EdWorkingPaper No. 24-986

Accelerating Opportunity: The Effects of Instructionally Supported Detracking

Thomas S. Dee
Stanford University

Elizabeth Huffaker
Stanford University

The pivotal role of algebra in the educational trajectories of U.S. students continues to motivate controversial, high-profile policies focused on when students access the course, their classroom peers, and how the course is taught. This random-assignment study examines an innovative district-level reform—the Algebra I Initiative—that placed 9th-grade students deemed well below grade level in Algebra I classes coupled with teacher training instead of a remedial pre-Algebra class. We find that this reform significantly increased grade-11 math achievement (ES = 0.2 SD) without lowering the achievement of classroom peers eligible for conventionally tracked Algebra classes. This initiative also increased attendance, district retention, and overall math credits. These results suggest the impact of higher expectations coupled with aligned teacher supports for the lowest-performing students.

VERSION: June 2024

Suggested citation: Dee, Thomas S, and Elizabeth Huffaker. (2024). Accelerating Opportunity: The Effects of Instructionally Supported Detracking. (EdWorkingPaper: 24-986). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/v492-1p91

High school mathematics attainment has significant consequences for postsecondary and labor-market outcomes (Altonji, 1995; Goodman, 2019; Kim, 2018; Long et al., 2012). Notably, Goodman (2019) estimates than an additional year of high school math completion improves future income by 10% for Black students (but not white) students. This effect is driven by improved access to lucrative STEM careers and implies that greater development of latent human capital in minoritized youth through high-school math coursework is a lever to reduce economic inequality. Yet, enrollment in advanced courses (i.e., those requiring completion of Algebra II) is starkly stratified by race, class, and ethnicity (Ayalon & Gamoran, 2000; Schiller & Hunt, 2011). Black, Hispanic, and poor students complete fewer college preparatory math classes than their white, Asian, and non-poor peers (Conger et al., 2009). Highly fragmented patterns of math enrollment also undermine school-level equity goals by exacerbating within-school ethnoracial segregation (Clotfelter et al., 2021; Dalane & Marcotte, 2020; Davis, 2014; Francis & Darity, 2021). Because stratification is primarily driven by within-school assignment practices (Antonovics et al., 2022; Asim et al., 2019; Betts, 2011; Clotfelter et al., 2021), "tracking"–the practice of sorting students on the basis of perceived ability–has long been critiqued as an inherently unequal method for distributing educational opportunities (e.g., Oakes, 2005).

Policy efforts to reduce these disparities have centered on the foundational high school math course: Algebra I. This attention is motivated by well-established correlations between the accelerated take-up of Algebra with stronger high school math test scores and course progression (Gamoran & Hannigan, 2000; Stein et al., 2011). Towards the end of the 20th century, the Algebra-for-All movement criticized Algebra I assignment practices for creating a bottleneck in student entry to rigorous math classes. Early Algebra I access was recognized as a concern grounded in fundamental fairness and educational "civil rights" (Moses & Cobb Jr., 2001).[1] Conversely, restricting early take-up of Algebra

---

[1] In recent years, some districts have implemented equity-motivated policies that universally *delay* Algebra I access (e.g., Huffaker, Dee, and Novicoff, 2023). However, these are exceptions in the history of Algebra I access policies. Policymakers have predominantly pursued equality through *acceleration*.

was considered "gatekeeping" (Stein et al., 2011; Gamoran & Hannigan, 2000; Oakes et al., 1990). The Algebra-for-All movement dramatically increased the prevalence of acceleration into early Algebra I in the U.S. For example, Chicago Public School eliminated pre-algebra in ninth grade so that all freshmen would enroll in Algebra I or higher (Allensworth et al., 2009). In California, the share of eighth graders enrolled in Algebra I grew from 16% in 1999 to 65% in 2013 (McEachin et al., 2020). Many districts opted to widely – and even universally – accelerate middle schoolers into Algebra I to reduce academic stratification across secondary math courses (i.e., as a method of "*de*tracking").

A substantial body of quasi-experimental evidence generally indicates that acceleration into Algebra I academically benefits well-prepared students but carries negative consequences for lower-performing students (Allensworth et al., 2009; Clotfelter et al., 2015; Dougherty et al., 2015; Heppen et al., 2011; Lafortune, 2018; McEachin et al., 2020). In isolation, these results suggest advantages to tracking students into more or less advanced math courses using baseline achievement measures. Such reasoning is also supported by recent causal analyses that found homogeneous student grouping carries benefits *across* the achievement distribution (Card & Giuliano, 2016; Cohodes, 2020; Collins & Gan, 2013; Cortes & Goodman, 2014; Duflo et al., 2011; Figlio & Page, 2002). A dominant explanation for these findings is the positive impact of efficient instructional targeting (e.g., Duflo et al., 2011). That is, tracking lessens the technical burden of instructional differentiation by reducing within-classroom variation in student preparedness. In a tracked class, "teaching to the middle" approximates "teaching at the speed of learning" for a larger share of the class (Good et al., 1978). Detracking, by contrast, introduces greater diversity in within-classroom student needs and can exacerbate pedagogical challenges for teachers (Rosenbaum, 1999).

However, there are equity-oriented concerns about tracking students into homogeneous classrooms according to baseline achievement. For example, because prior educational opportunity is correlated with ethnoracial and socioeconomic status, tracking necessarily increases within-school segregation (Clotfelter et al., 2021; Conger, 2005; Oakes, 1995). To the extent that such stratification reduces the prevalence of diverse and inclusive classroom environments, it impedes the democratic and social aspirations for schooling (Brighouse et al., 2018; Labaree, 1988). This dynamic can also create powerfully self-reinforcing and inequitable patterns in longer-run student engagement. When students lack role models with shared identities in rigorous courses, they are in turn less inclined to pursue advanced math classes (Francis & Darity, 2021). Similarly, Legette and Kurz-Costes (2021) note that, even after controlling for baseline achievement, students' sense of belonging and motivation in school is negatively associated with assignment to a low-status track. Tracking can also amplify inequity by influencing teacher expectations and effectiveness. A body of qualitative evidence indicates that teachers perceive students in lower-level tracks as having limited capacity for growth and accordingly reduce the rigor and richness of their pedagogy (Gamoran, 1989; Kelly, 2004; Oakes, 2005). Furthermore, students in lower-level classes are more likely to be taught by novice teachers (Kalogrides & Loeb, 2013).

In sum, the literatures on math acceleration and tracking surface vexing tensions for any effort that simultaneously seeks to support both mathematical excellence and broad opportunity. An effective solution would need to avoid the negative consequences of tracking students by prior achievement but still harness the academic benefits of appropriately differentiated instruction for math learners. For example, evidence from Chicago suggests that the targeted provision of additional instructional time (i.e., "double-dose" math), though expensive, protects low-performing students in Algebra classes from the academic harm of acceleration (Nomi & Allensworth, 2009). This partnership study provides evidence on a novel and innovative approach—the Algebra I (A1) Initiative—that bundles the acceleration of lower-performing students into Algebra with a different, classroom-focused strategy: capacity building for high-quality, differentiated instruction.

The A1 Initiative occurred within a diverse, medium-sized school district, which randomized eligible 9th grade students into either the control condition (i.e., their conventional assignment to a remedial pre-Algebra class or to Algebra, based on prior achievement) or to the treatment condition: an A1 Initiative classroom that featured both heterogeneously grouped students and teachers who received unique professional development (e.g., on strategies for instructional differentiation) and additional resources (e.g., more planning time). This novel program, piloted under random assignment, presents a unique opportunity for researchers and practitioners to better understand the promise of instructional improvement as a strategy for promoting both high expectations and inclusivity in math pathways.

The Algebra I Initiative

The Algebra Initiative was a response to persistent math achievement disparities in an ethnoracially and socioeconomically diverse suburban school district in California's Bay Area (i.e., the Sequoia Union High School District). The District serves students across four comprehensive high school and three alternative high schools. Roughly, forty percent of its students are socioeconomically disadvantaged, 13.5% are classified as English learners, 43% identify as Hispanic, 8% as Asian and 39% as White. While the District performed above state averages on English Language Arts and Mathematics assessments in years preceding the Initiative, scores were highly stratified by race, ethnicity, and residential zip code. Specifically, over three quarters of White and Asian students in the graduating classes of 2017 through 2020 met University of California admissions criteria in math, compared with fewer than half of Black and Hispanic students.

The District identified ninth-grade math assignment practices as a potential driver of these disparities, an inference consistent with the evidence on the pivotal role ninth grade plays in shaping student trajectories (Phillips, 2019). Prior to the Initiative, the District used a placement chart (Figure A1) to assign students to freshman math courses using middle-school transcripts and test scores. Incoming Black and Hispanic freshmen were disproportionately likely to be classified as having "below grade-level" proficiency, resulting in ethnoracially stratified math enrollment. The A1 Initiative aimed to redress disparities without reducing achievement for any student group using two key strategies. It combined all treatment-assigned students entering high school at-or-below grade level into Algebra I classes (i.e., acceleration). Critically, the District also provided targeted professional support and development for teachers of these classes to implement new pedagogical approaches.

Initiative pedagogy and resources

The "Initiative Pedagogy" aimed to cultivate effective Algebra I instruction within the specific context of a heterogeneous classroom. This foremost requires teachers engage in a continuous process of differentiated evaluation, reflection and adaptive instruction (Valiande & Koutselini, 2009). From interviews with district and school administrators, we identify three main ways the Initiative promoted mastery of this approach.

First, teachers were trained in specific instructional strategies such as "math language routines" to foster academic conversation and permit frequent assessment of student comprehension (Zwiers et al., 2017). Instructional leaders emphasized the importance of hearing and seeing student reasoning in Initiative classes. Second, to facilitate responsive pacing, teachers were afforded flexibility in executing the Algebra I curriculum (Rosenbaum, 1999). The Initiative cohort collaborated on unit planning and assessments and provided teachers with optional lesson planning resources but otherwise teachers retained autonomy over day-to-day activities.[2] By contrast, Algebra I control-condition

---

[2] Initiative teachers were encouraged to use Illustrative Math (IM), a curriculum and professional leaning open resource to retrieve or inspire their lesson materials. The problem solving emphasis of IM broadly aligns with IES suggested strategies of practice for algebra instruction (What Works Clearinghouse, 2014), but the efficacy IM has itself not yet been independently evaluated.

teachers followed a structured, district-defined pacing guide tied to a single textbook. Third, the Initiative cohort strongly emphasized high expectations and the promotion of a growth mindset for students. Specifically, they encouraged all students to proceed to Geometry and stay on the grade-level track.

To facilitate this approach, teachers received significant resources coordinated by the District as well as support from a math-education consultancy. In addition to approximately 15 full days of professional development, they received an extra planning period, four coaching days per site per semester, a District-wide professional learning community, and a partner teacher at their campus.[3] They also participated in "lesson studies" to share and learn promising practices within the cohort. The Initiative introduced considerable resources and training to promote high quality pedagogy in heterogeneous classrooms. Finally, an important note for the internal validity of our analysis is that teachers in the treatment group were no more experienced than control group teachers and were less likely to have obtained an advanced degree or National Board Certification (Table 1).

Heterogeneous classroom assignment

The second defining feature of the Initiative classrooms was their heterogeneity. Nearly two thirds of the freshman class were eligible to be 'detracked' under the Initiative.[4] Students who entered high school "above grade level", a third of the incoming cohort, were not included.[5] For those randomly assigned to treatment, the Initiative collapsed the number of ninth-grade math pathways from three to one. Students who would otherwise have been assigned, based on prior achievement, to "Algebra Readiness" (i.e., remedial pre-algebra), "Algebra I with Support" (i.e., a "double dose" option) or "Algebra I" (i.e., the standard grade-level track) were combined into A1 Initiative sections.

---

[3] The second semester of intended on-site coaching was interrupted by the COVID-19 pandemic.

[4] Some students with IEPs or who were English learners assigned to Algebra I sections specifically for students deemed to have "Limited English Proficiency" were also not included in the randomization and are therefore excluded from our analytic sample.

[5] Middle school acceleration is administered by independent feeder districts.

As a result, assignment to an Initiative class exposed students to a different mix of peers than assignment to a "business-as-usual" section.

Table 1 illustrates these effects by reporting the results of auxiliary regressions that examine the impact of the intent to treat (ITT)—random assignment to an Initiative classroom—on measures of students' classrooms.[6] "Nearly at grade level" Initiative students experienced modest declines in peer economic disadvantage (i.e., -10 pp) and increases in peer achievement (i.e., 0.25 SD). Classroom composition changes for higher achieving students, the majority share of each class, were not statistically significant. Notably, lower-achieving ("below grade level") students saw a substantial (i.e., 1.28 SD) increase in the baseline math achievement of their peers. Furthermore, the share of their peers classified as poor or English learners declined by 30 percentage points and 25 percentage points, respectively.

We note that these changes in classroom peers, though substantial, likely contribute little to the academic gains we find among "below grade-level" students. The measured impacts of peer achievement on student academic outcomes are generally small (Angrist, 2014; Cohodes, 2020; Cortes & Goodman, 2014; Lefgren, 2004; Sacerdote, 2014). Feld & Zölitz, (2017) estimate that a one standard deviation increase in peer GPA translates, on average, to a 0.0126 improvement in student grades. This implies that assignment to the Initiative would boost "below grade level" academic achievement 0.016 SD solely through a peer-effects channel. However, this is likely an upper bound as Feld & Zölitz (2017) uncover effect heterogeneity suggesting that less skilled students may actually be harmed by a low position in the class distribution.

This finding is consistent with Nomi & Allensworth's (2013) observation that a change in achievement within a classroom can alter assigned course grade to the detriment of lower achievers. Furthermore, the salience of peer achievement depends on interaction quality between students at

---

[6] We describe the data construction, variables, and methods in detail below.

different levels of achievement and students tend to sort themselves by proficiency within classrooms (Kang, 2007; Feld & Zölitz, 2017; Murata, 2013). These dynamics suggest a major role for pedagogical practice (i.e., to facilitate collaborative learning across skill levels) in unlocking the benefits of having high achieving peers for less prepared students.

Features of the control conditions

We also note that, because students were sorted across three groups in the "business-as-usual" control condition, the treatment-control contrasts we study are differentiated by baseline achievement (Table A1).[7] Specifically, students *not* in the Initiative were assigned to ninth grade math course conditional on their eighth grade class (i.e., Common Core Math 8) and their best score from three assessments (i.e., the Smarter Balanced Assessment Consortium (SBAC) California state test taken in 7th grade, and two other diagnostic tests).[8] Although teachers and parents/guardians had influence to "level up" student assignment from the objective placement, fidelity with the placement chart was nearly 90% among the control group in fall 2019 (and our analysis focuses on the intent to treat).

"Below grade level" students *not* assigned to the Initiative were enrolled in a pre-algebra, remedial course called "Algebra Readiness". This course is slower paced and less rigorous than Algebra I. It also does not count towards either district or UC/CSU math requirements. These students cannot enroll in Algebra I until after ninth grade. Assignment to the Initiative allowed earlier access to high school level coursework among this group, which meaningfully changes the structure of future math opportunities for these students relative to the control condition. Control group students "nearly at" grade-level proficiency enrolled in Algebra I and a second (also *not* UC/CSU aligned) block of math

---

[7] We note that in fall 2019 only 79% of control students eligible for Algebra I with a support class ("nearly at grade level") took both of those courses in fall 2019. Control group compliance with assignment to Algebra I ("at grade level") and Algebra Readiness ("below grade level") was 99% and 93%, respectively.

[8] As an example, referring to Figure A1, a student who scored a 2490 on the SBAC ("Nearly Met") and a 6 on the DOMA, but a 20 on the MDTP should have been assigned to Algebra I *without* support, even though they only met the qualifying threshold for that stratum on one of the three tests (the MDTP, in this case).

instruction called “Support”. Students in this category therefore received a “double dose” of math. While “double dose” math has had demonstrably positive impacts on student achievement in other contexts (Cortes & Goodman, 2014; Nomi & Allensworth, 2009), administrators worried the course crowded students out from taking electives. It is also costly to staff. The treatment-control contrast is smallest amongst “at grade level” students, for whom instructional time and potential course credits are identical across conditions.

Data

We use administrative data from the District to examine the effects of the Algebra Initiative on an array of student outcomes. We observe enrollment, transcripts, test scores, and attendance for students in the pilot cohort from ninth grade (AY 2019-2020) through twelfth grade (AY 2022-2023). Twelfth grade data also include an exit indicator for on-time graduation status.

Sample Construction

Our analytic sample includes 1,039 students from a cohort of 2,124 ninth graders who entered any of the District’s four comprehensive high schools in fall 2019. We exclude students enrolled in any of the District’s small, alternative high schools as these campuses did not use random assignment for freshmen course placement. We do, however, follow the high school trajectory of students who initially enrolled at a comprehensive high school and then transferred to a different District site, inclusive of the alternative campuses.

From this cohort, we refine the sample until it is composed only of students included in the randomization. This encompasses nearly all students eligible for either Algebra I or Algebra Readiness (Figure A1). Our ITT population excludes ninth graders enrolled in math classes beyond Algebra I (i.e., Geometry, Algebra II; n=760), students with IEPs enrolled in either no math course or in a basic math skills class (n=31), and those enrolled in Algebra I courses designated for English learners (n=77). Students who took Algebra I in eighth grade but had to repeat it in ninth grade are also excluded (n=29). Two further freshmen were dropped for enrolling late in the fall semester, after randomization had occurred. Finally, we eliminate observations for students whose baseline proficiency under the "business-as-usual" placement scheme (Figure B1) could not be verified. These are students missing baseline data because they did not attend a feeder district in AY 2018-2019 (n=145) or whose eighth-grade math courses are not included in the assignment table (n=41).

Measures

We examine both academic (i.e., assessment and course progression) and nonacademic (i.e., attendance) student outcomes. Math proficiency, our focal outcome, is measured using multiple sources of assessment data. Of these, we privilege the Smarter Balanced Assessment Consortium (SBAC) assessment administered under the California Assessment of Student Performance and Progress (CAASPP) system in spring 2021, during our pilot cohort's 11th grade year. Scale scores are standardized using the published statewide mean and standard deviation (California Department of Eduction, 2023). Secondary test metrics include the fall 2020 and fall 2021 results of a District-created and administered Interim Comprehensive Assessment (ICA).[9] We apply a hybrid Item Response Theory (IRT) model to item-level response data to create a standardized measure of achievement across the analytic sample for each ICA (StataCorp, 2023).

Exploratory outcomes including course enrollment and attainment are parsed from transcript data that denote, for each enrollment, the course name, a teacher identification number, credit type (i.e., math, ELA, elective), credits attempted (one semester of a core course is equivalent to 5 credits), and any credits or letter grade received. Due to the disruption of the COVID-19 pandemic, grading was on a "credit" or "no credit" basis in spring 2020. Summative measures are constructed from

---

[9] The 10th grade ICA score was the outcome we pre-registered when we only expected to follow this cohort for two years. However, the ICA included only 11 closed-response items (i.e., multiple choice, true or false, or numeric response), limiting its construct validity. We prioritize SBAC scores due to its validated psychometric properties and known state-wide distribution (CDE, 2023).

longitudinal ninth through 12th grade transcript data. These variables include binary indicators of whether a student ever enrolled in and/or earned credit in a course (i.e., Algebra II), and whether a student achieved math subject eligibility for entry to the University of California system.[10] Additional details on course classification are in the data appendix. We also calculate total credit attainment – by subject and overall – through 12th grade. Finally, grade-by-grade indicators of course enrollment and completion (i.e., earning at least 10 units, equivalent to two semesters of credit) are used to compare the *pace* of course progression across the treatment and control groups.

We use annual indicators of district enrollment and attendance to measure student engagement. Absence rates are constructed by dividing days reported absent by total enrollment days in each year. We also code a binary indicator for chronic absenteeism that takes on a value of one if a student's absent rate exceeds 10%. Chronic absenteeism is less sensitive to distortion from outliers than absence rate and is used by California as a primary indicator of student connectedness. However, impacts on the continuous absence measure can be more precisely estimated. We therefore present results using both measures. Enrollment files describe student status at the beginning and end of each academic year and denote timing plus explanatory codes for entry or exit events (see data appendix for more detail on treatment of student exits). These data allow us to explore student persistence in the District. District administrators expressed high confidence in the fidelity of attendance and enrollment data, citing extra care taken during the COVID-19 pandemic to maintain accurate records.

Sample Description

Table A2 reports summary statistics for the ITT sample of 1,039 students upon ninth grade entry. These administrative categories are not necessarily reflective of the full diversity of student identities. All students are listed under a single racial/ethnic category, which does not include a

---

[10] In the simple case, this means a student has passed Algebra I, Geometry and Algebra II with at least a "C-" course grade, although our variable construction accounts for the full range of alternative course validations sanctioned by the University of California system.

multiracial indicator. In fall 2019 about half of the students in our sample were eligible for free-or-reduced-price meals (FRPM) and 16% were classified as being English learners (EL). A slight majority are identified as Hispanic, nearly a third as White, 12% as Asian or Pacific Islander (API), four percent as Black, and one percent as Native American or Alaska Naive. Eligible students are dispersed relatively evenly across the four comprehensive high schools in the district and exactly half were randomly assigned to the Initiative. Following Figure A1, 64% of the sample entered ninth grade "at grade level" for math, while 20% were "nearly at grade level" and 16% were "below grade level".

Turning to the key dependent variables, students in the analytic sample scored 0.12 SD above the statewide mean of the 11th grade math Smarter Balanced assessment. While most passed their ninth-grade math class and received credit for two Algebra I semesters by the end of tenth grade, only 62% of students completed a full year of Geometry by the end of AY 2020-21. By the end of 11th grade, 79% of students had received two semesters of Geometry credit and just over half the sample had completed Algebra II. This increased to just over 60% after first-time 12th grade attendance. Similarly, among the four-year sample including dropouts, half of the remaining students completed at least one advanced math course and 62% met the UC/CSU "C" admission requirement for math. Students earn an average of 33 credits (i.e., they complete just over three courses) during high school. Of the four-year sample, 90% graduated from high school within that time span. Finally, we observe a decline in attendance related outcomes over the high school years. Chronic absenteeism rates were 14% in AY 2020-21, 25% in AY 2021-22, and 29% in AY 2022-23, compared with only four percent in AY 2019-20. The underlying absence rate increased from three to five to nearly ten percent across as students progressed from ninth through 12th grade.

Estimation Strategy

Causal effects of the Initiative are estimated by leveraging the random assignment of eligible students into treatment. The strong internal validity of our causal claim is derived from both institutional information and empirical evidence consistent with successful randomization. From interviews, we know randomization was conducted by campus-level administrators during the course assignment process preceding the fall 2019 academic term. Students at three of four campuses were individually randomized between the Initiative and the "business-as-usual" condition. They took varied approaches to apportion the "mix" of students from each placement-level stratum. For example, one school stopped assigning students with "below grade-level" achievement to Initiative sections after reaching a 15% threshold. At other campuses, the baseline achievement composition of Initiative classes matched the ITT student distribution. At the only campus that did not randomize at the student level, an administrator used a two-stage process. First, "business-as-usual" Algebra I and Algebra Readiness sections were creating using the District's standard procedure to balances classes by demographic traits. Second, a randomly selected half of these sections were dissolved, and then recombined as Initiative classes.

Empirically, randomization is assessed through inspection of the balance in pre-treatment student traits across conditions conditional on campus membership and baseline achievement group. Table A3 presents the results of these tests for individual characteristics, and jointly using seemingly unrelated regression (SUR). There is no evidence of systematic imbalance. We therefore turn to our preferred specification for estimating ITT Initiative effects by baseline achievement group:

$$y_{ics} = \beta_1(ITT * atgradelevel)_i + \beta_2(ITT * nearly)_i + \beta_3(ITT * below)_i + \beta_4 nearly_i + \beta_5 below_i + \alpha_s + \varepsilon_{ics} \qquad (1)$$

Where y_{ics} is an outcome for student *i* who, in ninth grade, was in class *c* at school *s*. The parameters, $\beta_1, \beta_2,$ and β_3, estimate the impact of assignment to the Initiative for each achievement group. To guide interpretation of the estimand, we note that, while $\beta_1, \beta_2,$ and β_3 technically represent the *intent-to-treat* effects of assignment to the Initiative, compliance with this assignment was high. By the end of the fall 2019 semester, 99% remained in their assigned condition, and 97% remained through spring 2020. All regressions include indicator variables for ninth-grade campus membership and baseline placement level, with "below grade level" students in the control condition being the reference category. Standard errors are clustered by class section in our preferred model.

This error selection attends to evolving guidance from recent conceptual and econometric observations on the specification of standard errors for causal inference. Influential work by Abadie et al., (2023) encourages researchers to consider a design-based rationale for error selection. In the case of clustered sampling and randomization, heterogeneity-robust standard errors can be too small, but in other cases (i.e., when random sampling and assignment occur predominantly at the unit level) clustering is overly conservative. In this case, randomization was largely the unit-level in a procedural sense, but treatment status is almost perfectly correlated within classrooms. Therefore, we cluster standard errors by the ninth-grade classroom membership in our main analysis.

In additional analysis, we investigate the sensitivity of our results to alternative empirical specifications and across teacher and student-defined subgroups. This includes exploring the implications of alternative standard error specifications on our main test-score outcome in several ways. First, we present less conservative heterogeneity-robust Eicker-Huber-White standard errors. Second, we adjust standard errors for the relatively small number of class sections in our sample, following recommendations from Pustejovsky & Tipton (2018) for N<50 clusters. For each of $\beta_1, \beta_2,$ and β_3 , effective counts of clusters range from 38 to 47.

Third, we follow the encouragement of econometricians (e.g., as in Abadie et al., 2020) to consider an alternative to frequentist standard errors. Like many randomized control trials (RCTs) we use a convenience sample and observe all relevant units. Under the conventional framework, standard errors confer uncertainty in our estimated parameters relative to the true values that would exist in a hypothetical "super-population". However, the source of uncertainty in our estimates comes not from random drawing, but from the random assignment of treatment itself. Standard errors are then derived by simulating permutations of the treatment indicator. We use "ri_test" (Heß, 2017) to execute the randomization inference procedure.

We also report ITT effects on our main outcomes after the addition of controls for student-level and teacher-level characteristics. In Equation 2 $\boldsymbol{X_{ics}}$ is a vector of student traits which includes race/ethnicity, gender, EL, FRPM status and sometimes a middle school test score. $\boldsymbol{T_{cs}}$ includes indicators of two observable teacher characteristics: years of experience, and advanced certification status (i.e., a postgraduate degree or National Board Certification). Equation 1 results are privileged in our main analysis over Equation 2 estimates due to superior performance under Akaike information criterion (*AIC*) cross-validation for most outcomes.

$$y_{ics} = \beta_1(ITT * atgradelevel)_i + \beta_2(ITT * nearly)_i + \beta_3(ITT * below)_i + \beta_4 nearly_i + \beta_5 below_i + \gamma\boldsymbol{X_{ics}} + \delta\boldsymbol{T_{cs}} + \alpha_s + \varepsilon_{ics} \qquad (2)$$

The inclusion of $\boldsymbol{T_{cs}}$ in some specifications is one component of our strategy to address an implementation detail that warrants further discussion. That is, because *most* of the Initiative teachers volunteered to join the pilot, it is possible our analysis is detecting the efficacy of the eight Initiative teachers rather than the impact of the program itself. It is plausible that these eight are particularly enthusiastic and committed instructors. As discussed, we control for observable teacher traits that correlate with teacher quality (e.g., Harris & Sass, 2011) in the specification of Equation 2. Additionally, we test the sensitivity of our findings to a "leave-one-out" exercise. Under this test, we repeat our main analysis eight times, in-turn sub-setting the data to exclude students enrolled in each teacher's ninth grade class. This check allows us to observe whether any singular teacher is driving our results. We further discuss interpretation of our findings in light of this concern in subsequent sections.

Results

Tables 2 through 4 present our preferred estimates from Equation 1 across a rich collection of exploratory (i.e., course progression and student engagement) and confirmatory (i.e., test scores) student outcomes.

Exploratory Analysis of Initiative Mechanisms

Student Engagement

Attendance is an important determinant of student achievement (e.g., Lamdin, 1996; Liu & Loeb, 2021). Table 2 summarizes our findings on the impacts of assignment to the Initiative on student engagement with school. Across all four years, assignment to the Initiative reduced absenteeism for students who entered high school below grade level. Their absence rate was between two and seven percent lower than for control group students assigned to Algebra Readiness. This translated to a five to nine percent reduction in chronic absenteeism, although this is only measured precisely in ninth grade. Attendance is largely unaffected for higher achieving students. Taken together, these findings suggest that features of remedial pathways (i.e., Algebra Readiness) such as isolation with a homogeneous, low-achieving peer group, stigma, and weakened academic expectations reduce student engagement with school.

Additionally, for the highest and lowest achieving students at baseline, the Initiative improved their chances of remaining in the district rather than transferring to a different school.[11] For the "below grade level" group, assignment to Initiative improved their likelihood of remaining in the district for all four years of high school by 13 percentage points from a control-group base rate of 77%. This is a substantive outcome consistent with the Initiative generating heightened levels of belonging and satisfaction for an academically vulnerable population. While full explication of the mechanism

---

[11] We do not detect any evidence of systematic differences in dropout rates across the treatment conditions.

underlying this result is beyond our study's scope, it stands to reason that parents may be less inclined to remove a child from their current school if the student is content and succeeding in their educational environment. Given the negative consequences of reactive moves (e.g., Welsh, 2017), this is another academically protective feature of the Initiative. Interestingly, we see in Column 10 that students who enter the district "at grade level" are about six percentage points more likely to remain there through twelfth grade if assigned to the Initiative. It is possible their improved *relative* classroom position boosted satisfaction and belonging.

These results carry notable implications for interpreting effects on dependent variables measured in 11th and 12th grade. Specifically, downstream outcomes are likely biased by the differential rate of district exit (i.e., attrition). We explore the likely direction of this bias conceptually and empirically. First, we hypothesize that Initiative students who would be marginal out-transfers in the unobserved counterfactual are likely to have lower levels of expected achievement. This is because mobility is negatively correlated with achievement (e.g., Mehana & Reynolds, 2004). The disproportionate persistence of these students in our sample's treatment arm exerts a downward pressure on the average outcomes of "below grade level" Initiative student relative to the observed control group. This reasoning is supported by the observed baseline achievement of attriters – the average control group attriter scored worse on middle school math assessments than the average treatment group attriter. We do not observe pre-treatment differences in Initiative and non-Initiative attriters across demographic dimensions. Therefore, the estimates we report for the impact of Initiative assignment on downstream outcomes (e.g., 11th and 12th grade course-taking) likely reflect a *lower* bound on the true ITT effects.

Course Progression

In Figure 1 and Table 3 we summarize ITT effects on student course progression, attainment, and on-time graduation. First, we use a Sankey graph (Figure 1) to illustrate the role assignment to the

Initiative plays in shaping math trajectories for students achieving "below grade level" at baseline.[12] Considering the ninth to 10th grade transition depicted in Figure 1, it is evident that placement in Algebra I rather than Algebra Readiness was challenging for many of these students. Ninth grade math failure rates were greater in this stratum for Initiative students. Approximately half had to re-take Algebra I or enroll in a pre-Geometry bridge course as sophomores. However, the other half continued the college-preparatory grade level pathway to Geometry in 10th grade, while only one Readiness student did so. And by the end of 12th grade, academically underprepared Initiative students were more likely to pass Algebra II than their control group peers (Table 3, Column 5). So, while the pipeline observed in Figure 1 is "leaky", students who are held back in their progression after 9th grade do no worse than students whose progression is delayed *before* 9th grade. In other words, track stability – the propensity for sustained placement on a particular pathway over time (Domina et al, 2017) – is very high for students on the remedial track. Once students are assigned a ninth-grade remedial course, we observe that acceleration onto the "standard" pathway is very rare.

Table 3 supplements the visual patterns in Figure 1 with estimates from Equation 1 for a series of transcript outcomes. Specifically, it shows that by the end of tenth grade, "below grade level" Initiative-assigned students were no less likely to have passed two semesters of Algebra I when compared with the control group. Furthermore, they were 22 percentage points more likely to have passed two semesters of Geometry. This accelerated progression continues through eleventh grade, when "below grade level" students assigned to the Initiative are 14 percentage points more likely to have earned Algebra II credit than their peers in the control group. The 11-percentage point magnitude of the analogous grade-12 estimate reflects a *more than doubling* of the likelihood of completing Algebra II. Twenty-one percent of "below grade level" students assigned to treatment earned full Algebra II course credit compared with only nine percent of comparable control students. Treated students in

---

[12] Similar figures for the other strata are not included because of the null impacts observed in Table 3.

this group also obtained more math credits (Column 8) but were no more likely than their Algebra Readiness-assigned peers to complete an optional "fourth" (i.e., post-Algebra II) math course. The collection of results captured by Figure 1 and Table 3 underscore how critical a student's ninth grade placement in the hierarchical math sequence is in structuring opportunity for the entirety of their high school career.

Turning to the other baseline-achievement groups, those "at grade level" achieve comparably regardless of their treatment status. We observe a slight divergence in medium-term course progression and graduation within the "nearly at grade level" stratum. Specifically, these students are less likely to have completed two semesters of Geometry by the end of tenth grade if assigned to the Initiative, despite passing Algebra I at equivalent rates to control group students. We rule out baseline imbalance or differential attrition as explanations per the results in Tables 2 and A2. Notably, control students from this stratum qualify for "double dose" math, so treated students receive less math instructional time. Given the sturdy evidence base in support of "double dose" math for borderline proficient students (e.g., Allensworth et al., 2009), its removal may explain these results. However, by the end of 11th and 12th grade, there is no difference by treatment assignment in cumulative math attainment for students with "nearly at grade level" middle school achievement.

Confirmatory Analysis of Test Scores

Our main, pre-registered confirmatory analysis aims to measure the effect of Initiative assignment on student math achievement as directly as possible. Specifically, we use assessment outcomes as the main metrics for student proficiency (Table 5). Otherwise, course outcomes alone are imprecise proxies for cognitive achievement. This is because course content may vary across similarly titled classes. And, depending on the grading autonomy and strategies afforded to teachers, credit attainment may not indicate content mastery. That is, absent other measures of proficiency, the superior course-taking outcomes we observe *could* reflect social promotion rather than learning.

Despite this concern, for students entering high school with "below grade level" math proficiency, assignment to the Initiative yields large and positive test score effects. On the 11th grade SBAC – the psychometrically validated exam for high school math proficiency in California – we detect a substantial and precisely estimated +0.19 SD (Table 5, Column 3) to +0.20 SD (Table 5, Column 4) ITT impact for these students using state-normed scores. We do note that, because this study was originally intended to run through the pilot cohort's tenth grade year, our pre-registration plan focused on an 11-item District-constructed assessment (i.e., the ICA) taken in the fall following the pilot year. However, project delays related to the COVID-19 pandemic implied that the superior grade-11 assessments, which were instead intended for a separate follow-up study, became available to us. We note that, for the same group of underprepared students, the relevant ITT estimate using the 11-item grade-10 assessment was nearly as large as the SBAC-based effect (i.e., +0.14 SD), though not statistically significant.

Following the spirit of our pre-registration, in Table 5, we also present Romano-Wolf p-values that implement a multiple-comparison correction for the three confirmatory estimates on the test-score outcomes. After this correction the intent-to-treat SBAC effect for the "below grade level" group takes on a p-value of 0.0589 conditional on student traits, within the threshold for marginal statistical significance. In sum, the results presented in Table 4 provide strong evidence that the Initiative improved math learning for low proficiency students, as well as – and possibly because of – superior course attainment and school attendance.

The sizeable magnitudes of these estimates can be contextualized in several ways. Kraft (2020) deems 0.2 SD to be a "large" effect in the distribution of estimates across 700 education RCTs. Given a normal distribution of scores, this implies a seven percentile-point effect, placing it in the most effective third of educational interventions (von Hippel, 2024). While a math-specific meta-analysis of randomized interventions by Williams et al., (2022) finds an average effect size of 0.24 SD for Algebra-related interventions, the authors note that available estimates are based on outcomes measured soon after the intervention. So, the persistence of the effect we detect from a ninth-grade Initiative on a grade-11 assessment is notable. This impact is also large relative to learning trajectories at this age. Specifically, Bloom et al. (2008) find that a 0.19 SD effect across grades nine to 11 roughly translates to an entire additional year of math learning. Finally, we reiterate that, because of the asymmetric rate of District exit observed across the treatment and control conditions, our estimate is plausibly a *lower* bound on the true impact of the Initiative on math achievement.

Additional Analysis

Effect Heterogeneity

Table A2 presents results for exploratory analyses of Initiative assignment effects by student gender and socioeconomic (i.e., as proxied by FRPM) status. Given small sample sizes – which preclude further sub-setting across racial and ethnic categories – and the type I error risk associated with estimating so many effects, we consider these results to be merely suggestive of directions for future research. Still, we note a few intriguing patterns. First, estimates of the positive influence of the Initiative on attendance and district retention for "below grade level" students are consistently larger when the sample is subset to only include girls, versus when effects are estimated among boys. Prior research *has* identified girls as more sensitive to negative behavioral peer effects than boys (Imberman et al., 2012). This result is consistent with a symmetrical propensity for positive peer influence. Conversely, the test-score benefits of the Initiative for "below grade level" students are concentrated among boys, as well as poorer students. It is possible that acceleration may benefit students through different mechanisms, and this heterogeneity warrants future investigation.

Impacts on Other Academic Outcomes

There is only limited evidence of spillover effects into performance in other subjects. At the 10 percent level we detect negative impacts from assignment to the Initiative on credit accumulation for “nearly at grade level” students in science and elective courses, as well as overall (Table A9, Columns 4 and 5). Their inferior performance in science constitutes more evidence that such students were impacted by the removal of “double dose” math for Initiative students. The deficit in elective credit is likely mechanical and the result of their non-enrollment in the elective-credit-bearing “support” course. Similarly, the negative impact on elective credit attainment for “below grade level” students can be explained by their acceleration *out* of the pre-algebra “Algebra Readiness” elective credit course.

Sensitivity & Robustness Checks

Our main findings, especially those we highlight for the “below grade level” proficiency group are robust to a variety of alternative specifications. Table A6 present the results of Equation (2) where student (Panels B and C) and teacher characteristics (Panel D) are controlled for. As we would expect given successful randomization, the magnitudes of the ITT estimates do no differ substantially across specifications. Table A7 similarly show that our findings are unlikely to be driven by a “superstar” teacher (or its opposite), as dropping each Initiative teacher from our sample in-turn leaves our main findings broadly unaltered. And in Table A4, the standardized test-score effect for our focal student group remains large and at least marginally statistically significant using alternate standard errors, except for when randomization inference is combined with our non-preferred strategy of including a vector of baseline student characteristics.

Conclusions

In contrast to much of the extant research on the acceleration of very low-proficiency students (e.g., Clotfelter et al., 2015; Domina et al., 2019; Lafortune, 2018; McEachin et al., 2020), this study identifies academic and nonacademic benefits from a program that randomly assigned academically underprepared ninth graders into Algebra I rather than a remedial pre-Algebra course. Given established links between high school math attainment and postsecondary as well labor market attainment (e.g., Altonji, 1995; Goodman, 2019), these findings carry significant implications for these students' long-run outcomes. Our results indicate that the Initiative positively shaped the trajectories of students who were the lowest achievers in middle school. On the state high school math assessment, we detect a substantial 0.19 effect size. Furthermore, the Initiative improved math course attainment and credit accumulation, as well as student attendance throughout high school. And, because District retention of treatment-assigned students was 13 percentage points higher than of control students, our estimates likely reflect a lower bound on the true impacts of the Initiative. We find no evidence that students entering high school "at grade level" were negatively influenced by assignment to the Initiative.

Given this strikingly positive collection of results for students deemed to have low levels of proficiency, several implementation and policy details of the Initiative deserve particular attention. The Initiative provided support for the development (i.e., dedicated professional learning) and execution (i.e., additional planning time) of an appropriate instructional approach for a mixed-achievement classroom environment. For example, teachers were provided flexibility to responsively pace their courses, strategies to help surface student misconceptions (e.g., math language routines), and community support including a partner teacher and on-site coaching. Initiative trainings also strongly emphasized that teachers hold high expectations for *all* students to continue progressing through a college preparatory math sequence.

Our study of the Initiative yields important implications for the evidence base on the student-level effects of acceleration and detracking. First, pedagogical quality should be a central implementation concern for policies that expand access to rigorous content for less prepared students and/or increase within-classroom variation in baseline achievement. The findings of this study are consistent a protective dynamic wherein supportive practices for high quality instructional differentiation offset the academic risks of acceleration. Second, the superior engagement, as measured by attendance and District retention, of accelerated Initiative students relative to control students indicates that remedial pathways diminish students' sense of belonging in school, possibly due to the isolation, demotivation, and stigma. Third, the persistence of impacts in eleventh and twelfth grade underscore the status of ninth grade as a "make or break year" for the rest of high school (Phillips, 2019). Most empirical studies of math acceleration and tracking focus on practices in middle school grades (e.g., eighth grade Algebra I). However, the patterns we observe in this high-school based study suggest that ninth grade acceleration or remediation play a significant role in how math educational opportunities are structured in a time-limited setting (i.e., generally within four years to complete a hierarchical sequence). This study demonstrates that high school is not "too late" for interventions to positively transform academic trajectories.

Furthermore, while our analysis focuses on student-level outcomes, the Initiative did advance the District's broader equity and inclusion goals. Ethnoracial and socioeconomic diversity was greater in Initiative sections relative to remedial sections where poor and minoritized students were disproportionately concentrated (see Table 1). We can quantify the differences in Initiative versus control classroom-level segregation using a dissimilarity index. This is a measure of the evenness in distribution of students and can be interpreted as the share of students within each treatment arm that would have to switch classes for all sections to include a balanced proportion of the indicated student groups. A higher ratio indicates more severe between-class segregation. The dissimilarity index between poor (i.e., qualifying for FRPM) and non-poor students among Initiative sections is .212 compared with .479 for control sections. The same measures comparing segregation of EL and non-EL students are .243 and .580. Ethnoracial segregation is also lower in the detracked condition, with the distribution of white to not-white, white to Hispanic and Hispanic to not-Hispanic students being 1.7 to 1.9 times more uneven across control classes versus across Initiative classes.

Another notable issue is the comparative cost effectiveness of the reform. One sensible benchmark is to use the average salary of certified District teachers and the number of teacher planning periods funded to estimate the total cost of the Initiative at nearly $173,000. Given the comparatively small number of "below grade level" students assigned to Initiative classes (i.e., the only students with clear academic benefits), the implied cost per relevant student would then be nearly $2,600. The estimated 0.19 SD test-score gain among these students implies the Initiative generated nearly 0.08 SD in test-score gains per $1,000 spent. Notably, this far exceeds the return on general increases in spending. Specifically, Jackson & Mackevicius (2024) conclude that a $1000 increase in annual spending per-pupil repeated over four years increases test scores by only 0.032 SD (i.e., four times the cost for less than half the gain in learning). This comparison ignores other costs of the Initiative (e.g., an instructional coach and consulting fees). However, overall, it is likely to understate the cost of the Initiative because these classes were somewhat larger and required less supplementary staff (i.e., coteaching and support sections). In general, these results suggest the Initiative, which generated quite large gains targeted among a uniquely important subgroup of students is comparatively cost-effective.

Still, the promising academic, non-academic, and inclusion benefits we observe should be contextualized against the limitations of generalizing from this study. For example, policies that promote detracking often face pushback from some parents who oppose the removal of selective tracks (Kariya & Rosenbaum, 1999; Tucker, 2023; Wells & Serna, 2010). In this respect, the Initiative was a *moderate* policy because the highest-achieving third of the student body was unaffected by the reform. Additionally, we note that our analysis did not bear out concerns that that the addition of "below grade level" students to Algebra I classrooms would induce negative peer effects. A "partial" detracking model typified by the Initiative may engender greater buy-in than more aggressive reforms. These findings also may not generalize to populations with different distributions of baseline achievement. In the District, positive effects were concentrated among a small – and uniquely high-need – group of students. Similar policies *could* produce even larger average gains in Districts where a larger share of students would otherwise be placed in remediation. Conversely, it is possible that in classes with large shares of the lowest-proficiency students (i.e., a reverse of the composition observed in Table 1), high achievers would experience negative peer impacts and low achievers would not see benefits from having more proficient peers. In sum, our findings therefore do not necessarily generalize to policies that would *fully* detrack 9th grade math, such as by removing middle school acceleration pathways.[13] Such an expansion may risk the "collective effects" identified by Penner et al (2015): the broader the scope of a detracking program the greater the risk created for unintended consequences like political backlash, negative peer effects, and burdensome staffing and pedagogical demands.

A final category of scalability concerns is the feasibility of high-fidelity implementation in other contexts. That is, whether other districts can support and nurture reforms in the manner of districts that pioneered those reforms is uncertain. Interventions like the Initiative that rely on substantial shifts in within-classroom practice are challenging to promote successfully at scale (Elmore, 2010). Additionally, while our results are robust to multiple checks for teacher-selection effects (i.e., conditioning on observable traits, a "leave-one-out" exercise) it is possible the Initiative would be less effective if teachers were mandated to adopt it against their preferences. Optimistically, however, the program improved outcomes for low proficiency students *despite* the COVID-19 pandemic interrupting a planned program of teacher support in spring 2020. Sub-ideal implementation still generated outcomes above-and-beyond those of the status quo.

Our findings do highlight two specific areas for improvement of the Initiative program. First, in twelfth grade we observe a partial fadeout of the Initiative's relative influence on Algebra II

---

[13] Recent backlash to a broad range of equity-motivated policies, such as detracking in ELA courses, suggest a more comprehensive detracking Initiative would have drawn greater community pushback.

completion – and null impacts for coursework beyond Algebra II. Students who sit out senior math lose the "subject continuity" which may contribute to a successful college transition (Wainstein et al., 2023). So, there may be a role for proactively guiding students to enroll in a grade-12 math class above and beyond graduation requirements. Second, evidence of small negative impacts from assignment to the Initiative on academic attainment for students who would have received a "Support" class in the control group suggests the Initiative could be more effective for some students if paired with additional instructional time. However, staffing demands for co-implementation of the Initiative *and* "double dose" math could be prohibitive in some districts.

Overall, this study shows that "raising the floor" of academic expectations for even very low-proficiency students can be a successful strategy to link equity goals with improved achievement. Critically, the Initiative paired reforms to course assignment with aligned supports for students and teachers. The positive impacts of the Initiative on student engagement suggest that any academic benefits of tracking likely come at the cost of heightened stigma and isolation for remedial-track students. Furthermore, even the partial detracking of the Initiative decreased within-school segregation by race, ethnicity, and class. The Initiative, therefore, presents a provocative proof point for high-school math classes in which students with disparate levels of prior math achievement excel together.

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2020). Sampling-Based versus Design-Based Uncertainty in Regression Analysis. *Econometrica*, *88*(1), 265–296. https://doi.org/10.3982/ECTA12675

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2023). When Should You Adjust Standard Errors for Clustering?*. *The Quarterly Journal of Economics*, *138*(1), 1–35. https://doi.org/10.1093/qje/qjac038

Allensworth, E., Nomi, T., Montgomery, N., & Lee, V. E. (2009). College Preparatory Curriculum for All: Academic Consequences of Requiring Algebra and English I for Ninth Graders in Chicago. *Educational Evaluation and Policy Analysis*, *31*(4), 367–391. https://doi.org/10.3102/0162373709343471

Altonji, J. G. (1995). The Effects of High School Curriculum on Education and Labor Market Outcomes. *The Journal of Human Resources*, *30*(3), 409–438. https://doi.org/10.2307/146029

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, *30*, 98–108. https://doi.org/10.1016/j.labeco.2014.05.008

Antonovics, K., Black, S. E., Cullen, J. B., & Meiselman, A. Y. (2022). *Patterns, Determinants, and Consequences of Ability Tracking: Evidence from Texas Public Schools* (Working Paper 30370). National Bureau of Economic Research. https://doi.org/10.3386/w30370

Asim, M., Kurlaender, M., & Reed, S. (2019). 12th Grade Course-Taking and the Distribution of Opportunity for College Readiness in Mathematics. In *Policy Analysis for California Education, PACE*. Policy Analysis for California Education, PACE. https://eric.ed.gov/?id=ED600439

Ayalon, H., & Gamoran, A. (2000). Stratification in Academic Secondary Programs and Educational Inequality in Israel and the United States. *Comparative Education Review*, *44*(1), 54–80.

Baird, M. D., & Pane, J. F. (2019). Translating Standardized Effects of Education Programs Into More Interpretable Metrics. *Educational Researcher*, *48*(4), 217–228. https://doi.org/10.3102/0013189X19848729

Betts, J. R. (2011). Chapter 7—The Economics of Tracking in Education. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education* (Vol. 3, pp. 341–381). Elsevier. https://doi.org/10.1016/B978-0-444-53429-3.00007-7

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions. *Journal of Research on Educational Effectiveness*, *1*(4), 289–328. https://doi.org/10.1080/19345740802400072

Brighouse, H., Ladd, H. F., Loeb, S., & Swift, A. (2018). *Educational Goods: Values, Evidence, and Decision-Making*. University of Chicago Press.

California Department of Eduction. (2023). *California Assessment of Student Performance and Progress Smarter Balanced Summative Assessment 2021–22 Technical Report*. California Department of Education Assessment Development & Administration Division.

Card, D., & Giuliano, L. (2016). Universal screening increases the representation of low-income and minority students in gifted education. *Proceedings of the National Academy of Sciences*, *113*(48), 13678–13683. https://doi.org/10.1073/pnas.1605043113

Clotfelter, C. T., Ladd, H. F., Clifton, C. R., & Turaeva, M. R. (2021). School Segregation at the Classroom Level in a Southern 'New Destination' State. *Race and Social Problems*, *13*(2), 131–160. https://doi.org/10.1007/s12552-020-09309-w

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2015). The Aftermath of Accelerating Algebra Evidence from District Policy Initiatives. *Journal of Human Resources*, *50*(1), 159–188. https://doi.org/10.3368/jhr.50.1.159

Cohodes, S. R. (2020). The Long-Run Impacts of Specialized Programming for High-Achieving Students. *American Economic Journal: Economic Policy*, *12*(1), 127–166. https://doi.org/10.1257/pol.20180315

Collins, C. A., & Gan, L. (2013). *Does Sorting Students Improve Scores? An Analysis of Class Composition* (Working Paper 18848). National Bureau of Economic Research. https://doi.org/10.3386/w18848

Conger, D. (2005). Within-School Segregation in an Urban School District. *Educational Evaluation and Policy Analysis*, *27*(3), 225–244. https://doi.org/10.3102/01623737027003225

Conger, D., Long, M. C., & Iatarola, P. (2009a). Explaining race, poverty, and gender disparities in advanced course-taking. *Journal of Policy Analysis and Management*, *28*(4), 555–576. https://doi.org/10.1002/pam.20455

Conger, D., Long, M. C., & Iatarola, P. (2009b). Explaining race, poverty, and gender disparities in advanced course-taking. *Journal of Policy Analysis and Management*, *28*(4), 555–576. https://doi.org/10.1002/pam.20455

Cortes, K. E., & Goodman, J. S. (2014). Ability-Tracking, Instructional Time, and Better Pedagogy: The Effect of Double-Dose Algebra on Student Achievement. *American Economic Review*, *104*(5), 400–405. https://doi.org/10.1257/aer.104.5.400

Dalane, K., & Marcotte, D. E. (2020). The Segregation of Students by Income in Public Schools. In *EdWorkingPapers.com*. Annenberg Institute at Brown University. https://www.edworkingpapers.com/ai20-338

Davis, T. M. (2014). School Choice and Segregation: "Tracking" Racial Equity in Magnet Schools. *Education and Urban Society*, *46*(4), 399–433. https://doi.org/10.1177/0013124512448672

Domina, T., McEachin, A., Hanselman, P., Agarwal, P., Hwang, N., & Lewis, R. W. (2019). Beyond Tracking and Detracking: The Dimensions of Organizational Differentiation in Schools. *Sociology of Education*, *92*(3), 293–322. https://doi.org/10.1177/0038040719851879

Dougherty, S. M., Goodman, J. S., Hill, D. V., Litke, E. G., & Page, L. C. (2017). Objective course placement and college readiness: Evidence from targeted middle school math acceleration. *Economics of Education Review*, *58*, 141–161. https://doi.org/10.1016/j.econedurev.2017.04.002

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*, *101*(5), 1739–1774. https://doi.org/10.1257/aer.101.5.1739

Feld, J., & Zölitz, U. (2017). Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects. *Journal of Labor Economics*, *35*(2), 387–428. https://doi.org/10.1086/689472

Figlio, D. N., & Page, M. E. (2002). School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality? *Journal of Urban Economics*, *51*(3), 497–514. https://doi.org/10.1006/juec.2001.2255

Francis, D. V., & Darity, W. A. (2021). Separate and Unequal Under One Roof: How the Legacy of Racialized Tracking Perpetuates Within-School Segregation. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *7*(1), 187–202. https://doi.org/10.7758/RSF.2021.7.1.11

Gamoran, A. (1989). Measuring Curriculum Differentiation. *American Journal of Education*, *97*(2), 129–143. https://doi.org/10.1086/443918

Gamoran, A., & Hannigan, E. C. (2000). Algebra for Everyone? Benefits of College-Preparatory Mathematics for Students With Diverse Abilities in Early Secondary School. *Educational Evaluation and Policy Analysis*, *22*(3), 241–254. https://doi.org/10.3102/01623737022003241

Good, T. L., Grouws, D. A., & Beckerman, T. M. (1978). Curriculum pacing: Some empirical data in mathematics. *Journal of Curriculum Studies*, *10*(1), 75–83. https://doi.org/10.1080/0022027780100106

Goodman, J. (2019). The labor of division: Returns to compulsory high school math coursework. *Journal of Labor Economics*, *37*(4), 1141–1182. https://doi.org/10.1086/703135

Huffaker, E., Novicoff, S., & Dee, T. S. (forthcoming). Ahead of the game? Course-taking patterns under a math pathways reform. *Educational Researcher*.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics*, *95*(7), 798–812. https://doi.org/10.1016/j.jpubeco.2010.11.009

Heß, S. (2017). Randomization inference with stata: A guide and software. *The Stata Journal: Promoting Communications on Statistics and Stata*, *17*(3), 630–651. https://doi.org/10.1177/1536867X1701700306

Hippel, P. T. von. (2024). Multiply by 37 (or Divide by 0.023): A Surprisingly Accurate Rule of Thumb for Converting Effect Sizes from Standard Deviations to Percentile Points. In *EdWorkingPapers.com*. Annenberg Institute at Brown University. https://edworkingpapers.com/ai23-829

Imberman, S. A., Kugler, A. D., & Sacerdote, B. I. (2012). Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees. *American Economic Review*, *102*(5), 2048–2082. https://doi.org/10.1257/aer.102.5.2048

Kalogrides, D., & Loeb, S. (2013). Different Teachers, Different Peers: The Magnitude of Student Sorting Within Schools. *Educational Researcher*, *42*(6), 304–316. https://doi.org/10.3102/0013189X13495087

Kang, C. (2007). Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea. *Journal of Urban Economics*, *61*(3), 458–495. https://doi.org/10.1016/j.jue.2006.07.006

Kariya, T., & Rosenbaum, J. E. (1999). Bright Flight: Unintended Consequences of Detracking Policy in Japan. *American Journal of Education*, *107*(3), 210–230. https://doi.org/10.1086/444216

Kelly, S. (2004). Are Teachers Tracked? On what Basis and with what Consequences. *Social Psychology of Education*, *7*(1), 55–72. https://doi.org/10.1023/B:SPOE.0000010673.78910.f1

Kim, S. (2018). *Return to Algebra II: The Effect of Mandatory Math Coursework on Postsecondary Attainment* (SSRN Scholarly Paper ID 3351235). Social Science Research Network. https://doi.org/10.2139/ssrn.3351235

Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, *49*(4), 241–253.

Krone Phillips, E. (2019). *The Make-or-Break Year | The New Press*. https://thenewpress.com/books/make-or-break-year

Labaree, D. F. (1988). *The Making of an American High School: The Credentials Market and the Central High School of Philadelphia, 1838-1939*. Yale University Press.

Lafortune, J. (2018). *Essays on the Distribution and Effectiveness of Educational Resources* [UC Berkeley]. https://escholarship.org/uc/item/8qh075sj

Lamdin, D. J. (1996). Evidence of Student Attendance as an Independent Variable in Education Production Functions. *The Journal of Educational Research*, *89*(3), 155–162. https://doi.org/10.1080/00220671.1996.9941321

Lefgren, L. (2004). Educational peer effects and the Chicago public schools. *Journal of Urban Economics*, *56*(2), 169–191. https://doi.org/10.1016/j.jue.2004.03.010

Liu, J., & Loeb, S. (2021). Engaging Teachers: Measuring the Impact of Teachers on Student Attendance in Secondary School. *Journal of Human Resources*, *56*(2), 343–379. https://doi.org/10.3368/jhr.56.2.1216-8430R3

Long, M. C., Conger, D., & Iatarola, P. (2012). Effects of High School Course-Taking on Secondary and Postsecondary Success. *American Educational Research Journal*, *49*(2), 285–322. https://doi.org/10.3102/0002831211431952

McEachin, A., Domina, T., & Penner, A. (2020). Heterogeneous Effects of Early Algebra across California Middle Schools. *Journal of Policy Analysis and Management*, *39*(3), 772–800. https://doi.org/10.1002/pam.22202

Mehana, M., & Reynolds, A. J. (2004). School mobility and achievement: A meta-analysis. *Children and Youth Services Review*, *26*(1), 93–119. https://doi.org/10.1016/j.childyouth.2003.11.004

Moses, R. P., & Cobb Jr., C. (2001). Organizing Algebra: The Need to Voice a Demand. *Social Policy*.

Murata, A. (2013). Diversity and High Academic Expectations Without Tracking: Inclusively Responsive Instruction. *Journal of the Learning Sciences*, *22*(2), 312–335. https://doi.org/10.1080/10508406.2012.682188

Nomi, T., & Allensworth, E. (2009). “Double-Dose” Algebra as an Alternative Strategy to Remediation: Effects on Students’ Academic Outcomes. *Journal of Research on Educational Effectiveness*, *2*(2), 111–148. https://doi.org/10.1080/19345740802676739

Oakes, J. (1995). Two Cities’ Tracking and Within-School Segregation. *Teachers College Record*, *96*(4), 1–7. https://doi.org/10.1177/016146819509600418

Oakes, J. (2005). *Keeping Track: How Schools Structure Inequality*. Yale University Press.

Oakes, J., Ormseth, T., Bell, R. M., & Camp, P. (1990). *Multiplying Inequalities: The Effects of Race, Social Class, and Tracking on Opportunities to Learn Mathematics and Science*. RAND Corporation. https://www.rand.org/pubs/reports/R3928.html

Pustejovsky, J. E., & Tipton, E. (2018). Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models. *Journal of Business & Economic Statistics*, *36*(4), 672–683. https://doi.org/10.1080/07350015.2016.1247004

Rosenbaum, J. E. (1999). If Tracking Is Bad, Is Detracking Better? *American Educator*, *23*(4), 24.

Sacerdote, B. (2014). Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward? *Annual Review of Economics*, *6*(1), 253–272. https://doi.org/10.1146/annurev-economics-071813-104217

Schiller, K. S., & Hunt, D. J. (2011). Secondary Mathematics Course Trajectories: Understanding Accumulated Disadvantages in Mathematics in Grades 9–12. *Journal of School Leadership*, *21*(1), 87–118. https://doi.org/10.1177/105268461102100105

StataCorp. (2023). *Stata 18 Item Response Theory Reference Manual*. Stata Press.

Stein, M. K., Kaufman, J. H., Sherman, M., & Hillen, A. F. (2011). Algebra: A Challenge at the Crossroads of Policy and Practice. *Review of Educational Research*, *81*(4), 453–492. https://doi.org/10.3102/0034654311423025

Tucker, J. (2023, May 10). *S.F. could bring back algebra in eighth grade. Here's what you need to know about city's math wars*. San Francisco Chronicle. https://www.sfchronicle.com/bayarea/article/math-algebra-eighth-grade-sfusd-18085751.php

Valiande, S., & Koutselini, D. M. I. (2009). *APPLICATION AND EVALUATION OF DIFFERENTIATION INSTRUCTION IN MIXED ABILITY CLASSROOMS*.

Vygotsky, L. S. (2011). The Dynamics of the Schoolchild's Mental Development in Relation to Teaching and Learning. *Journal of Cognitive Education and Psychology*, *10*(2), 198–211. https://doi.org/10.1891/1945-8959.10.2.198

Wainstein, L., College, R., Miller, C. E., Phillips, M., Yamashiro, K., & Melguizo, T. (2023). *Twelfth Grade Math and College Success.*

Wells, A. S., & Serna, I. (2010). The Politics of Culture: Understanding Local Political Resistance to Detracking in Racially Mixed Schools. *Harvard Educational Review*, *66*(1), 93–119. https://doi.org/10.17763/haer.66.1.274848214743t373

Welsh, R. O. (2017). School Hopscotch: A Comprehensive Review of K–12 Student Mobility in the United States. *Review of Educational Research*, *87*(3), 475–511. https://doi.org/10.3102/0034654316672068

What Works Clearinghouse. (2014). *Teaching Strategies for Improving Algebra Knowledge in Middle and High School Students.* https://ies.ed.gov/ncee/wwc/PracticeGuide/20

Williams, R., Citkowicz, M., Miller, D. I., Lindsay, J., & Walters, K. (2022). Heterogeneity in Mathematics Intervention Effects: Evidence from a Meta-Analysis of 191 Randomized Experiments. *Journal of Research on Educational Effectiveness*, *15*(3), 584–634. https://doi.org/10.1080/19345747.2021.2009072

Zwiers, J., Dieckmann, J., Rutherford-Quach, S., Daro, V., Skarin, R., Weiss, S., & Malamut, J. (2017). Principles for the Design of Mathematics Curricula: Promoting Language and Content Development. *Understanding Language.*

Table 1. Intent-to-Treat (ITT) Effects of the Algebra Initiative on Peer and Teacher Traits

Variable	Peer Characteristics								Teacher Characteristics	
	% Female	% FRPM	% English Learner	% Hispanic	% White	% Asian/Pacific Islander	% Black	Baseline Math Achievement	Teacher Has Advanced Certification	Years of Teacher Experience
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ITT × At Grade Level	-2.6279	-5.7903+	2.3603	-0.6246	2.5092	-1.7735	-0.3048	0.0353	-0.2644*	-2.6740
	(1.8608)	(2.9749)	(1.8409)	(2.7136)	(2.7552)	(2.9930)	(0.8647)	(0.0925)	(0.1196)	(2.1168)
ITT × Nearly at Grade Level	-3.7042	-10.8411**	-0.6322	-4.0139	3.9343	0.2958	-0.5980	0.2565**	-0.0416	0.9208
	(2.2323)	(3.1882)	(2.6696)	(3.3230)	(3.1289)	(2.3231)	(1.0318)	(0.0863)	(0.0399)	(1.5760)
ITT × Below Grade Level	-1.6828	-30.0119***	-34.8696***	-17.2054***	21.7841***	-2.5376	-2.7275	1.2827***	0.0322	0.3685
	(4.4828)	(3.3715)	(3.4143)	(3.4435)	(2.7970)	(2.3027)	(2.7903)	(0.0650)	(0.0499)	(1.2256)
Nearly at Grade Level	2.0405	-17.5717***	-34.3724***	-11.4917**	17.0115***	-2.0038	-3.8589	1.0588***	0.0644	-0.5274
	(4.5031)	(3.3064)	(3.9787)	(3.5105)	(3.1126)	(2.3171)	(2.4007)	(0.0822)	(0.0398)	(1.5007)
Below Grade Level	1.3326	-23.6105***	-38.9858***	-16.8639***	20.3341***	0.0674	-3.9601	1.3384***	0.2303*	2.9618
	(4.3691)	(3.8138)	(3.4899)	(3.1306)	(2.7804)	(2.9842)	(2.5410)	(0.1014)	(0.0974)	(2.0187)
p value: (H_0: $\beta_1 = \beta_2 = \beta_3$)	0.4209	0.0000	0.0000	0.0000	0.0000	0.4581	0.7851	0.0000	0.1207	0.1257
Control Mean | Below Grade Level	49.5823	77.8134	50.2244	72.4320	8.5303	9.8620	9.0205	-1.2515	0.9375	9.5312
Observations	1039	1039	1039	1039	1039	1039	1039	1039	1036	1036

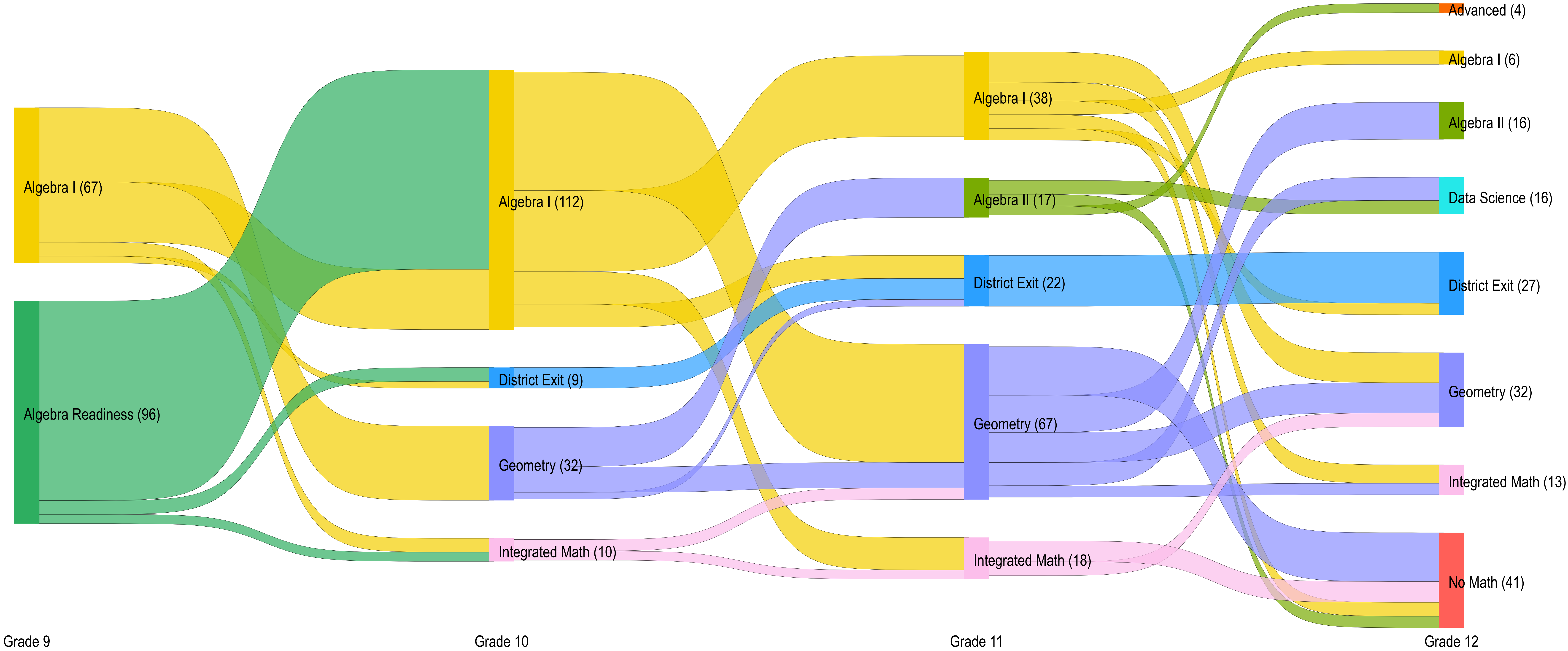
Notes: This table summarizes the effect of assignment to the Algebra I Initiative on peer and teacher characteristics. Columns 1 through 8 present the percentage point difference in classroom concentration of baseline student traits between treatment and control conditions, for each achievement statrum. The measure of math achievement in column 8 is constructed by standardizing high school math readiness asssesment (MDTP) scores and imputing a standardized 7th grade score for the small number of students (n=5) missing this data. Columns 9 presents the extent to which assignment to the Initiative changes a student's likelihood of receiving a teacher with a graduate degree and/or national board certificaiton, while column 10 summarizes the effect of Initiative assignment on average maximum years of teacher experience. All models control for 9th grade campus membership. Standard errors are clustered at the ninth grade classroom level.+ $p<.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Table 2. Intent-to-Treat Effects (ITT) of the Algebra Initiative on Measures of Student Engagement

Variable	9th Grade			10th Grade			11th Grade			12th Grade		
	Enrolled Through Academic Year	Chronically Absent	Absence Rate (%)	Enrolled Through Academic Year	Chronically Absent	Absence Rate (%)	Enrolled Through Academic Year	Chronically Absent	Absence Rate (%)	Enrolled Through Academic Year	Chronically Absent	Absence Rate (%)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ITT ✕ At Grade Level	0.0002	-0.0034	0.0919	0.0092	0.0002	-0.2906	0.0578*	-0.0067	-0.0169	0.0638*	-0.0306	-0.1830
	(0.0078)	(0.0121)	(0.2718)	(0.0283)	(0.0231)	(0.7048)	(0.0267)	(0.0266)	(0.6692)	(0.0280)	(0.0467)	(0.9597)
ITT ✕ Nearly at Grade Level	-0.0021	0.0351	0.8471	0.0374	0.0739	1.2713	-0.0202	-0.0130	-1.0298	0.0398	0.0993	3.0712
	(0.0239)	(0.0415)	(0.8444)	(0.0338)	(0.0493)	(1.5447)	(0.0516)	(0.0645)	(2.1515)	(0.0538)	(0.0862)	(2.0829)
ITT ✕ Below Grade Level	0.0337	-0.0747*	-1.9784*	0.0103	-0.0863	-7.4185***	0.0728	-0.0881	-5.9688*	0.1268*	-0.0495	-4.3176
	(0.0243)	(0.0349)	(0.8480)	(0.0329)	(0.0629)	(1.8294)	(0.0483)	(0.0888)	(2.5776)	(0.0486)	(0.0845)	(2.8075)
Nearly at Grade Level	0.0402+	-0.0892***	-2.6571**	0.0146	-0.2431***	-11.0243***	0.0761+	-0.3044***	-9.3166***	0.1122**	-0.2582**	-7.4899**
	(0.0229)	(0.0244)	(0.7796)	(0.0319)	(0.0523)	(1.6130)	(0.0438)	(0.0567)	(2.3614)	(0.0408)	(0.0763)	(2.7399)
Below Grade Level	0.0152	-0.0685*	-1.7787*	-0.0214	-0.2100***	-8.6745***	0.0570	-0.1402*	-3.9867	0.0659	-0.1820*	-5.4663+
	(0.0240)	(0.0323)	(0.8514)	(0.0311)	(0.0541)	(1.7562)	(0.0416)	(0.0607)	(2.6608)	(0.0459)	(0.0792)	(2.9752)
Control Mean | Below Grade Level	0.9490	0.1183	5.0258	0.9286	0.3626	15.6035	0.8061	0.5063	16.6826	0.7347	0.5139	15.4438
Observations	1039	1022	1022	1039	980	980	1039	935	935	1039	907	907

Notes: This table presents impacts of assignment to the Algebra I Initiative on measurs of student engagement and connection to school. For each grade level indicated, this table shows effects on enrollment through the end of that academic year, on the likelihood of chronic absenteeism (e.g., absent greater than 10% of all enrolled days and on the continuous percent rate of days absent over total days enrolled). All models control for 9th grade campus membership. Standard errors are clustered at the ninth grade classroom level.+ p<.1, * p<0.05, ** p<0.01, *** p<0.001.

Students Entering District Below Grade Level: Algebra I Initiative vs. Readiness Pathways



Notes: Grade 9 totals indicate initial assignment. Pathways with 2 or fewer students excluded

Table 3. Intent-to-Treat (ITT) Effects of the Algebra Initiative on Student Course Access and Attainment

	Through Grade 10		Through Grade 11		Through Grade !2				
Variable	Passed Algebra I	Passed Geometry	Passed Geometry	Passed Algebra II	Passed Algebra II	Passed Any Advanced Math	Met UC/CSU Math Admission Requirement	Total Math Credits	Graduated
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ITT × At Grade Level	0.0192	0.0458	0.0277	0.0409	0.0524	-0.0338	0.0107	0.7642	0.0024
	(0.0148)	(0.0424)	(0.0363)	(0.0408)	(0.0377)	(0.0513)	(0.0437)	(0.8993)	(0.0269)
ITT × Nearly at Grade Level	0.0152	-0.1298*	-0.0410	-0.0258	-0.0181	0.0274	-0.0163	0.3732	-0.0807+
	(0.0447)	(0.0580)	(0.0736)	(0.0734)	(0.0795)	(0.0673)	(0.0777)	(1.6619)	(0.0458)
ITT × Below Grade Level	0.0784	0.2187***	0.1204	0.1434**	0.1138+	0.0594	0.0041	2.7074+	0.0302
	(0.0819)	(0.0596)	(0.0786)	(0.0520)	(0.0644)	(0.0607)	(0.0618)	(1.3518)	(0.0602)
Nearly at Grade Level	0.3174***	0.6990***	0.4949***	0.6338***	0.5998***	0.5480***	0.6037***	15.0331***	0.1150*
	(0.0451)	(0.0534)	(0.0622)	(0.0557)	(0.0545)	(0.0669)	(0.0645)	(1.2274)	(0.0474)
Below Grade Level	0.2211***	0.3719***	0.2310**	0.2574***	0.2176***	0.1302*	0.1704*	6.0274***	0.0736
	(0.0551)	(0.0510)	(0.0678)	(0.0500)	(0.0601)	(0.0549)	(0.0649)	(1.3073)	(0.0479)
p value: (H_0: $\beta_1 = \beta_2 = \beta_3$)	0.5317	0.0006	0.3639	0.0218	0.1024	0.6593	0.9909	0.1232	0.3519
Control Mean | Below Grade Level	0.6413	0.0109	0.3250	0.0125	0.0909	0.0779	0.1429	20.1461	0.8052
Observations	991	991	940	940	923	922	922	923	923

Notes: This table presents estimates of the impact of assignment to the Algebra initiative on key academic outcomes, through the indicated grade level. All models control for 9th grade campus membership. Standard errors are clustered at the ninth grade classroom level.+ p<.1, * p<0.05, ** p<0.01, *** p<0.001.

Table 4. Intent-to-Treat (ITT) Effects of the Algebra Initiative on Standardized Math Test Scores

Variable	10th Grade District Assessmentn (ICA)		11th Grade State Asessment (SBAC)	
	(1)	(2)	(3)	(4)
ITT × At Grade Level	0.0270	0.0036	0.0745	0.0603
	(0.0626)	(0.0613)	(0.0739)	(0.0690)
Romano-Wolf Adjusted P-Value	[0.7163]	[0.9520]	[0.2857]	[0.3626]
ITT × Nearly at Grade Level	0.0309	0.0208	0.1332	0.1377
	(0.0856)	(0.0890)	(0.0857)	(0.0881)
Romano-Wolf Adjusted P-Value	[0.7682]	[0.8551]	[0.1698]	[0.1518]
ITT × Below Grade Level	0.1433	0.1613	0.1904*	0.2027*
	(0.1108)	(0.1051)	(0.0863)	(0.0768)
Romano-Wolf Adjusted P-Value	[0.2358]	[0.1768]	[0.0729]	[0.0589]
At Grade Level	1.0631***	0.8944***	1.2677***	1.0657***
	(0.0817)	(0.0889)	(0.1037)	(0.1117)
Nearly at Grade LeveL	0.2812**	0.1950*	0.3299***	0.2641*
	(0.0876)	(0.0934)	(0.0937)	(0.0992)
Includes Controls for Student Traits	No	Yes	No	Yes
p value: (H_0: $\beta_1 = \beta_2 = \beta_3$)	0.5375	0.4949	0.0498	0.0209
Control Mean | Below Grade Level	-0.8397	-0.8397	-0.8817	-0.8817
Observations	850	850	805	805

Notes: This table presents estimates of the impact of assignment to the Algebra initiative on standardized test score outcomes. Columns 1 and 2 present impacts on a 10th grade district-level ICA assesment, while columns 3 and 4 summarize Initiative impacts on the state SBAC assesment, taken in 11th grade. ICA scores are standardized over the sample, while SBAC scores are standardized using the mean and standard deviation of the statewide distribution. All models control for campus membership. Columns 2 and 4 also feature controls for baseline student race/ethnicity, gender, limited english proficiency status and free or reduced price lunch status. Standard errors are clustered at the ninth grade classroom level. + $p<.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Multiple comparison adjustments are indicated with Romano-Wolf p-values (resample n=10,000).

Data Appendix

Defining Advanced Coursework

Advanced courses are defined as any which proceed Algebra II. In a few cases, advanced courses are categorized based on either their position in the hierarchical course sequences or according to their a-g articulation domain. One of the District campuses offers International Baccalaureate (IB) courses for students who have completed either the Algebra I, Geometry or Algebra II sequence (or equivalent content). Standard level (SL) IB courses are grouped with Pre-Calculus as a "fourth year" high school course while Higher level (SL) courses are grouped with AP Calculus offerings as a "fifth year" course. This aligns with UC/CSU treatment of these classes. Additionally, the District began offering "Explorations in Data Science" to 12th graders in our focal cohort in AY 2022-23.

At the time they enrolled data science courses were categorized as a Statistics equivalent (i.e., an Algebra II-validating advanced course) by the UC/CSU High School Articulation Unit. However, this designation is pending clarification after being revoked, citing insufficient evidence on data science course rigor, by the Board of Admissions and Relations with Schools (Fensterwald, 2023). Our main analysis treats the data science course as advanced (i.e., Algebra II-validating) consistent with its status during the academic year of attendance. However, we test the sensitivity of potentially affected measures (i.e., meeting the "c" criterion and advanced course enrollment) to removing this status.

Defining Student Persistence

Once a student from the analytic sample no longer appears in the data (i.e., in the fall semester of AY 2020-2021) they are assumed to have exited the district. However, their inclusion in analysis of downstream outcomes depends on whether they are reported to have legitimately transferred to another high school, or if they are considered "truant" or a "dropout". If students transfer to a school outside of the district within an academic year, which occurs in approximately 10 cases per year after ninth grade, they are excluded from analysis using outcome data from that year. For example, if a student is enrolled in the District at the beginning of AY 2021-22 for 11th grade but is noted as moving to another state after 50 days of enrollment, they are dropped from the 11th grade analytic sample. If a student is classified as a dropout or as truant, however, they are retained in the sample. If a student exits the district between academic years we cannot identify whether they have dropped out of high school or just transferred districts, and they are classified as an attriter. A small number of students (n=8) exited and then returned to the district. For these students, we recover their complete high school transcript upon return, but not attendance data for the academic year(s) they were not attending the District. Discussion of Table 4 includes more detail on the implications of exit rates via transfers on our main findings.

Table A1. Summary of treatment-control contrasts by baseline achievement group

Student Group by Baseline Achievement According to middle school test scores and transcripts	**Control**	**Treatment**
At Grade level	Regular Algebra I (n=304)	Algebra I with Initiative pedagogy and mixed-achievement peers (n=366)
Nearly At Grade level	Regular Algebra I + support class (n=121)	Algebra I with Initiative pedagogy and mixed-achievement peers (n=83)
Below Grade level	Algebra Readiness (n=98)	Algebra I with Initiative pedagogy and mixed-achievement peers (n=67)

Note: Baseline achievement groups are determine according to Figure A1 using middle school test scores and transcripts.

Table A2. Descriptive Statistics

Variables	Mean	SD	Minimum	Maximum	*N*
Student Characteristics at High School Entry					
Female	0.51	0.50	0	1	1039
Free-and-Reduced Price Meals	0.49	0.50	0	1	1039
English Learner	0.16	0.36	0	1	1039
White	0.30	0.46	0	1	1039
Hispanic	0.54	0.50	0	1	1039
Asian or Pacific Islander	0.12	0.32	0	1	1039
Black	0.04	0.19	0	1	1039
Native American or Alaska Native	0.01	0.09	0	1	1039
Attend Campus #1 in Grade 9	0.28	0.45	0	1	1039
Attend Campus #2 in Grade 9	0.24	0.43	0	1	1039
Attend Campus #3 in Grade 9	0.25	0.43	0	1	1039
Attend Campus #4 in Grade 9	0.22	0.42	0	1	1039
Assigned Placement					
At Grade Level	0.64	0.48	0	1	1039
Nearly at Grade Level	0.16	0.37	0	1	1039
Not at Grade Level	0.20	0.40	0	1	1039
Assigned to the Initiative	0.50	0.50	0	1	1039
Math Test Scores (Standardized)					
Baseline Readiness Score	22.00	9.84	0	45	1039
10th Grade Interim Comprehensive Assessment (ICA)	0.00	0.85	-2	2	850
11th Grade State Assessment (SBAC)	0.12	0.84	-2	2	805
Transcript Outcomes					
Completed Algebra I by End of Grade 10	0.99	0.09	0	1	908
Completed Geometry by End of Grade 11	0.63	0.48	0	1	985
Completed Geometry by End of Grade 11	0.79	0.41	0	1	936
Completed Algebra II by End of Grade 11	0.55	0.50	0	1	940
Completed Algebra II by End of Grade 12	0.60	0.49	0		920
Completed Any Advanced Math by End of Grade 12	0.49	0.50	0	1	920
Met Math Requirement for UC/CSU	0.61	0.49	0	1	920
Total Math Credits	32.58	10.57	0	60	920
Graduated On Time	0.91	0.29	0	1	920
Attendance Outcomes					
Chronically Absent in Grade 9	0.04	0.21	0	1	1024
Chronically Absent in Grade 10	0.14	0.34	0	1	984
Chronically Absent in Grade 11	0.25	0.43	0	1	936
Chronically Absent in Grade 12	0.29	0.45	0	1	916
Absence Rate in Grade 9 (%)	2.87	4.27	0	50	1024
Absence Rate in Grade 10 (%)	5.48	12.36	0	86	984
Absence Rate in Grade 11(%)	8.72	11.35	0	90	936
Absence Rate in Grade 12 (%)	9.42	11.62	0	91	916

Notes: Our analytic sample includes 1,039 students who were enrolled in an Algebra class for Semester 1 of 9th grade in one of the four comprehensive high schools in the partner district, and who attended a feeder elementary district so they could be matched to their middle school academic records. All data are sourced from the partner districts or one of the feeder elementary districts.

Table A3. Balance in Student Characteristics by Intent-to-Treat (ITT) X Placement Group

Student Characteristic	Female (1)	FRPM (2)	English Learner (3)	Baseline Math Achievement (Standardized) (4)	Hispanic (5)	White (6)	Asian (7)	Black (8)
ITT ✕ At Grade Level	-0.0050	-0.0732	-0.0231	0.0540	-0.0229	0.0293	-0.0111	-0.0047
	(0.0437)	(0.0379)	(0.0276)	(0.0503)	(0.0382)	(0.0371)	(0.0270)	(0.0159)
ITT ✕ Nearly at Grade Level	-0.0832	-0.0146	0.1019*	0.0200	0.0241	0.0423	-0.0185	-0.0545*
	(0.0713)	(0.0618)	(0.0450)	(0.0820)	(0.0623)	(0.0606)	(0.0441)	(0.0260)
ITT ✕ Below Grade Level	-0.0705	-0.0260	-0.1150*	0.1131	-0.0209	0.0388	-0.0453	0.0388
	(0.0790)	(0.0685)	(0.0498)	(0.0908)	(0.0690)	(0.0671)	(0.0488)	(0.0288)
At Grade Level	-0.0818	-0.3730***	-0.4824***	1.9299***	-0.2352***	0.2913***	-0.0023	-0.0418
	(0.0620)	(0.0538)	(0.0391)	(0.0713)	(0.0542)	(0.0527)	(0.0383)	(0.0226)
Nearly At Grade Level	-0.0412	-0.0874	-0.3124***	0.5243***	0.0284	0.0442	-0.0495	-0.0190
	(0.0677)	(0.0587)	(0.0427)	(0.0779)	(0.0592)	(0.0575)	(0.0419)	(0.0247)
Control Mean | Below Grade Level	0.5612	0.8367	0.5510	-1.3923	0.7449	0.0510	0.1122	0.0816

Notes: This table shows the balance of characteristics of students in the randomized sample. Each column presents coefficients from a regression of a baseline characteristic on interactions between assignment to treatment and placement level at baseline, controlling for both campus membership and baseline achievement group. The p-value of a joint-significance test for all 24 coefficients of interest is 0.6553, indicating no systematic imbalance. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. N=1,039

Table A4. Intent-to-Treat (ITT) Estimates of Test Score Availability

Variable	10th Grade Interim Comprehensive Assessment (ICA)		11th Grade State Assessment (SBAC)	
	(1)	(2)	(3)	(4)
ITT × At Grade Level	0.0008	-0.0073	0.0467	0.0435
	(0.0352)	(0.0354)	(0.0357)	(0.0350)
ITT × Nearly at Grade Level	-0.1077	-0.0905	0.0419	0.0459
	(0.0666)	(0.0658)	(0.0525)	(0.0530)
ITT × Below Grade Level	0.1209	0.1069	0.1016	0.0939
	(0.0732)	(0.0724)	(0.0617)	(0.0620)
Includes Controls for Student Traits	No	Yes	No	Yes
p value: (H_0: $\beta_1 = \beta_2 = \beta_3$)	0.2412	0.3071	0.1731	0.2163
Control Mean | Below Grade Level	0.6224	0.6020	0.6224	0.6020
Observations	1039	1039	1039	1039

Notes: This table presents estimates of the impact of Algebra Initiative assignment on the availability of key test scores. All models control for campus membership and baseline proficiency group. Columns 2 and 4 also control for baseline student race/ethnicity, gender, limited english proficiency status and free or reduced price lunch status. Standard errors are clustered at the ninth grade the classroom level.+ p<.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A5. Intent-to-Treat (ITT) Effects of the Algebra Initiative on 11th Grade State Assessment (SBAC) Scores, Alternative Specifications

Variable	Classroom-Level Clustered Standard Errors		Robust Standard Errors		Small Sample Corrected Standard Errors		Design-Based Inference	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ITT × At Grade Level	0.0745	0.0603	0.0745	0.0603	0.0745	0.0603	0.0922	0.0669
	(0.0739)	(0.0690)	(0.0700)	(0.0675)	(0.0769)	(0.0721)	(0.0705)	(0.0687)
ITT × Nearly at Grade Level	0.1332	0.1377	0.1332	0.1377	0.1332	0.1377	0.1228	0.1359
	(0.0857)	(0.0881)	(0.0936)	(0.0939)	(0.0884)	(0.0914)	(0.0951)	(0.0964)
ITT × Below Grade Level	0.1904*	0.2027*	0.1904+	0.2027+	0.1904*	0.2027*	0.1842+	0.1783
	(0.0863)	(0.0768)	(0.1046)	(0.1049)	(0.0903)	(0.0803)	(0.1089)	(0.1113)
Includes Controls for Student Traits	No	Yes	No	Yes	No	Yes	No	Yes

Notes: This table presents alternative estimates of impact of assignment to the Algebra initiative on a standardized 11th grade state test scored (i.e., the SBAC). All models control for campus membership. Columns 2, 4,6 and 8 also feature controls for baseline student race/ethnicity, gender, limited english proficiency status and free or reduced price lunch status. Standard errors in columns 1 and 2 are clustered as the level of the 9th grade classroom. Heterogeneity-robust Eicker-Huber-White standard errors are used for columns 3 and 4. For Columns 5 and 6, standard errors are adjusted to improve-small sample performance using the reg_sandwich package. Standard errors in columns 7 and 8 are constructed using permutations of treatment status over repeated Monte Carlo simulations + p<.1, * p<0.05, ** p<0.01, *** p<0.001. N=805.

Table A6. Intent-to-Treat (ITT) Effects of the Algebra Initiative on Key Outcomes, with Alternative Controls

Variable	11th Grade SBAC Score (1)	9th Grade Absence Rate (2)	10th Grade Absence Rate (3)	11th Grade Absence Rate (4)	Enrolled Through Grade 12 (5)	Ever Passed Algebra II (6)	Total Math Credits (7)
(A) Main Specification							
ITT ✕ At Grade Level	0.0745	0.0919	-0.2906	-0.0169	0.0638*	0.0524	0.7642
	(0.0739)	(0.2718)	(0.7048)	(0.6692)	(0.0280)	(0.0377)	(0.8993)
ITT ✕ Nearly at Grade Level	0.1332	0.8471	1.2713	-1.0298	0.0398	-0.0181	0.3732
	(0.0857)	(0.8444)	(1.5447)	(2.1515)	(0.0538)	(0.0795)	(1.6619)
ITT ✕ Below Grade Level	0.1904*	-1.9784*	-7.4185***	-5.9688*	0.1268*	0.1138+	2.7074+
	(0.0863)	(0.8480)	(1.8294)	(2.5776)	(0.0486)	(0.0644)	(1.3518)
(B) Includes Controls for Student Traits							
ITT ✕ At Grade Level	0.0603	0.1094	-0.1347	0.0922	0.0593*	0.0414	0.4925
	(0.0690)	(0.2770)	(0.7169)	(0.6609)	(0.0278)	(0.0382)	(0.8614)
ITT ✕ Nearly at Grade Level	0.1377	0.9134	1.2624	-0.9302	0.0426	-0.0108	0.6381
	(0.0881)	(0.8770)	(1.6011)	(2.1912)	(0.0560)	(0.0706)	(1.4818)
ITT ✕ Below Grade Level	0.2027*	-2.0270*	-7.4371***	-5.9927*	0.1232*	0.1076+	2.4480+
	(0.0768)	(0.8522)	(1.9059)	(2.6225)	(0.0490)	(0.0549)	(1.3364)
(C) Includes a Control for Baseline Score							
ITT ✕ At Grade Level	0.0368	0.1404	-0.0659	0.1322	0.0588*	0.0354	0.3271
	(0.0538)	(0.2849)	(0.7144)	(0.6867)	(0.0276)	(0.0358)	(0.7972)
ITT ✕ Nearly at Grade Level	0.1253	0.9302	1.3063	-0.8910	0.0423	-0.0161	0.4935
	(0.0868)	(0.8765)	(1.5936)	(2.1972)	(0.0561)	(0.0705)	(1.4545)
ITT ✕ Below Grade Level	0.1656*	-1.9723*	-7.3106***	-5.8931*	0.1219*	0.0967+	2.1501
	(0.0785)	(0.8461)	(1.9014)	(2.6462)	(0.0491)	(0.0553)	(1.3668)
Observations	805	1022	980	935	1039	923	923
(D) Includes Controls for 9th Grade Math Teacher Traits							
ITT ✕ At Grade Level	0.1456+	0.1114	-0.0820	0.0632	0.0803*	0.0740	1.5462
	(0.0793)	(0.3083)	(0.8235)	(0.7673)	(0.0304)	(0.0469)	(1.1001)
ITT ✕ Nearly at Grade Level	0.1338	0.8478	1.3726	-0.9401	0.0427	-0.0162	0.4588
	(0.0870)	(0.8410)	(1.5264)	(2.1586)	(0.0535)	(0.0802)	(1.6794)
ITT ✕ Below Grade Level	0.1940*	-1.6757*	-6.8721***	-4.8937+	0.1154*	0.1116+	2.4768+
	(0.0860)	(0.6722)	(1.7617)	(2.5001)	(0.0487)	(0.0656)	(1.3610)
Observations	803	1019	976	931	1035	920	920

Notes: This table presents estimates for the impact of assignment to the Initiative on key outcomes using alternative student-level controls. Panel A presents estimates using our preferred specification (see notes for Tables 1-4). Panel B presents results with controls for student characteristics. Panel C presents results for models that add a standardized, continuous measure of baseline math proficiency to our preferred specification. Panel D includes controls for observed ninth grader teacher traits including years of experience and whether a teacher held an advanced degree or certification. Because not all students in our sample took the same pre-trest, this measure imputes missing MDTP scores (n=5) using 7th grade SBAC results. All models control for initial campus membership and baseline proficiency group. Standard errors are clustered at the ninth grade classroom level.+ p<.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A10. Teacher-Level Sensitivity of Intent-to-Treat (ITT) Effects of the Algebra Initiative on Key Outcomes

		Omitted Sections by Teacher:							
Dependent Variable: Standardized 11th Grade SBAC Score	Full Sample (1)	A (2)	B (3)	C (4)	D (5)	E (6)	F (7)	G (8)	H (9)
ITT × At Grade Level	0.0745	0.0514	0.0649	0.0974	0.0783	0.1088	0.0763	0.0577	0.1255+
	(0.0739)	(0.0735)	(0.0721)	(0.0911)	(0.0748)	(0.0773)	(0.0732)	(0.0757)	(0.0735)
ITT × Nearly at Grade Level	0.1332	0.1530+	0.1364	0.1644+	0.1433	0.1061	0.0829	0.1150	0.1509+
	(0.0857)	(0.0857)	(0.0941)	(0.0854)	(0.0975)	(0.0880)	(0.0895)	(0.0911)	(0.0843)
ITT × Below Grade Level	0.1904*	0.2113*	0.1584+	0.1889*	0.2087*	0.1973*	0.1690+	0.1789+	0.2017*
	(0.0863)	(0.0867)	(0.0908)	(0.0880)	(0.0953)	(0.0872)	(0.0888)	(0.0896)	(0.0882)
Observations	923	872	829	882	862	860	817	882	855
(B) Dependent Variable: 9th Grade Absence Rate									
ITT × At Grade Level	0.0919	0.0976	0.2334	0.2736	0.0365	0.0592	-0.0055	0.1627	0.0580
	(0.2718)	(0.2864)	(0.2935)	(0.3193)	(0.2809)	(0.2580)	(0.2839)	(0.2971)	(0.2902)
ITT × Nearly at Grade Level	0.8471	1.0296	1.2408	0.9153	1.2393	1.0313	-0.1351	0.8539	0.8528
	(0.8444)	(0.8941)	(1.0308)	(0.8702)	(0.9342)	(0.8921)	(0.6027)	(0.9078)	(0.8738)
ITT × Below Grade Level	-1.9784*	-2.2488*	-1.8757*	-1.9525*	-2.2215*	-1.8210*	-1.8264*	-1.8156*	-1.9798*
	(0.8480)	(0.8708)	(0.9143)	(0.8495)	(0.9242)	(0.8599)	(0.8495)	(0.8405)	(0.8619)
Observations	935	788	747	793	781	776	737	797	765
(C) Dependent Variable: Ever Passed Algebra II									
ITT × At Grade Level	0.0524	0.0506	0.0529	0.0538	0.0363	0.0476	0.0619	0.0593	0.0686+
	(0.0377)	(0.0363)	(0.0384)	(0.0465)	(0.0365)	(0.0395)	(0.0377)	(0.0385)	(0.0405)
ITT × Nearly at Grade Level	-0.0181	-0.0148	-0.0243	-0.0157	-0.0133	-0.0769	-0.0106	-0.0011	-0.0047
	(0.0795)	(0.0816)	(0.0875)	(0.0798)	(0.0872)	(0.0729)	(0.0859)	(0.0833)	(0.0797)
ITT × Below Grade Level	0.1138+	0.1413+	0.0829	0.1189+	0.1029	0.1621*	0.0877	0.0978	0.1037
	(0.0644)	(0.0710)	(0.0684)	(0.0650)	(0.0738)	(0.0657)	(0.0676)	(0.0615)	(0.0660)
Observations	923	872	829	882	862	860	817	882	855

Notes: This table presents estimates of the impact of assignment to the Algebra Initiative on key outcomes. Each of columns 2-9 drop one of the eight Initiative teachers from the sample. All models control for initial campus membership, and baseline proficiency group. Standard errors are clustered at the ninth grade classroom level. + p<.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A8. Intent-to-Treats (ITT) Effects of the Algebra Initiative on Key Outcomes by Student Subgroup

	Full Sample (n=1039) (1)	Female (n=530) (2)	Male (n=509) (3)	FRPM (n=512) (4)	Non-FRPM (n=527) (5)
(A) Dependent Variable: 11th Grade SBAC Score (Standardized over state)					
ITT × At Grade Level	0.0745	0.1212	0.0378	0.1607	0.0189
	(0.0739)	(0.0998)	(0.1016)	(0.1208)	(0.0907)
ITT × Nearly at Grade Level	0.1332	0.1673	0.1025	0.1995*	-0.1348
	(0.0857)	(0.1275)	(0.1053)	(0.0948)	(0.2513)
ITT × Below Grade Level	0.1904*	0.0159	0.3654*	0.2214*	-0.0046
	(0.0863)	(0.1240)	(0.1403)	(0.0930)	(0.2641)
Observations	805	405	400	380	425
(B) Dependent Variable: 9th Grade Absence Rate					
ITT × At Grade Level	0.0919	0.3785	-0.1458	0.0746	0.0862
	(0.2718)	(0.4641)	(0.2817)	(0.4065)	(0.3477)
ITT × Nearly at Grade Level	0.8471	0.0367	1.2934	1.0901	-0.1121
	(0.8444)	(0.5988)	(1.4591)	(1.0902)	(0.9767)
ITT × Below Grade Level	-1.9784*	-2.6245*	-1.2189	-1.9515*	-1.9556
	(0.8480)	(1.2023)	(0.8878)	(0.8640)	(2.0598)
Observations	1022	522	500	499	523
(C) Dependent Variable: Ever Passed Algebra II					
ITT × At Grade Level	0.0524	0.0921+	0.0192	-0.0269	0.0902*
	(0.0377)	(0.0540)	(0.0525)	(0.0683)	(0.0366)
ITT × Nearly at Grade Level	-0.0181	0.0131	-0.0106	-0.0688	0.1699
	(0.0795)	(0.1124)	(0.0773)	(0.0835)	(0.1556)
ITT × Below Grade Level	0.1138+	0.1614+	0.0662	0.0575	0.2497+
	(0.0644)	(0.0954)	(0.0789)	(0.0831)	(0.1264)
Observations	923	473	450	445	478

Notes: This table presents the heterogeneous estimates for the impact of assignment to the Initiative on key outcomes by student subgroup. Column 1 presents results for the full sample using our preferred specification (additional notes with Tables 1-4). All regressions control for initial campus membership and baseline proficiency group. Standard errors are clustered as the 9th grade classroom level.+ p<.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A9. Intent-to-Treat (ITT) Effects of the Algebra Initiative on Non-Math Academic Outcomes

	Cumulative Through Grade 12:					Ever Through Grade 12:			
Variable	Grade Point Average (1)	Total Credits Through Grade 12 (2)	English Credits Through Grade 12 (3)	Science Credits Through Grade 12 (4)	Elective Credits Through Grade 12 (5)	Taken AP/IB (6)	Switched Campus In-District (7)	Classified Dropout/Truant (8)	Graduated On-Time (8)
ITT × At Grade Level	0.0357	5.4682*	-0.0331	0.9509+	2.4630+	-0.0040	-0.0279+	-0.0053	0.0024
	(0.0565)	(2.4234)	(0.3705)	(0.5608)	(1.3718)	(0.0392)	(0.0162)	(0.0209)	(0.0269)
ITT × Nearly at Grade Level	-0.1416	-15.0335+	-1.9006	-2.0881+	-5.8821+	0.0151	0.0430	-0.0143	-0.0807+
	(0.1159)	(7.5917)	(1.5137)	(1.1315)	(3.1968)	(0.0162)	(0.0432)	(0.0310)	(0.0458)
ITT × Below Grade Level	0.0686	2.6938	2.4458+	1.5274	-8.8124*	-0.0115	-0.1174	0.0211	0.0302
	(0.1091)	(7.0859)	(1.2986)	(1.0874)	(4.0448)	(0.0139)	(0.0766)	(0.0315)	(0.0602)
At Grade Level	0.9292***	29.5964***	4.1901***	8.0076***	-11.5874***	0.1313**	-0.1804*	-0.0106	0.1150*
	(0.0818)	(5.2937)	(1.1183)	(0.8470)	(3.2408)	(0.0432)	(0.0681)	(0.0238)	(0.0474)
Nearly At Grade Level	0.1969*	13.6037*	2.0524	4.2375***	-7.5469*	-0.0149	-0.1272+	0.0153	0.0736
	(0.0938)	(5.9884)	(1.3049)	(0.9769)	(3.5624)	(0.0142)	(0.0751)	(0.0238)	(0.0479)
p value: (H_0: $\beta_1 = \beta_2 = \beta_3$)	0.5442	0.0734	0.2223	0.0536	0.0212	0.6234	0.0992	0.8596	0.3519
Control Mean | Below Grade Level	2.0624	211.3480	34.7222	18.4271	48.6389	0.0130	0.4533	0.0510	0.8052
Observations	922	916	907	907	907	923	923	1039	923

Notes: This table presents the impacts of Initiative assignment on additional academic outcomes through 12th grade using our preferred specification. All models control for campus membership. Standard errors are clustered at the ninth grade classroom level.+ p<.1, * p<0.05, ** p<0.01, *** p<0.001.