065 (0.049) [0.187]	-0.060 (0.037) [0.124]	0.030 (0.056) [0.241]
3. Boys, High Noncog (N=3,853)	0.119 (0.049) [0.428]	0.044 (0.042) [0.144]	0.028 (0.034) [0.075]	-0.026 (0.039) [0.181]	0.080 (0.032) [0.028]	-0.020 (0.016) [0.025]	-0.068 (0.040) [0.199]
4. Girls, Mid Noncog (N=3,732)	0.124 (0.050) [0.199]	0.003 (0.018) [0.014]	0.001 (0.023) [0.074]	0.067 (0.033) [0.053]	0.054 (0.045) [0.059]	-0.002 (0.039) [0.083]	-0.051 (0.044) [0.276]
5. Boys, Mid Noncog (N=3,853)	-0.008 (0.056) [0.216]	-0.031 (0.027) [0.080]	0.004 (0.020) [0.031]	0.010 (0.050) [0.078]	0.009 (0.023) [0.028]	0.003 (0.013) [0.016]	-0.024 (0.035) [0.120]
6. Girls, Low Noncog (N=3,732)	0.136 (0.049) [0.049]	0.015 (0.013) [-0.008]	0.042 (0.017) [0.005]	0.034 (0.020) [0.012]	0.048 (0.037) [0.041]	0.005 (0.022) [0.029]	-0.006 (0.029) [0.100]
7. Boys, Low Noncog (N=3,853)	-0.041 (0.021) [0.048]	-0.010 (0.012) [0.007]	-0.019 (0.009) [0.006]	-0.007 (0.020) [0.033]	-0.006 (0.013) [0.003]	-0.004 (0.007) [0.005]	0.019 (0.020) [0.012]

Notes: Table entries are 2SLS estimates for the effect of entry into GEM in 6th grade for the marginally eligible compliers (i.e., they are estimates of β in equation 3). All models include 4 dummy variables for the student's cohort. Standard errors (in parentheses) are clustered on the school where the student was enrolled in 6th grade. Square brackets contain estimated means for untreated compliers.

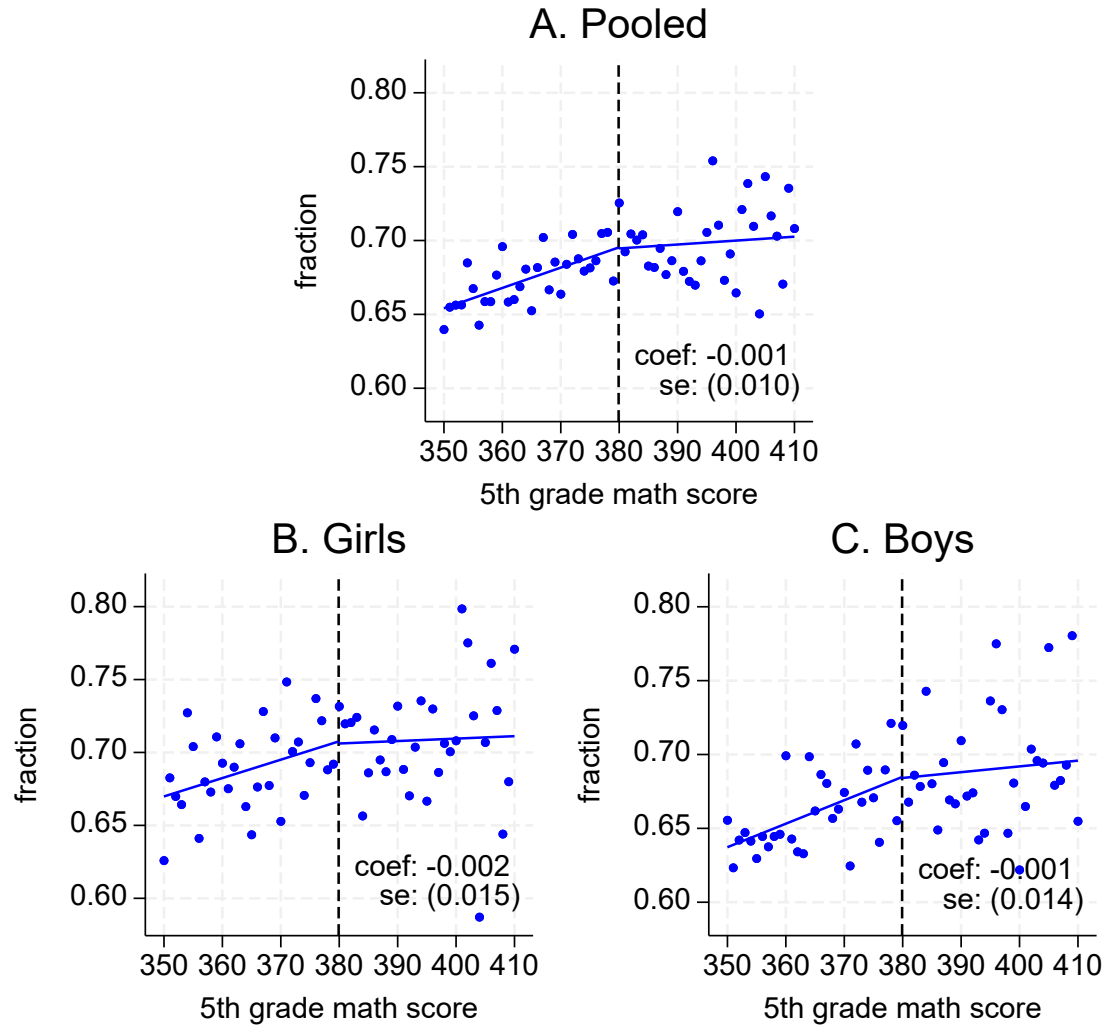
Online Appendix

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

David Card, and Laura Giuliano

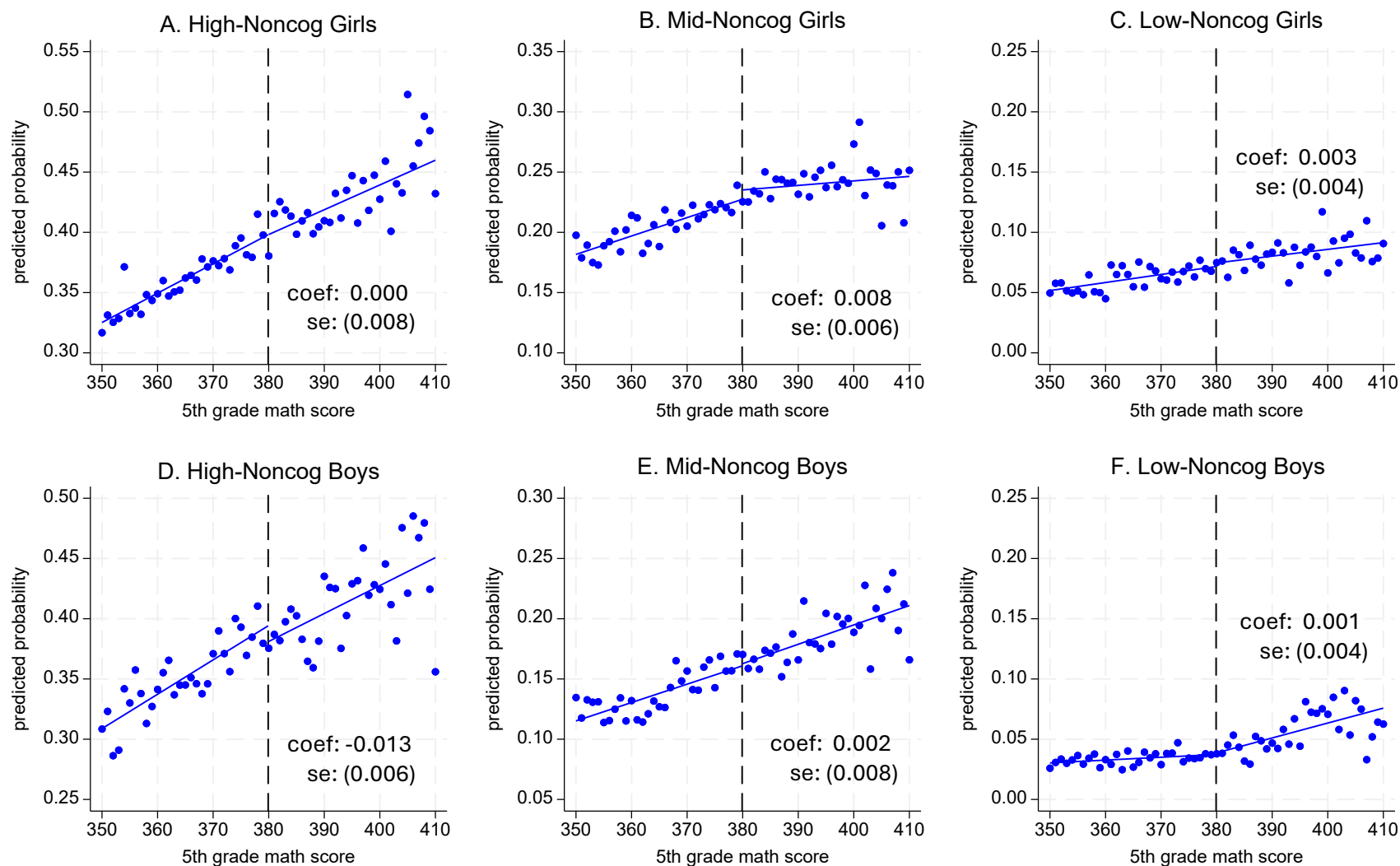
Appendix A.
Additional Figures and Tables

Figure A1. Reduced-form relationships for meeting sample criteria



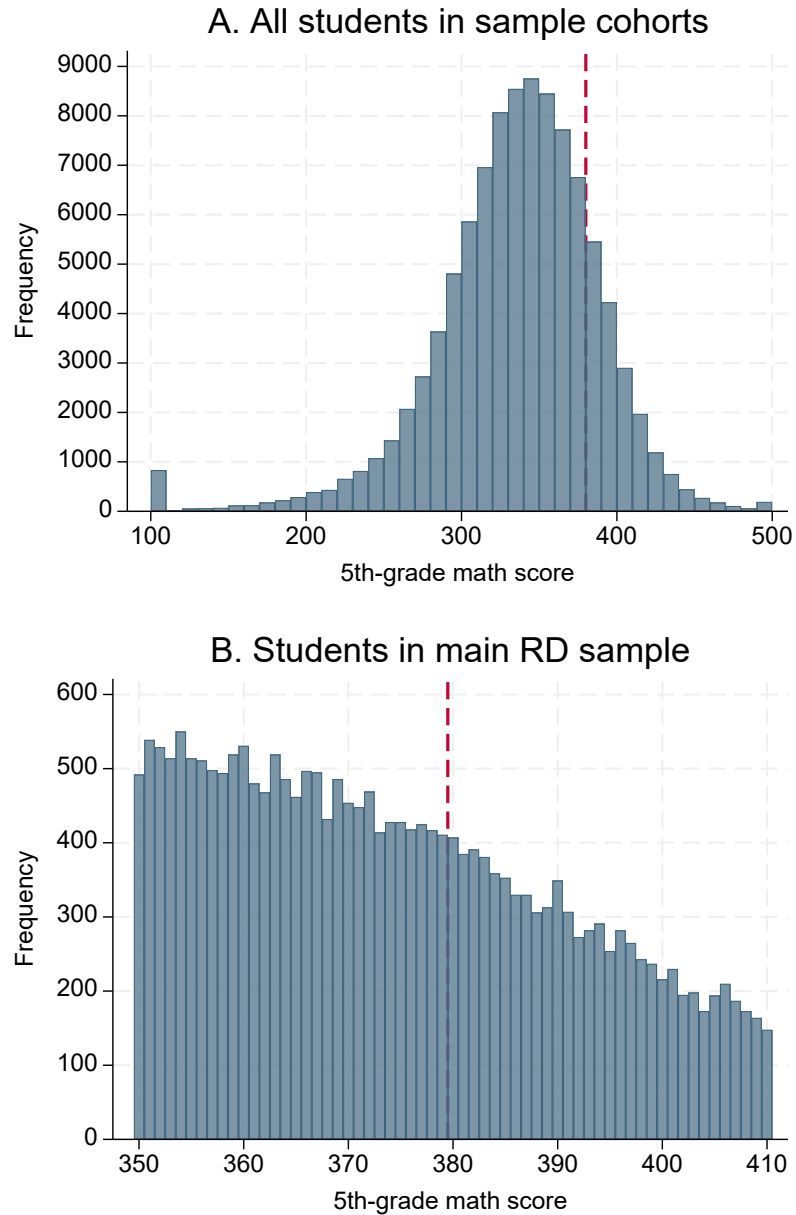
Notes: This figure plots means of an indicator for meeting the inclusion criteria for the main analysis sample (see Appendix Table A2), using all students who enrolled in the District for 5th grade between 2003 and 2007. It also shows fitted values from a local linear RD model with a symmetric bandwidth of 30. Panel A pools girls and boys while Panels B and C show the relationships separately by gender. Each panel also reports the estimated discontinuity at the GEM threshold of 380; parentheses contain standard errors, clustered on elementary school.

Figure A2. Reduced-form relationships for the predicted probability of completing a bachelor's degree in STEM or business, by gender and noncognitive skill



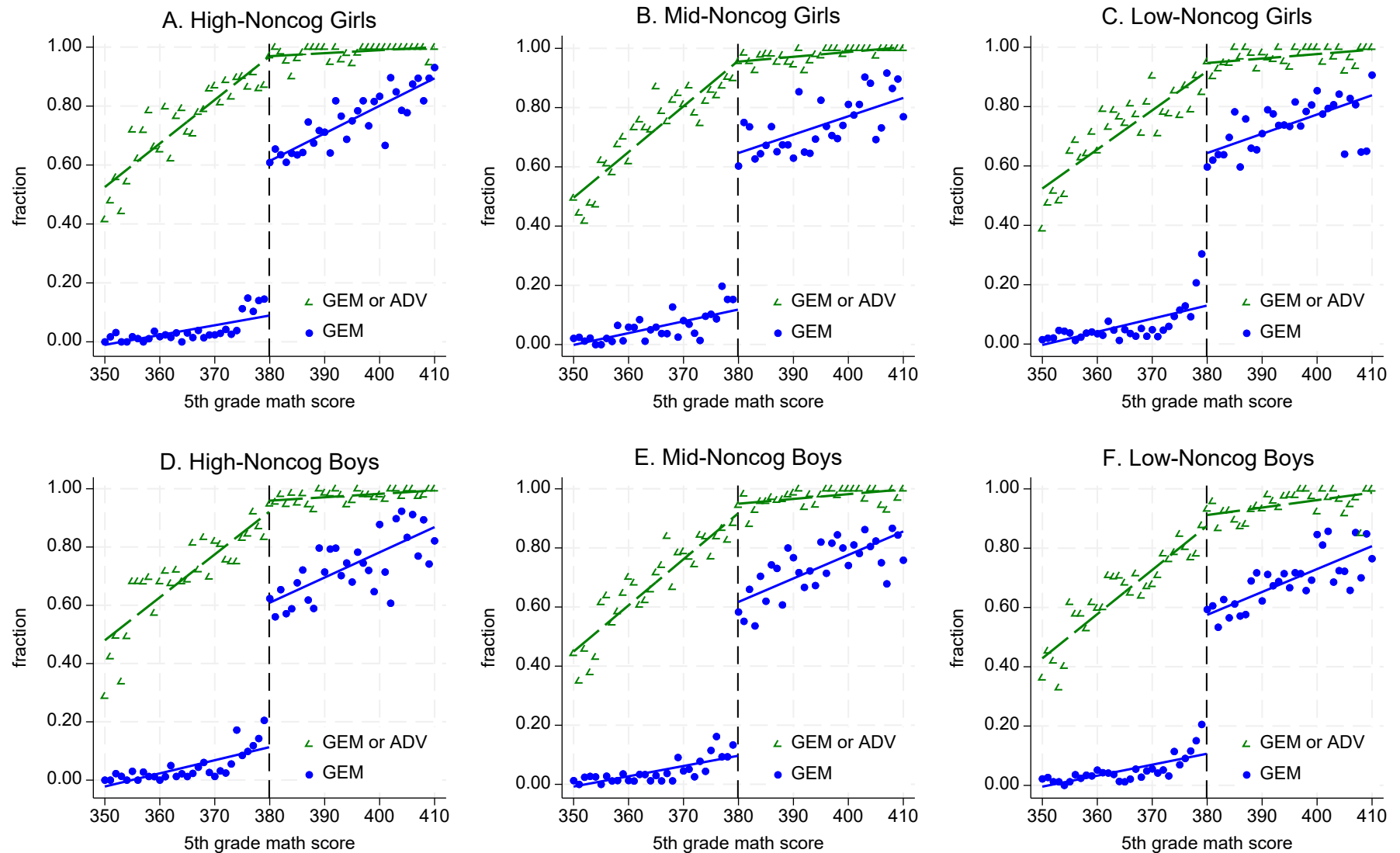
Notes: The figure plots sample means, along with fitted values from local linear RD models with symmetric bandwidths of ± 30 points, of an index of observable pre-treatment characteristics. The index is constructed as the predicted probability of graduating high school on time and completing a bachelor's degree in a STEM or business field within 6 years, from models fit to observations in our RD sample with 5th-grade math scores below the threshold, separately for each gender-NC group. The model includes: second-order polynomials in 4th-grade math and reading test scores; student age; and indicators for student gender, four race/ethnicity categories, FRL status, English Language Learner status, gifted status, and cohort.

Figure A3: Histograms of running variable



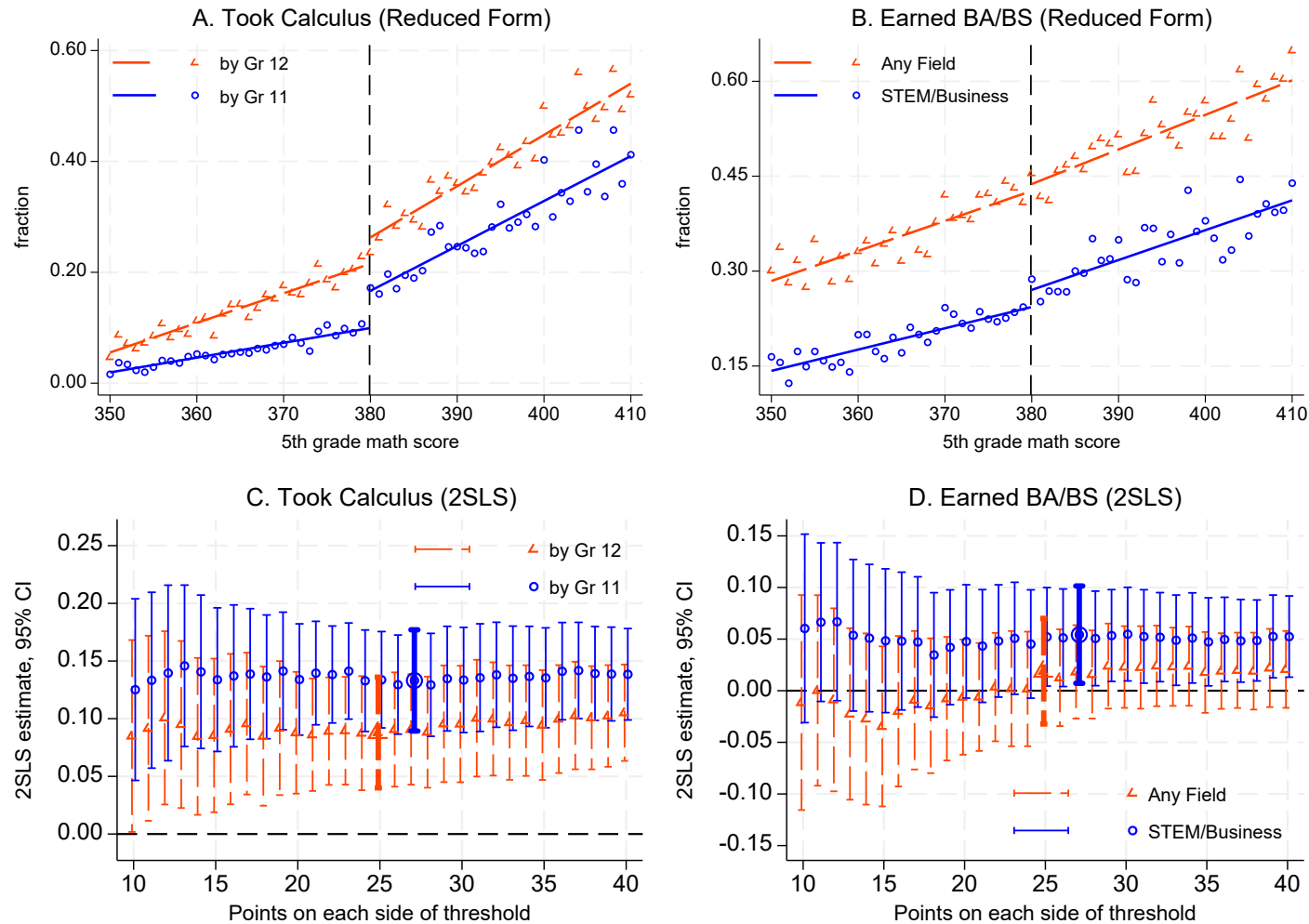
Notes: This figure shows the frequency distribution of 5th-grade math scores for all students in the studied cohorts who enrolled in a District school for 5th-grade, as in column (1) of Appendix Table A2, (Panel A) and for students in the main RD analysis sample (Panel B). The dashed vertical lines in each panel indicate the threshold.

Figure A4. First-stage relationships for GEM entry, by gender and noncognitive skill



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 5th-grade math score (the running variable), along with fitted values from local linear RD models with a symmetric bandwidth of 30.

Figure A5. Reduced-form relationships and 2SLS estimates by bandwidth



Notes: Panels A and B plot sample means of the dependent variable at each value of 5th-grade math score (the running variable), along with fitted values from local linear RD models with a symmetric bandwidth of 30. Panels C and D plot 2SLS estimates and 95% confidence intervals, for the treatment effect of GEM entry on the dependent variable, for symmetric bandwidths between 10 and 40. Estimates that use the CCT symmetric bandwidth are highlighted with larger markers and thicker lines.

**Appendix Table A1. Math Course Enrollment Distributions in Middle and High School
by Grade Level and Middle School Math Track/GEM Status**

Course	A. Enrolled in regular track in 6th grade							B. Enrolled in advanced track in 6th grade						
	Grade Level						All	Grade Level						All
	7 th	8 th	9 th	10 th	11 th	12 th		7 th	8 th	9 th	10 th	11 th	12 th	
Pre-algebra/middle school math	99.6	92.1						98.0	52.6					
Remedial	5.2	5.8	9.6	5.6	2.3	0.8	19.6	0.6	0.7	1.8	1.3	0.3	0.1	3.6
Algebra I	0.4	7.7	89.7	24.3	6.2	2.7	96.0	2.0	46.5	60.7	5.4	1.3	0.9	97.8
Geometry		0.1	13.1	72.9	29.4	13.3	92.7		0.9	49.4	51.8	9.6	4.2	97.2
Algebra II			1.5	16.6	45.1	22.7	67.7			9.7	43.9	39.4	12.5	90.2
Other math (see note)			10.0	5.2	23.9	49.5	58.4			3.9	7.9	19.4	32.1	46.9
Other advanced math (see note)			0.0	0.7	4.5	16.4	15.4			0.0	4.2	14.6	26.8	33.9
Pre-calculus			0.0	1.4	5.5	6.8	10.7			0.0	7.3	18.7	10.7	31.8
AP statistics			0.0	0.0	0.2	1.7	1.4			0.0	0.1	1.1	6.5	6.3
Non-AP calculus			0.0	0.1	0.8	1.4	1.8			0.0	0.9	3.9	5.0	8.4
AP calculus AB			0.0	0.0	0.3	1.1	1.1			0.0	0.0	2.9	6.2	7.6
AP calculus BC			0.0	0.0	0.0	0.2	0.2			0.0	0.0	0.4	2.1	2.0
Dual enrolment math			0.0	0.0	0.3	2.3	1.8			0.0	0.0	1.9	8.2	7.3
<i>No math course</i>			<i>0.6</i>	<i>1.2</i>	<i>3.8</i>	<i>17.0</i>				<i>0.3</i>	<i>0.7</i>	<i>4.0</i>	<i>15.9</i>	

Course	C. Entered GEM, did not complete							D. Completed GEM successfully						
	Grade Level						All	Grade Level						All
	7 th	8 th	9 th	10 th	11 th	12 th		7 th	8 th	9 th	10 th	11 th	12 th	
Pre-algebra/middle school math	51.4	21.5						0.1	0.0					
Remedial		0.2	0.4	0.3	0.1		1.0	0.1	0.1	0.1				0.3
Algebra I	48.6	54.4	31.2	1.8	0.3	0.3	98.7	99.9	0.0	4.2	0.1			100.0
Geometry		24.1	63.1	24.2	4.0	2.4	98.6		99.9	2.4	0.6		0.1	100.0
Algebra II			32.1	52.2	22.8	7.2	95.4		0.1	96.8	6.4	1.5	0.7	98.6
Other math (see note)			4.5	14.3	17.8	21.1	43.1			8.1	8.0	3.2	2.5	17.1
Other advanced math (see note)			1.1	10.6	21.8	28.2	43.7			11.9	20.6	14.9	17.4	47.9
Pre-calculus			0.6	18.9	25.4	12.0	47.9			4.7	65.1	15.5	4.3	81.6
AP statistics				0.2	2.4	11.0	10.5			0.1	6.5	7.7	27.2	34.4
Non-AP calculus				2.3	9.1	7.3	15.6				13.9	18.6	5.0	34.2
AP calculus AB				0.2	7.6	9.2	13.7				5.4	41.6	12.4	53.1
AP calculus BC					1.0	4.4	4.3				0.4	10.3	26.0	29.7
Dual enrolment math				0.1	3.1	10.6	8.9				0.2	8.2	20.0	17.6
<i>No math course</i>			<i>0.4</i>	<i>1.0</i>	<i>6.5</i>	<i>22.3</i>				<i>0.1</i>	<i>1.0</i>	<i>6.9</i>	<i>22.6</i>	

Notes: Table entries show, for students in each group and grade level as indicated by the column headers, the percentage enrolled in each high school math course or group of courses. The category "other math" (four row) includes an "Integrated Math" course that covers topics from algebra and trigonometry plus data analysis, and courses such as "Liberal Arts Math" and "Math College Readiness" that emphasize practical math skills or college preparation for non-STEM majors. The category "other advanced math" (fifth row) includes probability and statistics (excluding AP statistics), math analysis, trigonometry, and other related courses. Because students can enroll in multiple math course in the same year, the column entries sum to >100. Cell shading is relative to the proportion of sample observations in a given column who are enrolled in each course, with the modal course being shaded in dark green and also highlighted in bold font. Statistics are calculated using all students who meet the sample criteria prior to bandwidth selection (see Appendix Table A1, column 3), of whom N=28,220 enrolled in the regular track in 6th grade (Panel A); N=16,344 enrolled in the advanced track (Panel B); N=4,185 entered GEM in 6th grade but did not complete the full sequence by taking geometry in 8th grade and earning a B or better (Panel C); and N=5,475 completed the GEM sequence successfully (Panel D). Students who enroll only in remedial math in 6th grade (N=1,861) are not represented in the table.

Appendix Table A2. Sample Criteria & Sample Means

	All in Grade 5	Observed in Grades 4-6	Followed Through HS Grad Year
	(1)	(2)	(3)
Demographic characteristics			
female	0.50	0.50	0.51
FRL	0.43	0.43	0.42
White	0.34	0.34	0.35
Black	0.36	0.36	0.35
Hispanic	0.23	0.23	0.23
Gifted	0.07	0.07	0.07
Baseline achievement			
grade 4 math score (standardized)	0.11	0.11	0.14
grade 4 reading score (standardized)	0.10	0.10	0.13
School peer characteristics (6th grade)			
avg grade 4 math & reading scores	0.05	0.05	0.06
fraction FRL	0.43	0.43	0.43
fraction GEM eligible	0.18	0.18	0.19
Math track in 6th grade			
GEM (treatment)	0.16	0.16	0.17
advanced math	0.28	0.28	0.29
regular math	0.50	0.50	0.51
remedial math only	0.05	0.05	0.03
Number of observations	106,170	82,249	56,003

Notes: The sample in column 1 includes five cohorts of students who entered 5th 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 is limited to students who were enrolled in (or had graduated from) a District high school six years later (2010-2014) -- allowing their college enrollment to be tracked with the District's NSC data -- and for whom math course enrollment is observed for grades 6-8 plus at least one year of high school.

Appendix B: Classification of majors and constructions of main outcome variable

To construct our primary outcome variable, we start by classifying CIP codes associated with the bachelors' degrees in our sample as either STEM or non-STEM using the NSF/NCSES taxonomy at the 2-digit level (NSF, 2023). We then combine 2-digit CIP codes into 18 detailed categories, 6 broader categories, and an indicator for majors that are classified as either STEM or Business. Our main outcome is an indicator for earning a bachelor's degree within 6 years in any STEM or Business-related field.

Appendix Table B1 lists our 18 detailed major categories and their corresponding 2-digit CIP codes, along with their classification into one of six broader categories that we use for subfield analysis and the binary classification on which we base our main outcome variable: STEM/Business or non-STEM.

The indicated (STEM/Business) group in our main outcome includes all 2-digit CIP codes that fall under NSF's definition of STEM plus one 2-digit CIP code (52, Business) that is not considered STEM. It does not include majors that are considered "Science/Engineering related" by NSF – the most populous of which is health (51). Other S/E related majors are multi-disciplinary fields that are very sparsely populated in our sample.

The decision to bundle business (but not health) with STEM in our main outcome variable is motivated by the relative math intensity of these majors and STEM majors. We measure math intensity using O*NET ratings of the math requirements for occupations that each major feeds into. O*NET provides two Mathematics Knowledge ratings: (1) a job's "math level," measuring the degree of complexity and advancement of math knowledge required, and (2) it's "math importance," measuring how crucial mathematical facts and principles for daily job performance. (Appendix Table B2 shows the descriptions of each rating.) For each 6-digit CIP code, and for each math knowledge concept, we construct the mean rating across all 2018 SOC occupations that are linked to the CIP code via the NCES/BLS CIP-SOC crosswalk. (Pre-2020 CIP codes were normalized to CIP-2020 using the NCES vintage crosswalk.)

Figure B.1. plots, for each of the 18 categories in Table B.1, the average math importance score against the average math level score, averaging over all bachelor's degrees in our main analysis sample. The markers are weighted by the frequency of each major group in our sample. On average, business majors in our sample have a higher math importance score and a higher math level score than some of the fields officially classified as STEM. Health, on the other hand, has a lower math level score than any STEM field and its math importance score lies below that of STEM fields except for the social sciences. The dashed red lines formalize our decision rule for constructing our main outcome: all major groups with a math level rating of 4 ("High-Intermediate") or higher and a math importance rating of 3 ("Important") or higher are classified as "STEM/Business." All others are classified as non-STEM.

Appendix Table B.1: Detailed major categories and broader classifications for outcome variables

Binary classification	Classification for subfield analysis	Detailed major category	CIP code(s)
STEM/Business	Math, Engineering, Physics & CS	Engineering	14.xxxx
STEM/Business	Math, Engineering, Physics & CS	Computer science	11.xxxx
STEM/Business	Math, Engineering, Physics & CS	Mathematics & statistics	27.xxxx
STEM/Business	Math, Engineering, Physics & CS	Physical sciences	40.xxxx
STEM/Business	Life sciences	Agriculture	01.xxxx
STEM/Business	Life sciences	Natural resources	03.xxxx
STEM/Business	Life sciences	Biology	26.xxxx
STEM/Business	Business	Business	52.xxxx
STEM/Business	Social sciences	Economics	45.06xx
STEM/Business	Social sciences	Psychology	42.xxxx
STEM/Business	Social sciences	Other social sciences (excl. economics)	45.xxxx (~45.06xx)
Non-STEM	Health	Health professions	51.xxxx
Non-STEM	Other non-STEM	Communication	09.xxxx
Non-STEM	Other non-STEM	Security/protective svcs.	43.xxxx
Non-STEM	Other non-STEM	Visual/performing arts	50.xxxx
Non-STEM	Other non-STEM	Education	13.xxxx
Non-STEM	Other non-STEM	All other non-STEM	all others

Figure B.1. Average O*NET math knowledge scores by 2-digit CIP code for sample with bachelor's degree earned within 6 years of on-time high-school graduation

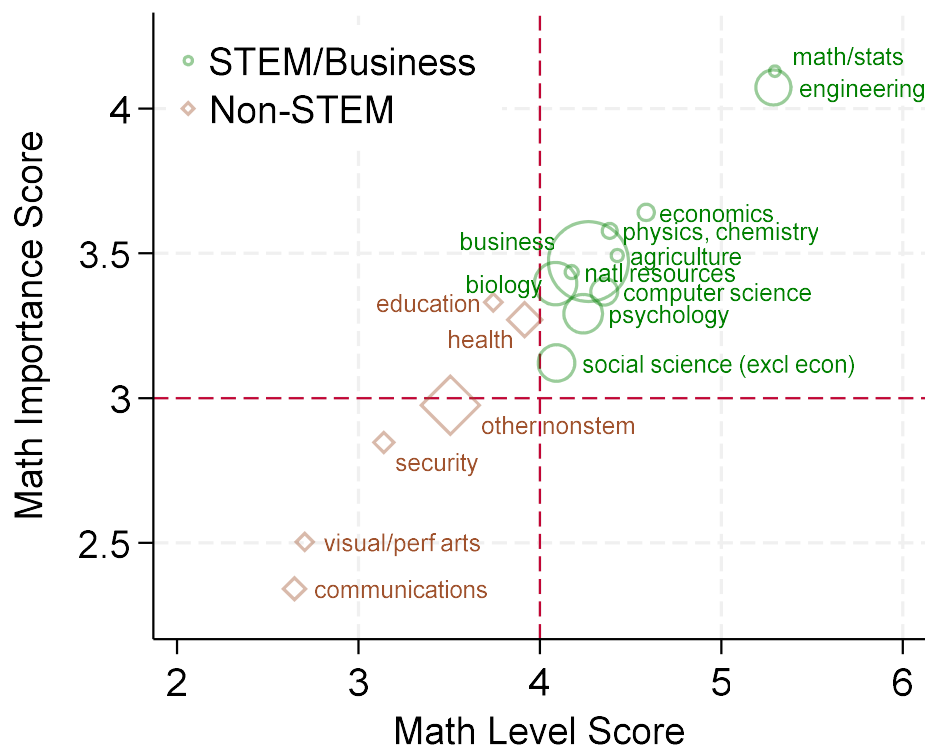


Table B.2. Definitions of O*NET math knowledge ratings

Rating	Short description	Detailed description
A. Math Level Scale		
1	Basic	Required to perform simple, repetitive tasks like adding two numbers or counting inventory.
2	Low-Intermediate	Required to handle routine tracking, balance registers, or perform basic shop arithmetic.
3	Intermediate Level	Required to calculate standard dimensions, compute simple percentages, or maintain basic business logs.
4	High-Intermediate	Required to apply basic statistics, analyze local trend data, or compute structural areas.
5	Advanced	Required to use algebra, geometry, or operational formulas to solve standard technical and financial problems.
6	Very Advanced	Required to apply calculus, differential equations, or complex regression models to system design and predictive forecasting.
7	Expert	Required to derive new mathematical equations, develop cutting-edge statistical theories, or solve abstract operational theorems.
B. Math Importance Scale		
1	Not Important	Not required for job performance.
2	Fairly Important	Occasionally used for minor tasks.
3	Important	A moderate requirement for performing day-to-day duties.
4	Very Important	A significant part of the core responsibilities.
5	Extremely Important	A primary, constant requirement for the job.

Source: <https://www.onetonline.org/help/online/scales>

Appendix C: Distribution of noncognitive skill index and tests to rule out sample selection when stratifying on noncognitive index

The noncognitive index is constructed as the residual from a regression of high school GPA in academic subjects, excluding math, on quadratics in 5th-grade math & reading test score, their interaction, and cohort fixed effects.

In Figure C.1, Panel A shows the distribution of the index separately for girls and boys with vertical lines indicating the gender-specific terciles used to stratify our main RD analysis sample into six subgroups. Panel B illustrates that the gender gap in the noncognitive distribution is fairly stable across values of the 5th-grade math score.

Figure C.1. Distribution of noncognitive index by gender

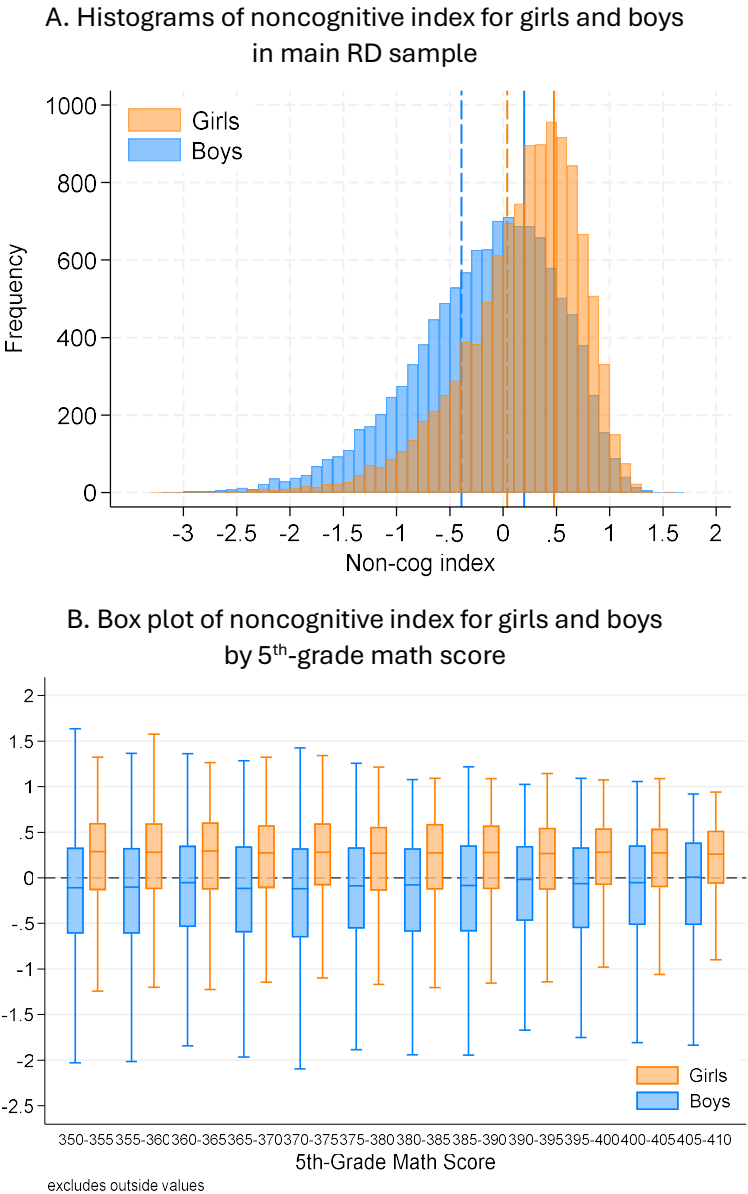
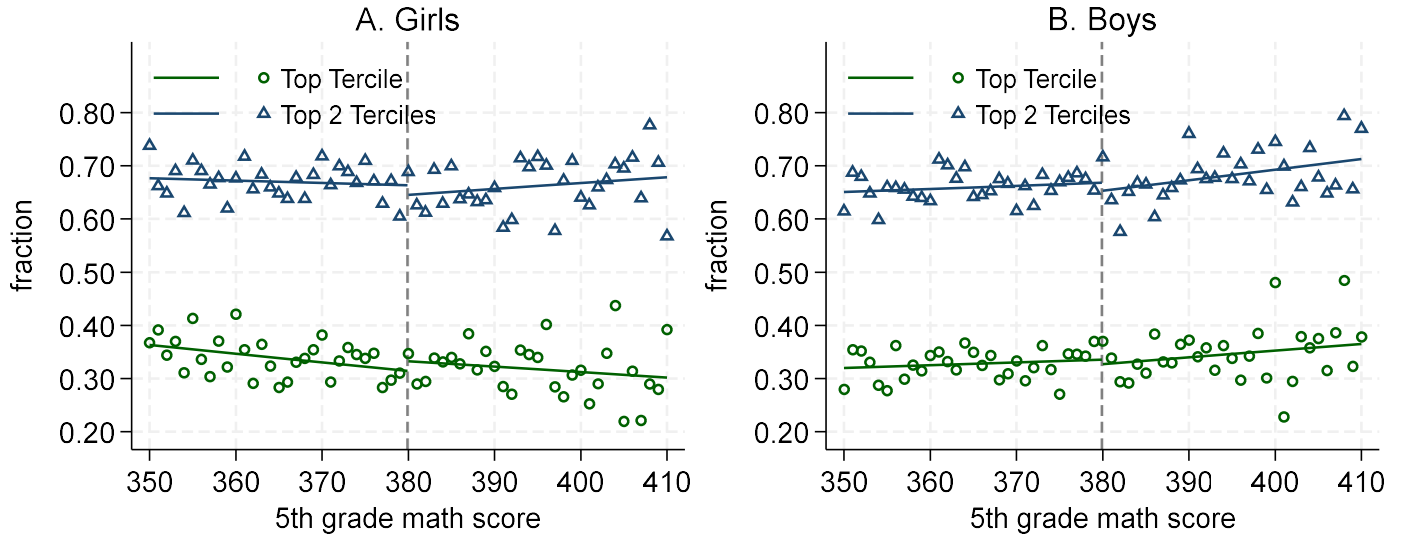
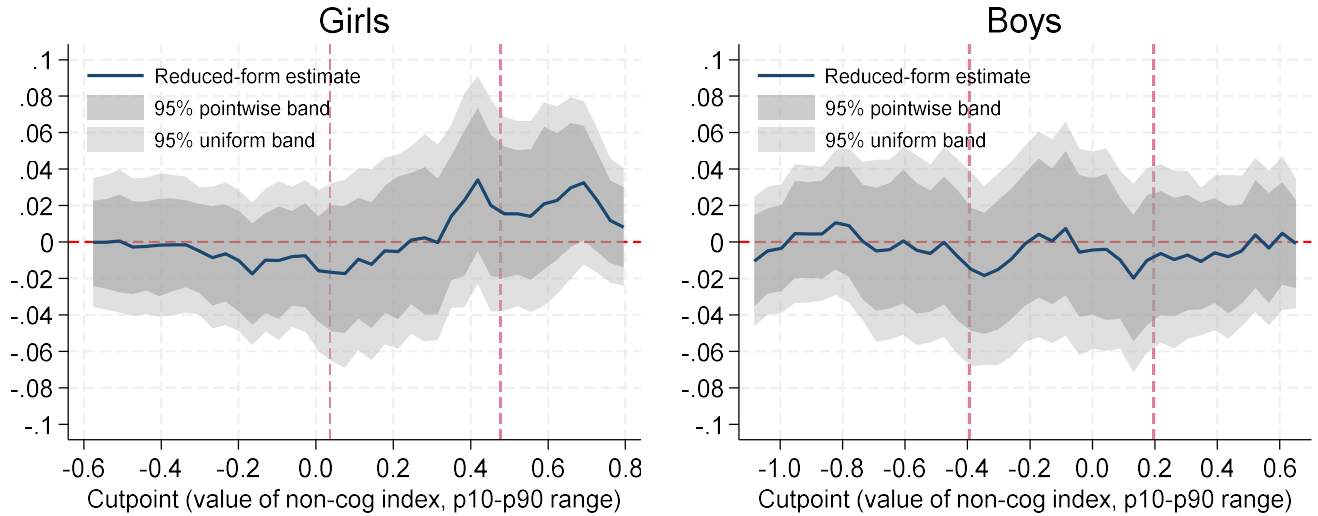


Figure C.2. Reduced-form estimates for the effect of GEM threshold crossing on having a noncognitive score in the top (top 2) gender-specific tercile(s)



Notes: The figure plots means of an indicator for being in the top tercile (dark green circles) or in the top two terciles (dark blue triangles) of the gender-specific distribution of the noncognitive index against 5th-grade math scores, separately for girls (Panel A) and boys (Panel B) in the main RD analysis sample. It also shows fitted values from local linear RD models with a symmetric bandwidth of 30.

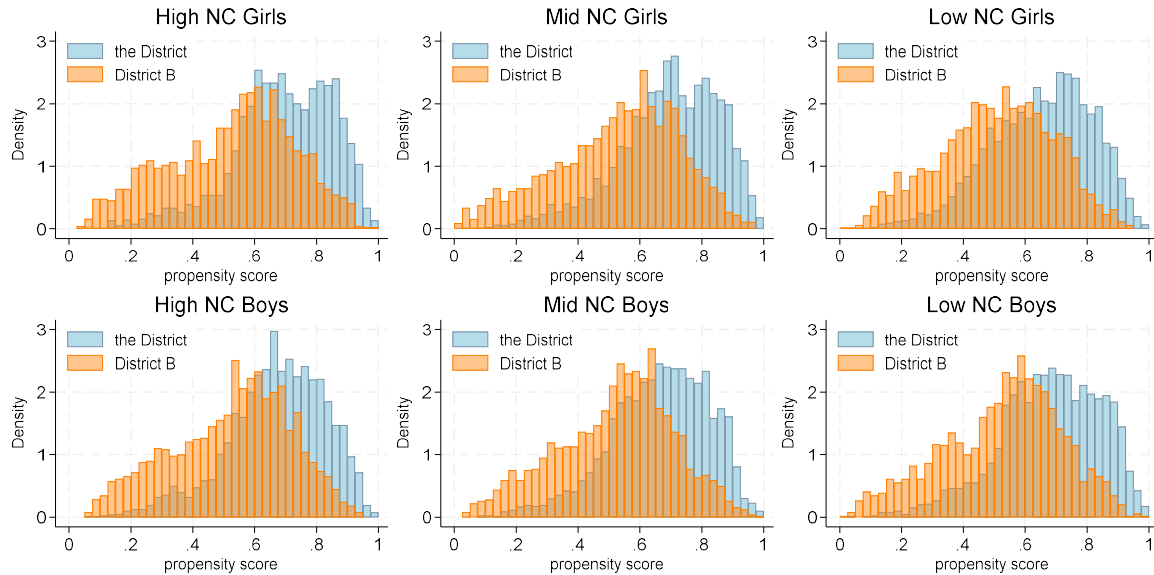
Figure C.3. Reduced-form effect of GEM placement on $\Pr(\text{Noncog} > \text{cutpoint})$



Notes: This figure tests whether GEM assignment shifts the likelihood that a student's non-cognitive skills index exceeds an arbitrary cutpoint, across a range of cutpoints. Each panel sweeps 41 cutpoints spanning the 10th–90th percentile of the noncognitive index within that gender's RD-bandwidth sample (± 30 points). At each cutpoint, the navy line plots a local linear, uniform-kernel reduced-form estimate of GEM assignment on an indicator for exceeding that cutpoint, clustered by school-grade; values above zero mean GEM raises the probability of scoring above the cutpoint. Shaded bands show pointwise (dark) and uniform sup-t (light) 95% confidence intervals from 500 cluster bootstrap replications. Dashed red lines mark that gender's own 33rd and 67th percentile — the tercile cutpoints used in the heterogeneity analysis.

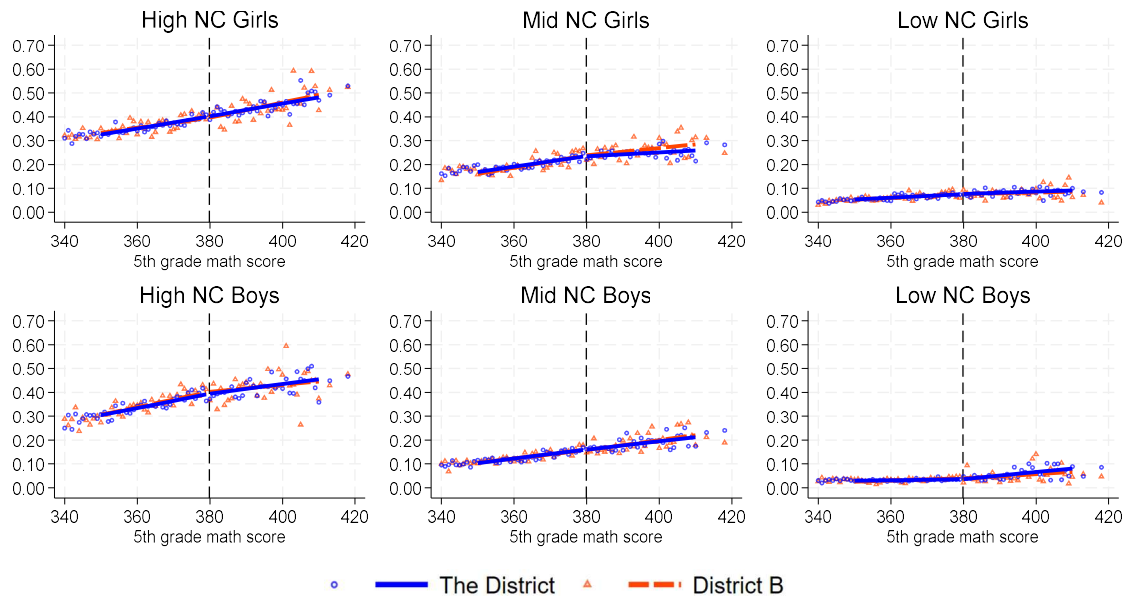
Appendix D: Descriptive Statistics for Between-District Analysis

Figure D1. Propensity score distributions by district and by gender/tercile of noncognitive index



Note: Each panel shows the district-specific distributions of a propensity score for being in the District from a logit model, estimated within 5th-grade math score bins, separately for each gender-noncog subgroup. The model includes quadratics in 4th-grade reading and math scores, student age, and dummies for FRL status, race/ethnicity, English-language learner status, gifted status, and cohort fixed effects.

Figure D2. Predicted probability of earning a STEM/business degree by district and by gender/tercile of noncognitive index



Note: Each panel plots the mean of an index of observables, constructed as the predicted probability of earning a STEM or business degree within 6 years of on-time high school graduation, against 5th-grade math score, by district, for the indicated subgroup. The outcome is constructed from models fit separately for each subgroup to District observations in the RD sample below the GEM threshold (i.e., 5th-grade math scores between 350-420); it includes quadratics in 4th-grade reading and math scores, a quadratic in age, indicators for gender, free lunch status, race and Hispanic ethnicity, English learning status, and gifted status, and cohort dummies (as in Appendix Figure A2). District B observations are reweighted to match the District using inverse probability weights $w = p/(1 - p)$ where p is the propensity score for the District (see Appendix Figure D1 note); observations with $p > 0.9$ are trimmed. Lines are weighted linear fits estimated separately on either side of the GEM cutoff (dashed vertical line), within a ± 30 -point bandwidth.

Appendix Table D1. Sample means for between-district analysis

	The District	District B
	(1)	(2)
A. Baseline characteristics		
Female	0.49	0.49
Free or reduced price lunch (FRL) recipient	0.30	0.31
White (non-Hispanic)	0.44	0.60
Black (non-Hispanic)	0.23	0.10
Hispanic	0.25	0.20
Gifted	0.10	0.28
<i>Fourth grade (baseline) test scores</i>		
Math scale score (standardized at state level)	0.82	0.77
Reading scale score (standardized at state level)	0.54	0.63
B. Mean intermediate outcomes		
<i>Math track in 6th grade</i>		
GEM (treatment)	0.28	0.00
Advanced	0.53	0.96
Regular	0.22	0.05
<i>Middle school course enrollments and grades</i>		
Take algebra 1 by 8th grade	0.56	0.86
Complete algebra 1 with A or B by 8th grade	0.40	0.69
Take algebra I by 7th grade	0.21	0.03
Take geometry in 8th grade	0.18	0.03
6th grade math GPA	3.10	3.35
7th grade math GPA	3.00	3.30
8th grade math GPA	2.88	2.95
<i>High school course enrollments & grades</i>		
Take algebra 2 in 9th grade	0.24	0.10
Number of advanced math courses	1.51	1.94
Take pre-calculus	0.42	0.03
Take any calculus	0.22	0.24
Take any calculus by 11th grade	0.14	0.02
Take any AP/IB calculus	0.15	0.24
Take AP calculus - BC	0.05	0.04
Take AP statistics	0.10	0.22
High school math GPA (all math courses)	2.49	2.58
High school GPA (non-math academic courses)	2.93	2.93
High school GPA (all academic courses)	2.83	2.86
C. Mean post-secondary outcomes		
<i>Transition to college</i>		
Graduate high school on time	0.93	0.89
+ enroll in any college within 1 year	0.74	0.74
+ enroll in selective college within 1 year	0.45	0.44
<i>College degree(s) earned within 6 years</i>		
Any degree	0.54	0.52
Bachelor's (BA or BS) in any field	0.41	0.38
Bachelor's in STEM or Business field	0.24	0.23
<i>Bachelor's degree with major in:</i>		
math, engineering, physical sciences, comp. science	0.046	0.046
life sciences	0.049	0.046
business	0.082	0.074
social sciences	0.068	0.065
health (nonSTEM)	0.038	0.042
nonSTEM (all other)	0.129	0.108
<i>Number of observations</i>	18,775	11,649

Note: Column 1 reports variable means for students in the District in the main RD analysis sample (math scores between 350 and 410) excluding the 2004 cohort. Column 2 reports means for students in the same cohorts in District B who meet the same sample criteria.

Appendix Table D2: Characteristics of overall population, children, and mothers in the District and District B

	Characteristics of Population		Characteristics of Children Age 6-17		Characteristics of Mothers w/ Children Age 6-17	
	The District	District B	The District	District B	The District	District B
	(1)	(2)	(3)	(4)	(5)	(6)
White	64.8	74.8	55.7	68.1	58.6	70.5
Black	25.2	16.4	32.1	21.3	29.9	19.6
Hispanic (any race)	24.3	23.3	27.1	27.4	29.6	25.9
Immigrant	30.7	14.7	12.0	6.0	46.0	22.4
Immigrant Mother	--	--	42.5	20.3	--	--
Has 2 parents in household	--	--	62.7	61.3	--	--
Mother married	--	--	--	--	67.0	66.4
<i>Education (age 24+ only)</i>						
Mean years of education	10.4	10.4	--	--	10.6	10.4
Share with BA	14.3	14.2	--	--	15.2	14.7
Share with grad degree	7.7	7.4	--	--	7.6	7.0
Share BA+ STEM/Bus major	56.8	56.8	--	--	49.1	49.4
<i>Household income (current \$)</i>						
Mean HH income	77,979	74,211	81,700	79,799	84,038	80,820
Share < 30k	23.3	25.0	22.2	25.2	20.0	24.1
Share 30-50k	18.4	20.3	18.3	19.3	18.2	19.2
Share 50-75k	19.8	19.3	19.9	17.4	20.0	17.2
Share 75-110k	17.9	17.2	17.7	17.6	18.8	18.4
Share over 110k	20.6	18.2	21.9	20.6	23.1	21.1
<i>Occupation (earners only)</i>						
Share in science/engineering	2.3	2.9	--	--	2.0	2.7
Share in finance and related	1.9	2.0	--	--	3.2	3.6
<i>Industry (earners only)</i>						
Share in construction	5.4	5.1	--	--	1.6	2.0
Share in manufacturing	3.8	4.2	--	--	4.2	4.2
Share in profession services	9.6	9.2	--	--	11.9	11.1
Share in education	5.5	5.8	--	--	11.9	13.3
Share in health-related ind's	7.3	7.1	--	--	18.7	17.8

Notes: based on tabulations of 2007-2010 American Community Survey microsamples. Sample sizes for overall populations are 65,602 for the District and 49,114 for District B. Household income is calculated by summing incomes of all residents of the household. Mothers are identified by an ACS variable ("PAOC") that indicates if a female resident of a household lives with own children in different age ranges.