The Unintended Consequences of Test-Based Remediation

David Figlio and Umut Ozek*

* Figlio: University of Rochester, david.figlio@rochester.edu; Ozek: RAND Corporation. The authors are grateful to numerous seminar participants and to the twelve anonymous Florida school districts for providing the confidential administrative data for this project. All errors are their own.

The Unintended Consequences of Test-Based Remediation

Abstract: School systems around the world use achievement tests to assign students to schools, classes, and instructional resources, including remediation. Using a regression discontinuity design, we study a Florida policy that places middle school students who score below a proficiency cutoff into remedial classes. Students scoring below the cutoff receive more educational resources, but they are also placed in classes that are more segregated by race, socio-economic status, and prior achievement. Increased tracking occurs not only in the remedial subject, but also in other core subjects. These tracking effects are significantly larger and more likely to persist beyond the year of remediation for Black students.

Keywords: test-based remediation; tracking; classroom segregation

The Unintended Consequences of Test-Based Remediation

1. Introduction

Schooling systems around the world use achievement tests to assign students to schools, classes, and instructional resources. These practices range from determining admission to elite exam or other selective-enrollment schools (e.g., Abdulkadiroglu, Angrist, and Pathak, 2014; Clark, 2010; Dobbie and Fryer, 2011; Estrada and Gignoux, 2017; Lucas and Mbiti, 2014; Ozier, 2018; Pop-Eleches and Urquiola, 2013) to different tracks within schools (e.g., Betts and Shkolnik, 2000; Card and Giuliano, 2016; Duflo, Dupas, and Kremer, 2011; Epple, Newlon, and Romano, 2002; Figlio and Page, 2002; McEachin, Domina, and Penner, 2020) to assignment to gifted or remedial education (e.g., Bui, Craig, and Imberman, 2014; Cortes and Goodman, 2014; Cortes, Goodman, and Nomi, 2015; Dougherty, 2015; Nomi and Allensworth, 2009; Ozek, 2021; Taylor, 2014).[1]

The argument in favor of using achievement tests for remediation is that doing so provides additional targeted resources that can promote student mastery of the subject. Indeed, there exists some evidence that providing remediation can boost student test scores and longer-run outcomes (Cortes, Goodman, and Nomi, 2015; Nomi and Allensworth, 2009; Ozek, 2021) but some studies find even negative short-run effects (e.g., Dougherty, 2015) or positive short-run effects that fade over time (e.g., Taylor, 2014).

But these additional resources may come at a cost if they lead to students receiving remediation being segregated into classes with lower-achieving peers, and that might inhibit students' ability to pursue more advanced study in the topic. Moreover, identifying students in need of remediation in one subject might lead to a labeling-related spillover effect in which students deemed as requiring remediation in one subject end up segregated into classes in entirely different subjects. This may lead to within-school segregation by race and/or socio-economic status (SES): In addition to the degree to which structural inequities in American society produce racial/ethnic test score gaps, there is evidence of racial bias in track placements (Mickelson 2015; Grissom and Redding 2016) and that higher-SES parents are more likely to

[1] Schooling systems also frequently use test performance to determine grade promotion and retention; recent evaluations include Eren, Depew, and Barnes (2017); Eren, Lovenheim, and Mocan (2018); Figlio and Ozek (2020); Greene and Winters (2007); Greene and Winters (2012); Larsen and Valant (2018); Mariano et al. (2018); and Schwerdt, West, and Winters (2017).

challenge their children's placement in lower tracks (Barg 2012) and avoid remediation (LiCalsi et al. 2019).[2] And this within-school segregation may influence students' access to educational resources (e.g., Kalogrides and Loeb 2013, Mickelson 2015) and might contribute to stereotype threat or otherwise influence students' sense of self-concept as an academic achiever (e.g, Chu et al, 2018; Domina, Penner, and Penner, 2017; Chmielewski, Dumont, and Trautwein, 2013; Wouters et al., 2012).

In this paper, we study what happens when school systems impose test-based remediation. We employ a regression discontinuity design and highly detailed student-level data from 12 anonymous county-level school districts in Florida to examine the effects of Florida's middle school remediation policy, which requires students who score below the proficient level on prior year reading or math tests to be placed on a remedial schedule in that subject. On the one hand, the test-based remediation program delivers more instructional resources to students deemed in need of remediation. Those students assigned to remedial classrooms are taught in smaller classes with more experienced, higher-value-added (in terms of contributions to student test scores) teachers who are more likely to be of the same race/ethnicity as the students themselves; given the documented benefits of same-race teachers on outcomes of students from underrepresented groups (e.g., Gershenson et al. 2018) this could be a beneficial consequence of the remediation policy.

On the other hand, the policy also leads to reduced opportunities for students, in terms of track placement – even in subjects unrelated to the subject where the student is deemed in need of remediation – and further exacerbates within-school segregation of students by race, ethnicity, and SES. Specifically, we find significant negative effects of scoring below the cutoff on the probability of taking advanced courses not just in the subject of remediation, but in all other core academic subjects, suggesting a labeling effect of remediation. Unsurprisingly, therefore, these students also are assigned to classrooms with lower-achieving peers in both the subject of remediation and *other core subjects*. These spillover effects are large: Scoring below the remediation cutoff in reading reduces the average prior performance of classroom peers by one-

[2] The existing evidence on within-school classroom racial segregation points toward a story that remediation efforts might lead to within-school segregation, since within-school racial segregation increases as students age. Clotfelter et al. (2003) find that within-school segregation accounted for just a fifth of all White/non-White segregation in elementary grades, but over half in middle and high school. Similarly, Conger (2005) and Kalogrides and Loeb (2013) find that classroom segregation within schools account for a much larger share of total school segregation in middle and high school compared to elementary schools.

fourth of the standard deviation (0.25σ) in English/Language Arts (ELA) courses, which might be expected (though is not pre-ordained) given the nature of remediation, *but also 0.07σ in math courses, 0.09σ in science courses, and 0.11σ in social studies courses* in sixth grade. In other words, test-based remediation in one subject apparently induces a form of global tracking in *all subjects*.[3]

Perhaps unsurprisingly given the prior literature on uneven implementation of universal policies (e.g., LiCalsi et al, 2019), we find that remediation-induced tracking patterns are especially pronounced for Black students: Scoring below the remediation cutoff in reading reduces a Black student's likelihood of taking an advanced ELA course by 17 percentage points (54 percent of the control mean at the cutoff). But the spillover effects on non-remediated subjects are also substantial: Black students targeted for remediation in reading are 6 percentage points (20 percent) less likely to take advanced math courses, 9 percentage points (40 percent) less likely to take advanced science courses, and 10 percentage points (50 percent) less likely to take advanced social studies courses. In contrast, the same effects for White students are still present but smaller: 7 percentage points (33 percent of the control mean at the cutoff) in ELA, 1.5 percentage points (6 percent) in math, 3 percentage points (15 percent) in science, and 3 percentage points (30 percent) in social studies. These Black/White gaps persist even after conditioning on student test score history and school setting. Further, while these tracking effects dissipate beyond the year of remediation for White students, they decline in magnitude, yet persist, for Black students.

We also examine the effects of the remediation policy on student outcomes in middle and high school using the cohorts of students who entered 6th grade in the two years after the policy took effect and hence were appropriately aged to graduate from high school by the end of our sampling frame (assuming on-time grade progression). We find positive test score effects in the year of remediation that dissipate over time and precisely estimated zero effects (or modest positive effects on high school advanced course-taking) on non-test outcomes including

[3] However, it is important to note that, unlike some of the tracking policies in Europe (for example, see Krause-Pilatus and Schüller 2014 for a review of the literature on secondary school tracking policies in Germany), it is possible for students placed in the remedial track to take more advanced courses after the year of remediation. In fact, Ozek (2021) finds no effect of being placed in the remedial schedule in middle school on the likelihood of taking an advanced middle school course in ELA or other core subjects after the year of remediation. There is, therefore, a limit to the degree to which these findings can be compared to European-style tracking.

disciplinary problems, attendance, high school graduation, and college credit-bearing course-taking in high school.

In summary, therefore, we find that test-based remediation delivers some remedial resources to students but does so at a considerable cost to those same students, especially for Black students. We conclude that a key challenge for policymakers is therefore to find opportunities to allocate educational resources in a compensatory fashion (e.g., with teachers and other classroom resources) but conducted in settings that do not exacerbate within-school segregation or place students into tracks that could inhibit subsequent success. There exists dramatic variation in the degree of classroom segregation and academic tracking in other subjects as a result of the remediation policy across school districts in our study, suggesting that it is possible to offer additional remedial resources to students without increasing within-school segregation.

2. Florida's Middle School Remediation Policy

In 2004, the state of Florida enacted legislation (s.1003.415) requiring all entering sixth grade students who scored below the proficient level (lowest two achievement levels out of five) on the prior year Florida Comprehensive Assessment Test (FCAT) reading test to be administered a personalized middle school success plan to attain state and school district expectations in academic proficiency. The FCAT score is a continuous measure, and provides the running variable for the analyses we conduct. A 2006 legislation (s.1003.4156) further expanded this policy and required that each year a middle school student (grades 6 through 8) scores below the proficient level on reading assessment in the previous year, the student takes two English language arts (ELA) courses—a regular course and a remedial course—instead of just one course. A similar requirement was also implemented in math, mandating low-performing students to receive remediation in the following year, which may be integrated into the student's regular math course.

While there are several exemptions to this requirement[4], the legislation significantly increased remedial course-taking in middle school. Panel (A) in Figure 1 presents the share of

[4] For example, low-performing students who do not have intervention needs in the areas of foundational reading skills (e.g., decoding, fluency) or students who have consistently scored above the proficient level in the past may be exempt from ELA remediation. Schools also have discretion over the final schedule of the student—low-performing

sixth graders taking remedial ELA or math courses between 2002-03 and 2010-11 school years in the 12 anonymous county-level school districts in Florida that provided us the data for this study. In the year before the legislation (2003-04 school year), 7 percent of 6th graders took remedial ELA courses. This number steadily increased after the legislation and reached 23 percent in 2010-11. A similar, yet flatter, trend is apparent in remedial math, with the share of 6th graders taking these courses increasing from 3 percent in 2003-04 to roughly 10 percent in 2010-11.

Panel (B) in Figure 1 examines how classroom achievement segregation (i.e., segregation of 6th grade classrooms based on student prior achievement, as measured by averaged reading and math scores) has changed over the same time frame in remediation subjects (ELA and math) and core subjects that were not subject to remediation (science and social studies). In this exercise, for each school-year, we first obtain the R^2 from regressions where we regress the averaged prior year test scores of 6th graders on classroom fixed effects in ELA, math, and science/social studies classrooms. We then regress these R^2s on year indicators -- with the year before the policy (2003-04 school year) serving as the baseline -- to examine how classroom achievement segregation changed over time by subject. Panel (B) in Figure 1 plots the estimated coefficients on these year indicators (along with their 95% confidence intervals obtained using standard errors clustered at the school level). The findings suggest that the share of within-school variation in prior achievement explained by cross-classroom variation increased dramatically (by nearly 12 percentage points in ELA classrooms, from 42 percent to 54 percent), suggesting increased classroom prior achievement segregation after the policy took effect. Further, this segregation is not confined to remediation subjects (ELA and math): Between 2003-04 and 2010-11, the share of within-school variation in prior achievement explained by cross-classroom variation in science and social studies classrooms in 6th grade increased by a statistically significant 10 percentage points (from 34 percent to 44 percent).

While these trends provide suggestive evidence that the remediation policy increased classroom achievement segregation, it is hard to fully attribute these trends to the causal effects of the policy due to secular trends in classroom segregation or other concurrent policies. Florida, for example, introduced a large-scale school accountability policy in 1999 that had several

---

students may be exempt, for example, if the need for remediation in math is greater than the student's need for reading remediation.

refinements during the study period, and also introduced school choice policies (both based on these school accountability measures [Chiang, 2009; Figlio and Rouse, 2006; Rouse et al, 2013] as well as targeted to family income [Figlio and Hart, 2014; Figlio, Hart, and Karbownik, 2020]) in the years immediately before and during our study period.[5] To address these concerns, since none of these policies are related to student performance around the cutoffs considered by the remediation policy[6], we rely on a regression discontinuity design and compare students whose prior year test scores fell right below the remediation cutoff (and hence were subjected to the policy) and those above the cutoff. We detail this empirical strategy in Section 4.

3. Data

In our analysis, we make use of detailed longitudinal, student-level administrative data that cover the school years between 2002–03 and 2010–11 from 12 anonymous, county-level school districts in Florida. These data cover all students enrolled in grades K–12 and include FCAT reading and math scores for all students between grades three and ten in these districts, and a wealth of student characteristics including student demographics (e.g., race/ethnicity, gender), whether the student is eligible for subsidized meals, country of birth, and special education status. These data also include detailed information on course enrollment such as unique classroom identifiers, time spent in each classroom (minutes per week), course type (e.g., subject; remedial, regular, or advanced), and teacher characteristics (experience, race/ethnicity, gender). The school districts linked these student-level administrative data with birth records for all children born in Florida between 1992 and 2002; these birth records contain maternal characteristics such as educational attainment, marital status, and age at birth, which provide measures of SES not typically observed in school records. In our analysis, we focus on 9 cohorts of 6th graders between the 2002-03 and 2010-11 school years, with two cohorts entering 6th grade before the remediation policy and 7 cohorts after the policy took effect.

[5] Another relevant policy in Florida that was implemented during this time frame was the gifted education program in a large school district which required schools to set up separate classrooms for all gifted students in 4th and 5th grades (Card and Giuliano, 2016). We do not expect this policy to be a meaningful driver of our main results as (1) it was targeting elementary grades and (2) it relied on thresholds based on Naglieri Non-verbal Ability Test (NNAT) scores rather than FCAT scores. That said, we check the robustness of our main results to the exclusion of one district at a time to ensure that they are not driven by an individual district in the online appendix.

[6] Many of these policies are not linked to student test scores per se. Those that are based on student test scores are either not related to proficiency thresholds or, if they are, are focused on a different (lower) proficiency threshold than the proficiency threshold at play in the policy being addressed in this paper.

Table 1 presents the descriptive statistics for students whose prior year test scores fell below and above the remediation cutoff in reading and math for the seven post-policy cohorts of 6th graders. The results reveal stark differences between these student groups in terms of race/ethnicity and measures of SES: Students whose prior year scores fell below the remediation cutoff in reading or math are significantly more likely to be eligible for subsidized meals, more likely to be Black, less likely to have a college-educated mother, and more likely to have a mother who was not married and/or who was a teenager at the time of the student's birth. These differences imply that separating classrooms by achievement will likely lead to more segregated classrooms by race/ethnicity and SES as well.

4. Empirical Framework

We employ a regression discontinuity design to estimate the causal effect of Florida's middle school remediation policy on students' track placement, classroom composition, and medium and long-run outcomes. Figure 2 presents the discontinuity in remedial course-taking around the cutoff in ELA and math, before and after the policy took effect. (There are 95 percent confidence bands around the estimated relationships.) These figures reveal that students whose prior year reading (math) scores fell right below the remediation cutoff are roughly 40 percentage points (10 percentage points in math) more likely to take a remedial course compared to students right above the cutoff.[7] We find no such discontinuities in the years before the policy.

In our study, we are mainly interested in the overall effect of Florida's middle school remediation policy. As such, we follow a "reduced-form" approach and estimate the effect of scoring below the remediation cutoff in the main analysis. Let S_i denote the difference between the FCAT scale score of student *i* on prior year reading or math test and the remediation cutoff—with negative values indicating scores below cutoff—and B_i denote an indicator for students below the cutoff. In this setting, the effect of failing the prior year reading or math test is given by:

$$\beta = lim_{S_i \uparrow 0} E[Y_i | S_i] - lim_{S_i \downarrow 0} E[Y_i | S_i]. \quad (1)$$

[7] There are also statistically significant, albeit much smaller, discontinuities in remedial course-taking at the cutoffs between achievement levels 1 and 2 on reading and math tests. In this study, we focus on the cutoff between achievement levels 2 and 3 mainly because the policy identifies students in need of remediation based on the 2/3 cutoff.

where Y_i is the outcome of interest. We are interested in 3 main outcomes: (1) classroom-peer and teacher characteristics in 6th grade; (2) course-taking in middle school; and (3) test scores, suspensions, absences, college credit-bearing course-taking in high school, and high school graduation. For classroom-peer characteristics, we consider 3 attributes: (1) prior achievement (as measured by averaged reading and math scores on prior year tests); (2) race/ethnicity (% same race/ethnicity); and (3) SES as proxied by a regression-based index we create using subsidized meal eligibility, maternal characteristics for students observed in birth records (maternal education, marital status, age at birth, and mother's country of birth), student country of birth, language spoken at home, and the subsidized meal eligibility rate of the elementary school the student attended.[8] In our analysis, for ease of interpretation we standardize this index to zero mean and unit variance. To calculate classroom peer attributes, we first identify classrooms in the 4 core subjects in 6th grade (ELA, math, science, and social studies), and obtain the average peer attribute in the classroom for each student. We then calculate the weighted peer/teacher characteristic for each student using the time the student spends in each classroom per week.

In our core specification, we estimate β using the following equation and OLS:

$$Y_i = \gamma + \beta B_i + k(S_i) + k(S_i) * B_i + \upsilon_i \tag{2}$$

where $k(S_i)$ is a polynomial function of the relative prior year test score. We estimate this model using the linear polynomial specification, and cluster the standard errors at the prior year test score level as suggested by Lee and Card (2008).[9] In our main analysis, we use a bandwidth of 10 points based on the range of bandwidths suggested for various outcomes by the bandwidth selection procedure in Calonico et al. (2017), and check the robustness of our findings to different bandwidths in Online Appendix Figures 1-4, described later in the paper. To improve the precision of the estimates and to account for any school-level differences at the remediation cutoff, we also control for student baseline characteristics and school-by-year fixed-effects

[8] In particular, similar to Figlio et al. (2014), we regress averaged prior year test scores on the given SES measures and use the predicted values as the SES index.
[9] This clustering approach has been recently questioned (Kolesár and Rothe 2018), and we check the robustness of our findings and compare the 95% confidence intervals obtained using standard errors clustered at the running variable level with confidence intervals obtained using Eicker-Huber-White heteroskedasticity-robust standard errors in Online Appendix Figures 3 and 4. These two approaches yield very similar conclusions using different bandwidths for our main outcomes of interest.

although the exclusion of these covariates makes no meaningful difference in the estimated effects.

In this empirical framework, $\hat{\beta}$ will yield unbiased estimates of the causal effect of failing the prior year test if all other student attributes are smooth around the cutoff. While this condition cannot be definitively proven, we conduct several tests. First, we examine if the observable characteristics of students are continuous by estimating the following models:

$$X_i = \alpha + \varphi B_i + k(S_i) + k(S_i) * B_i + \varepsilon_i \quad (3)$$

where X_i represents baseline student characteristics. In Online Appendix Table 1, we present the results of these falsification tests where each row represents a separate regression using the identified variable as the dependent variable estimated using the bandwidth of 10 points and the linear polynomial specification, and the estimated coefficient $(\hat{\varphi})$ indicates the size of the discontinuity, separately for reading and math cutoffs. Of the 32 estimated coefficients, only 4 are statistically distinguishable from zero, and none of these estimates represent a discontinuity larger than 5 percent of the dependent variable mean at the cutoff.

We then check for the possibility of selection variable manipulation as noted in McCrary (2008), even though this is very unlikely in this context since scores are assessed without any teacher, student, or principal involvement. Online Appendix Figure 5 presents the distribution of students around the remediation cutoff in reading and math, and present no unusual discontinuity at the cutoff even though there appears to be a modest discontinuity in the density of the distribution at the cutoff, with p-values of 0.042 and 0.026 in reading and math respectively (Frandsen, 2017). That said, there are 12 instances in the reading score distribution and 28 instances in the math score distribution where the magnitude of the discontinuity in density exceeds the discontinuity at the cutoff.

5. Results

5.1. Remediation is Differentially Enforced

The prior literature (e.g, LiCalsi et al, 2019) demonstrates that universal policies are not always implemented evenly, and that there may be different implementation by race, ethnicity and SES quartile. In Figure 3 and Online Appendix Figure 6 we investigate this prospect directly and examine the differential effect of scoring below the remediation cutoff on the likelihood of being placed in the remedial schedule in 6th grade whereas Online Appendix Table 2 presents the estimates broken down by race/ethnicity and SES quartiles.

Overall, the results suggest that the remediation policy was implemented for every racial, ethnic, and SES group, yet there is considerable variation. For example, the implementation was apparently stronger for lower-SES students especially in math: Failing the 5th grade math test increases the likelihood of being placed in the remedial math schedule in 6th grade by 13 percentage points among students in the lowest SES quartile compared to 9 percentage points for students in the highest SES quartile (p-value of the difference is less than 0.01). A similar pattern emerges in ELA, yet the differences are smaller and not statistically distinguishable from zero at conventional levels.

Further, the breakdown by race/ethnicity reveals stark differences between the remediation subjects. In ELA, failing the 5th grade test increases the likelihood of taking a remedial course in 6th grade by 40 percentage points for White students compared to 36 percent for Black and Hispanic students (p-value of 0.012 for the White-Black difference and 0.015 for the White-Hispanic difference). In contrast, the effect of failing the 5th grade math test is 12 percentage points for Black and Hispanic students compared to 7 percentage points for White students (p-value less than 0.01 for both the White-Black and White-Hispanic difference). The results also suggest that these racial/ethnic differences are more pronounced for lower-SES students: We find much smaller and statistically insignificant differences in enforcement by race/ethnicity among students in the top SES quartile. These differences in enforcement might have significant distributional consequences depending on what resources, peer groups, track placement, etc. accompany this enforcement: a point we will revisit throughout the paper.

5.2. Remediation Generates Additional Classroom Resources

We observed that scoring just below the test score threshold triggers a substantially increased chance of taking a remedial course. But does this bring with it increases in classroom resources? We begin by investigating the effect of failing the prior year reading or math test on (1) the average leave-out-year teacher value-added score experienced by the student in ELA and math courses[10]; (2) the likelihood of being assigned to at least one teacher with 10+ years of

[10] In particular, we calculate leave-out-year value-added scores separately for reading and math similar to Chetty, Friedman, and Rockoff (2014) using the STATA command *vam* and student data linked to their teachers in middle school between 2002-03 and 2011-12 in the anonymous districts. We use the value-added scores calculated using reading test scores for ELA courses and math scores for math courses, and calculate the average value-added score

experience; (3) class size; and (4) the likelihood of being assigned to at least one same race/ethnicity teacher for the remediation subject and others. Figure 4 presents these results graphically, by race/ethnicity, for the remediation subject (Online Appendix Figure 7 presents the results for other core subjects) and Online Appendix Table 3 presents the estimated treatment effects in tabular form by race/ethnicity and SES quartile.

The results indicate that failing the prior year test has a significant positive effect on the educational resources the student receives in the current year, especially in the remediation subject. For example, the average leave-out-year value-added score of the teachers assigned to the ELA courses of students right below the reading cutoff is 0.6 percent of the standard deviation in student reading scores (or roughly 10 percent of the standard deviation in teacher value-added scores) higher than students right above the cutoff. Similarly, students scoring right below the reading cutoff are 3.9 percentage points (or 8 percent of the control mean at the cutoff) more likely to be assigned to a teacher with at least 10 years of experience, and are assigned to smaller classrooms (by about 1 student) compared to students on the other side of the cutoff. Looking at heterogeneous effects by race/ethnicity, we find that failing the reading test leads to a somewhat larger effect on teacher value-added scores for White students, but the opposite is true for teacher experience in the remediation subject (roughly 12 percent of the control mean at the cutoff for Black and Hispanic students versus a statistically different 2 percent for White students). The effect sizes on class size are comparable across race/ethnicity.

Further, failing the reading test significantly increases the likelihood that Black and Hispanic students are assigned to teachers of the same race/ethnicity. In particular, Black students right below the remediation cutoff are 5.7 percentage points (or 13 percent) more likely to be assigned to a Black teacher in the remediation subject compared to Black students right above the cutoff. There is a similar effect on Hispanic students (11 percent), yet we find no significant effect on White students. These findings suggest that the remediation policy leads to more homogeneous classrooms not only based on student race/ethnicity, but also based on student-teacher racial match. This could yield positive effects for students, given the recent evidence on the benefits of having a same race/ethnicity teacher on student outcomes (e.g., Gershenson et al. 2018).

for each student using the time the student spends with each teacher in that subject as the weight. For more information on the procedure, please see Appendix A and B in Chetty, Friedman, and Rockoff (2014).

Finally, while we find positive effects on educational resources for almost all SES groups, the effects on teacher experience are slightly larger for low-SES students. For students in the lowest SES quartile, scoring below the reading cutoff increases the likelihood of being assigned to an experienced teacher by 14 percent compared to 5 percent for students in the top quartile. Similarly, the effect on being assigned to a same race/ethnicity teacher is larger for low-SES students, primarily driven by the fact that Black and Hispanic students are disproportionately represented in the lower quartiles of the SES index. We find similar, yet smaller effects of failing the math test on educational resources, which also point to improved educational resources for students identified for remediation. Taken together, it appears that the 12 school districts studied in this paper implemented the test-based remediation policy in a manner that enhanced some measures of educational resources for students identified for remediation.

5.3. Remediation Leads to Tracking in Tested Subjects and Other Core Academic Subjects

An important concern with remediation programs is that they could have a labeling effect on low-performing students, which may lead to classroom tracking for these students and hinder their ability to take more advanced courses. Track assignment changes the peer group in the classroom and could also signal something to students about their capabilities and promise. Furthermore, students flagged for remediation in one subject may end up being tracked in other subjects as well, which could reinforce these signals.

We first examine the effect of scoring below the remediation cutoff on a student's track placement. In Florida, middle school courses are categorized into three levels based on difficulty: remedial (or intensive), regular, and advanced courses. There are stark differences in student attributes between these classrooms: Compared to students in advanced courses, students in remedial courses score 1.45 standard deviations lower on prior year tests, are twice as likely to be Black, and have significantly lower SES indices (80 percent of the standard deviation). Similar, yet smaller differences emerge between students in regular and advanced courses (0.8σ difference in prior year test scores, 10 percentage point difference in being Black, and 0.4σ difference in SES index). As such, and given the above-cited racial biases in track placements, scoring below the remediation cutoff could have profound effects on classroom-peer composition if it affects the type of courses the student takes.

The top panel of Figure 5 examines the effects of failing the prior year reading or math test on the likelihood of taking an advanced course in that subject in sixth grade, broken down by race/ethnicity (the top panel in Online Appendix Figure 8 provides the breakdown by SES quartile and Online Appendix Table 4 presents these results tabularly). Unsurprisingly, being flagged for remediation substantially reduces a student's likelihood of taking advanced courses in that subject. But noteworthily, the rate at which being flagged for remediation and the probability of taking same-subject advanced coursework differs dramatically by race/ethnicity. Failing the 5th grade reading test reduces Black students' likelihood of taking advanced ELA courses by 17 percentage points (54 percent of the control mean at the cutoff given in brackets), a much larger change than the 7 percentage point (33 percent) reduction for White students and the 9 percentage point (36 percent) reduction for Hispanic students (p-values < 0.001 for both White-Black and Black-Hispanic differences). In math, the estimated treatment effect for Black students is 13 percentage points (36 percent), again substantially larger than that observed for White students (5 percentage points, or 21 percent), or Hispanic students (6 percentage points, or 23 percent), and both differences are statistically different than zero at 1 percent level. Strong gradients are apparent with regard to SES as well, with lowest-SES students 20 percentage points less likely to take an advanced ELA course when flagged for remediation in reading, and 12 percentage points less likely to take an advanced math course when flagged for remediation in math, as compared with 9 percentage points in reading and 6 percentage points in math for highest-SES students (the differences between top and bottom quartiles are statistically different than zero at 1 percent level).

Being flagged for remediation does not only substantially reduce the likelihood of being placed in an advanced class in the remediated subject. The bottom panels of Figure 5 and Online Appendix Figure 8 (along with Online Appendix Table 4) reveal that scoring below the cutoff leads to being placed in a lower track in other, non-remediated core subjects as well. For example, failing the prior year reading test reduces the likelihood of taking an advanced course in 6th grade by 3 percentage points in math (or by 12 percent), by 5 percentage points in science (25 percent), and by 6 percentage points in social studies (by 43 percent).

These figures also reveal significant heterogeneity by race/ethnicity and SES, with Black students and students in the lowest SES quartile being significantly more likely to be placed in a lower track if they fail prior year tests. For example, for Black students, scoring below the remediation cutoff in reading reduces the likelihood of taking an advanced course by 6 percentage points (20 percent) in math, 9 percentage points (40 percent) in science, and 10 percentage points (50 percent) in social studies. In contrast, the same effects for White students are 1.5 percentage points (6 percent) in math, 3 percentage points (15 percent) in science, and 3 percentage points (30 percent) in social studies. We observe similar discrepancies between Black and White students in math as well. These estimated effects are statistically distinguishable at 5 percent level or higher in all cases. We do not find any significant differences between the effect sizes for White and Hispanic students.

The estimated effects on advanced course-taking are also significantly larger for students in the lowest SES quartile compared to other groups. Specifically, failing the prior year reading test reduces likelihood of taking an advanced course by 59 percent in ELA, 19 percent in math, 34 percent in science, and 58 percent in social studies among students in the lowest SES quartile. The effect sizes for students in the top SES quartile, on the other hand, correspond to 33, 8, 20, and 35 percent of the control mean at the cutoff in ELA, math, science, and social studies respectively. Once again, the estimated coefficients for students in the bottom SES quartile are statistically distinguishable from those obtained using the students in the top SES quartile.

Race and SES are highly correlated: roughly 57 percent of lowest-SES students are Black while only 13 percent of highest-SES students are Black in our sample. Are there differences in the probability of track placement based on the intersection of race and SES? In Online Appendix Table 5, we address this question and find that the same-subject tracking effects in 6th grade are stronger for Black students compared to White and Hispanic students even after accounting for differences in student SES. While we have weaker statistical power in this exercise, the differences in effect sizes between Black students and non-Black students are statistically different than zero at 5 percent level in 5 out of 8 cases. Further, the results once again reveal significant differences in effects of failing the 5th grade test by SES quartile for each racial/ethnic group. In terms of spillover effects in other subjects, Online Appendix Table 6 suggests that the tracking effects are larger for Black students in 6 out of 8 cases, yet these differences between Black and non-Black students are only statistically distinguishable from zero at 10 percent level in 3 cases.

While these findings provide evidence that the remediation policy has profound effects on the types of courses students take and that this effect varies considerably across student groups, it does not necessarily imply that the policy is differentially enforced (or interpreted) due to differences between these groups such as prior achievement and school settings. For example, if low-SES students at the remediation cutoff have lower test scores on prior math or reading tests than high-SES students at the remediation cutoff, scoring below the cutoff could lead to different effects on the two groups.

Online Appendix Tables 7 and 8 examine this possibility and explore heterogeneous effects on the likelihood of taking at least one advanced course in core subjects by race/ethnicity and SES conditional on (1) students' averaged test scores in reading and math in grades 3 through 5, and (2) the average achievement levels of the incoming students as proxied by the averaged student test scores aggregated at the school level respectively. The overarching conclusion from these two tables is that while the effects of the policy vary within each racial/ethnic or SES quartile (often declining as we move from lower-performing to higher-performing students or school settings), in many cases the estimated effects are larger (and statistically distinguishable from other groups) for Black students and students in the lowest SES quartile compared to the effects for other student groups with similar prior achievement or in similar school settings.[11] These findings provide further evidence that the policy is differentially enforced based on student SES and race/ethnicity.

Do these tracking effects persist beyond the year of remediation? This is an important question to assess the extent of the labeling effect and whether it varies along racial and socioeconomic lines. Online Appendix Table 9 presents the estimated effects of failing the 5th grade reading (top panel) or math (bottom panel) test on course-taking in 7th and 8th grades, broken down by student race/ethnicity and SES quartile. While the tracking effects decline for all student groups (and mostly dissipate for White, Hispanic, and high-SES students) beyond the year of remediation, they still exist for Black students. For example, failing the 5th grade reading test increases the likelihood of taking a remedial ELA course by 4 percentage points (or 10 percent of the control mean at the cutoff), decreases the likelihood of taking an advanced ELA course by 5 percentage points (roughly 20 percent), an advanced course in other subjects by 4 percentage points (13 percent) in 7th grade, and an advanced ELA course in 8th grade by 3

[11] One might be curious about how the results look in schools in which all, or almost all, students are flagged for remediation. However, this rarely, if ever, occurs in the data. In fewer than 0.2 percent of cases are more than 90 percent of students in a school flagged for remediation under the policy, and in fewer than 6 percent of cases are more than 70 percent of students flagged for remediation.

percentage points (10 percent) for Black students. Similarly in math, failing the 5th grade test reduces the likelihood of taking an advanced math course by 3 percentage points (10 percent) in 7th grade, by 2 percentage points (15 percent) in 8th grade, and the likelihood of taking an advanced course in other subjects in 8th grade by 2.5 percentage points (5 percent). In contrast, we find precisely estimated zero effects on tracking in 7th and 8th grade for White students. These findings provide evidence that remediation policies could have labeling effects that persist after the year of remediation for Black students and students from socioeconomically disadvantaged backgrounds.

Can these discrepancies between Black versus non-Black and low-SES versus high-SES students be explained by differential enforcement (i.e., differential assignment to the remedial schedule in 6th grade)? Or are they driven by differential exclusion from advanced courses, which is not required by the remediation policy, but is an unintended consequence? In Online Appendix Table 10, we try to address this question by estimating the effect of taking a remedial course in 6th grade in a fuzzy RD framework and examining whether the racial and socioeconomic discrepancies observed in the reduced-form estimates presented thus far still exist. In ELA, the 2SLS estimates for Black students are larger than those for non-Black (White or Hispanic) students in all cases, and these differences are statistically different than zero in all but one case at 5 percent level or higher. Similarly, the effect sizes are larger for students in the bottom SES quartile than those for students in the top quartile (differences are statistically significant at 5 percent level in 3 out of 6 cases). In math, we find no consistent differences, yet it is important to note that the first stage (i.e., the effect of failing the prior year test on remedial course-taking in current year) is much weaker in math (compared to ELA) especially for White and higher-SES students (as reported in Online Appendix Table 2), which makes the 2SLS estimates less reliable. Overall, these findings provide evidence that remediation policies could have unintended tracking consequences especially for Black students and students from disadvantaged backgrounds.

5.4. Remediation Changes Classroom Segregation by Prior Performance, SES, and Race

We next turn to the question of whether the remediation policy apparently induced changes in classroom peer composition. Given the differences in student composition between different courses (i.e., remedial, regular, and advanced) and the tracking effects documented in the previous section, it is plausible to expect that failing the 5th grade test could have a profound effect on classroom-peer composition.

The top panel of Figure 6 presents the discontinuity in average 6th grade classroom-peer prior year achievement in the remediation subject at the ELA remediation cutoff and math remediation cutoff in the years after the policy (the top panel in Online Appendix Figure 9 presents the discontinuities in years before the policy). A tabular version is presented in Online Appendix Table 11 as well. The findings reveal significant effects of failing the prior year test on current year peer prior test performance after the policy took effect. For example, the classroom peers of students right below the remediation cutoff in ELA (math) scored 0.2σ (0.1σ) worse on prior year tests compared to students right above the cutoff. We find no discontinuity in peer achievement at the same cutoffs prior to the policy, providing evidence that the observed effect of failing the test on peer prior test performance is indeed driven by the policy change.

The bottom panel of Figure 6 presents the results for classroom-peer prior achievement in other core subjects in 6th grade and reveals significant spillover effects (the bottom panel in Online Appendix Figure 9 presents the discontinuities in years before the policy). While prior to the remediation policy's implementation, there was no spillover effect on core subjects not flagged for remediation, following the policy's implementation a substantial gap emerged: Failing the prior year reading test decreases average peer prior achievement by 0.07σ in math classrooms, 0.09σ in science classrooms, and 0.11σ in social studies classrooms. While these estimated effects are one-third the size of the 0.25σ effect estimated for ELA classrooms, they represent a substantial spillover, suggesting that flagging a student for remediation in one subject often leads to tracking in *all* subjects. A similar pattern emerges at the math cutoff, with significant effects on peer prior test performance in ELA, science, and social studies classrooms.

The top panel of Figure 7 further break down this analysis by student race/ethnicity (the top panel of Online Appendix Figure 10 presents the SES breakdown). There are several important takeaways. For example, the estimated effects on 6th grade peer prior achievement are larger for Black students than for than for White or Hispanic students not only in the remediation subject but also in other subjects, especially in math. Failing the prior year math test reduces current year peer prior achievement in math classrooms by roughly 0.18σ for Black students (versus 0.09σ for White students or 0.12σ for Hispanic students). In addition, looking at the heterogeneity in treatment effects by SES, we find that the decline in peer prior achievement is strongest for lower-SES students: Scoring below the remediation cutoff in reading reduces peer prior achievement in ELA classrooms by 0.30σ for students in the bottom SES quartile versus 0.26σ for students in the top SES quartile, and scoring below the remediation cutoff in math reduces peer prior achievement in math classrooms by 0.17σ for students in the bottom SES quartile versus 0.12σ for students in the top SES quartile. These racial/socioeconomic discrepancies also exist in other subjects and are statistically distinguishable from zero at the 5 percent level.

Further, the results presented in Online Appendix Table 11 suggest that these effects on peer composition decline, yet still exist, in 7th grade. In ELA, the effects are once again statistically larger for Black students compared to White or Hispanic students in ELA, yet we find no significant effect in 8th grade for all student groups. In math, the results are more mixed with larger effects for Black students in 7th grade whereas the effects are larger for White students in 8th grade. We do not observe any significant differences between students in different SES quartiles after the year of remediation.

There are, of course, other ways to measure classroom segregation besides average peer prior achievement. Online Appendix Figure 11 repeats the same analysis, this time looking at the effect on classroom peer SES by course subject whereas Online Appendix Table 12 provides the results in 6th, 7th and 8th grades. The findings reveal significant negative effects of scoring below the remediation cutoff on peer SES. For example, failing the prior year reading test reduces the SES index of classroom peers in ELA classrooms by 7 percent of the standard deviation of the index. To put this number in context, the difference in SES index between the average student whose mother has a 2- or 4-year college degree and the average student whose mother has a high school diploma or less is 1.1 standard deviations of the SES index. Assuming classrooms with 25 students on average, a 0.07σ decline in peer SES in ELA classrooms is equivalent to replacing roughly 2 students whose mothers have a 2- or 4-year college degree with 2 students whose mothers have a high school diploma in every ELA classroom.

Another important takeaway from Online Appendix Table 12 is that the effect on peer SES does not vary significantly based on students' own SES, except for the effect of failing the prior year reading test on peer SES in ELA classrooms. In that case, students in the top SES quartile experience a decline of 0.092σ of the SES index whereas the effect for students in the bottom quartile is 0.059σ of the index (the difference is statistically significant at 5 percent level). Further, we find that the effects on classroom-peer SES decline, yet still exist, in 7th grade, and vanish in 8th grade.

Likewise, a possible consequence of classroom segregation by prior achievement is segregation by race/ethnicity as students from certain races/ethnicities are disproportionately represented among students scoring below the remediation cutoff, given the societal factors that have produced large and sustained racial/ethnic test score gaps. Figure 8 examines the effect of failing the prior year reading or math test on the share of classroom peers of the same race/ethnicity by course subject and own race/ethnicity for ELA remediation (upper panel) and math remediation (lower panel) whereas Online Appendix Table 13 presents the estimated effects. We observe significant effects of failing the test on classroom-peer racial/ethnic composition, especially for White and Black students in remediation subjects. For example, scoring below the remediation cutoff in ELA increases the share of Black classroom peers by 3.7 percentage points (or 7 percent of the control mean at the cutoff) for Black students in ELA classrooms. We find similar, yet smaller, effects for Black students in non-remedial courses. On the other hand, for White students, failing the prior year test significantly reduces the share of same race/ethnicity classroom peers. These effects persist in 7th grade for Black students while they mostly dissipate for White students. We find no significant effect for Hispanic students.

5.5. Remediation Has Some Positive Effects on High School Outcomes

Our findings thus far point to two competing effects of the remediation policy: the positive effect on educational resources in the year of remediation versus the possible adverse effects of tracking in the year of remediation and the years that follow. Therefore, a natural next place to look is whether students targeted for remediation experience negative effects on academic and behavioral outcomes in middle and high school. Here, the answer appears to be no, at least in the medium-to-long run. In fact, we find modest positive effects on the number of college credit-bearing courses taken in high school. Online Appendix Table 14 presents the estimated effects of failing the prior year reading (top panel) or math test (bottom panel) on middle school test scores, student mobility after 6th grade, the likelihood of receiving a suspension and the average absence rates between grades 6 and 12, the likelihood of taking college-credit bearing courses (Advanced Placement or International Baccalaureate) in high school, and high school graduation. In this exercise, we use the first two post-policy 6th grade cohorts (i.e., 6th graders in 2004-05 and 2005-06) who were old enough to graduate from high school by the 2011-12 school year and break down the analysis by student race/ethnicity to explore whether the effects are different for student groups for whom the classroom segregation effects are more pronounced (e.g., Black students).[12] We present some of these results graphically in Figure 9 (test scores in the year of remediation and average test scores in grades 7 through 10), Figure 10 (number of college credit bearing courses in high school in the subject of remediation and other subjects), and Figure 11 (suspensions and high school graduation).

The results reveal no deleterious effects of the remediation policy on the student academic outcomes we observe. While the positive test score effects in the year of remediation fade out over time, we find modest positive effects on advanced course-taking in high school. For example, being flagged for remediation in ELA increases the number of college credit-bearing courses taken in high school by 0.07 (or by slightly more than 10 percent of the dependent variable mean at the cutoff) overall, by 0.02 (25 percent) in science, and by 0.04 (10 percent) in social studies. While we do not have sufficient statistical power to distinguish the estimated effects across racial groups, the results also suggest that these positive effects are more pronounced for Hispanic students: Failing the 5th grade reading test increases the number of college credit-bearing taken in high school by 20 percent overall, 30 percent in ELA, and nearly 40 percent in science for these students.[13] We find precisely estimated zero effects on student mobility at the end of 6th grade, suspensions and absences in middle and high school, and high school graduation, with the exception of a marginally significant positive effect on high school graduation among White students, which corresponds to roughly 5 percent of the control mean at the cutoff.[14] In sum, these findings suggest that the potential benefits from additional educational

[12] We are unable to break this analysis down by the SES index as the index relies on the subset of students that we can link to birth records (i.e., 6th grade cohorts in 2005-06 and beyond). Therefore, this breakdown would have to rely on a single cohort of 6th graders and hence does not have sufficient statistical power.

[13] That said, it is important to note that the statistical significance of the high school course-taking results is somewhat sensitive to multiple hypothesis testing correction. For example, when we conduct the correction for the reading remediation effects on high school advanced course-taking estimates using all students (top panel, first column), we find that only the effect on advanced course-taking in science remains to be statistically significant at conventional levels (p-value of 0.064), while the overall effect (p-value: 0.112) and the social studies effect (p-value: 0.123) are no longer significant. Similarly, the advanced course-taking effects on Hispanic students are marginally significant after correcting for multiple hypothesis testing.

[14] These findings are consistent with two other studies that examine the effects of remedial courses in Florida context: (1) Özek (2021), who examines the effects of remedial ELA courses in middle school on postsecondary outcomes and finds short-term test score gains that fade out over time, positive effects on advanced course-taking in high school that are smaller for 6th grade remediation, and null effects on suspensions, absences, and high school graduation and (2) Taylor (2014), who examines the effects of remedial math courses in 6th grade on test scores in the short-term and finds positive effects that fade out over time.

resources in the year of remediation offset (and sometimes outweigh) the adverse effects of tracking in middle school.

5.6. Bandwidth Selection and Robustness

Online Appendix Figures 1-4 check the sensitivity of the main findings (effects on 6th grade educational inputs in the remediation subject in Online Appendix Figure 1, effects on tracking in Online Appendix Figure 2, effects on classroom peers in Online Appendix Figure 3, and the effects on student outcomes in Online Appendix Figure 4) to bandwidth selection, and provides the effects (and the 95% confidence intervals) of failing the prior year reading (upper panels) or prior year math test (lower panels). In particular, each spike presents the β coefficient estimated on the outcome of interest using the linear polynomial specification and bandwidths ranging from 5 points to 30 points. These figures also compare the 95% confidence intervals obtained using standard errors clustered at the running variable level (solid spikes) with confidence intervals obtained using Eicker-Huber-White heteroskedasticity-robust standard errors (dashed spikes). The findings suggest that the estimated discontinuities at the cutoff are robust to bandwidth selection in almost all cases.

Another concern in the student outcome analysis is dynamic treatment assignment. Specifically, under Florida's middle school remediation policy, students can be subjected to remediation multiple times through middle and high school unless they score above the remediation cutoff. As a result, a student's treatment in a later grade is partly a function of treatment assignment in the current grade (as future treatment is a function of current test scores). This could imply that the long-term effects of failing prior year tests presented in Online Appendix Table 14 and Figures 9-11 could underestimate the true effects if failing on 5th grade tests reduces the likelihood of failing the following years. We check this possibility in a regression discontinuity design, replacing the outcome variables in Online Appendix Table 14 with indicators for failing the 6th and 7th grade tests. The results (available upon request) reveal precisely estimated zero effects of failing 5th grade tests on the likelihood of failing in subsequent grades. For example, we find that failing the 5th grade reading test reduces the likelihood of failing the reading test in 6th grade by 1.6 percentage points (4 percent of the control mean at the cutoff, p-value: 0.163) and increases this likelihood in 7th grade by 1.6 percentage points (3 percent, p-value: 0.173). We find similar results for the effects of failing the 5th grade math test.

These findings provide suggestive evidence that our estimates are not influenced by dynamic treatment assignment.

Online Appendix Figure 12 examines the extent of differential attrition from the sample in middle and high school around the remediation cutoffs, which can be particularly problematic when examining the effects on high school outcomes. We find no significant discontinuity in the likelihood of leaving the sample in the years following 6^{th} grade, with estimated discontinuities ranging between -0.012 to 0.0003 (all p-values greater than 0.15) in reading and between -0.007 and 0.003 (all p-values greater than 0.4), providing evidence that differential attrition is not a major driver of our results in high school.

Online Appendix Figures 13-16 examine the robustness of our main findings to the exclusion of individual districts similar to Online Appendix Figures 1-4 and show that the results are not driven by the effects observed in a single district. Finally, in Online Appendix Tables 15 and 16, we present the discontinuities in our outcomes of interest (educational resources in 6^{th} grade; advanced course-taking and classroom-peer characteristics in 6^{th}, 7^{th}, and 8^{th} grades; and middle and high school outcomes) estimated non-parametrically by race/ethnicity using the triangle kernel and the optimal bandwidths chosen for various outcomes by the bandwidth selection procedure in Calonico et al. (2017).[15] The results point to similar conclusions: significant positive effects on educational resources in the subject of remediation in 6^{th} grade, differential effects on tracking and classroom-peer composition in middle school by race/ethnicity, and null to modest positive effects on academic and behavioral outcomes in middle and high school.

6. Concluding Remarks

In this study, we take a closer look at test-based remediation policies and examine how they affect educational resources, opportunities, and outcomes for low-performing students, with emphasis on heterogeneities along racial/ethnic and socioeconomic lines. On the one hand, we find that test-based remediation, at least as implemented in Florida, had the intended

[15] Calonico et al. (2015) show that typical bandwidth selection procedures may lead to a choice of bandwidth that is too large, resulting in an asymptotic bias in the RD estimator and inappropriate confidence intervals. To address this issue, we use the bias-corrected RD estimator developed by Calonico et al. (2015) in this analysis. While the running variables in our analysis take a large number of values, we use the discrete running variable adjustments proposed in Calonico, Cattaneo and Farrell (2021) and present the Eicker-Huber-White heteroskedasticity-robust standard errors as suggested by Kolesár and Rothe (2018).

consequence of increased classroom resources for students flagged for remediation. On the other hand, we find significant unintended tracking consequences of this remediation with students flagged for remediation being significantly less likely to take advanced courses not only in the subject of remediation, but also in other core courses. We also show that these unintended effects are significantly larger and persist beyond the year of remediation for Black and low-SES students. Overall, we find precisely estimated zero to modest positive effects on middle and high school outcomes including test scores, suspensions, attendance, college credit-bearing course-taking in high school, and high school graduation, possibly pointing to these two competing mechanisms offsetting each other.

Within-school segregation of students – at least to the degree observed – as a consequence of the remediation policy is not a foregone conclusion. To investigate this, we estimate tracking effects at the cutoff separately for each of the 12 school districts in this study. We find that the effects of taking a remedial ELA course on the likelihood of taking an advanced course in other subjects (math, science, or social studies) in 6th grade (estimated using a fuzzy RD design) range from +0.019 to -0.783 with a median of -0.106. Some school districts, therefore, clearly implemented the remediation policy in ways that led to much greater tracking than did others. Interestingly, when we look at the effects of a remedial ELA course on high school outcomes separately for low-tracking (i.e., districts with a tracking effect below the median of -0.106) and high-tracking districts, we find a positive effect of 0.079 (p-value: 0.078) on the likelihood of taking a college-credit bearing course in high school in any subject and a positive effect of 0.08 (p-value: 0.016) in ELA or social studies for the former group. In contrast, these effects are -0.055 (p-value: 0.305) overall and -0.043 (p-value: 0.515) in ELA and social studies for the latter group of districts. These findings provide suggestive evidence that test-based remediation could be more beneficial in settings where it does not have unintended tracking effects although it is important to note that the differences in effects on outcomes could also be driven by differences in other district attributes including student composition or educational resources.

A key challenge, therefore, for policymakers is to find opportunities to allocate educational resources in a compensatory fashion (e.g., with teachers and other classroom resources) but conducted in settings that do not exacerbate within-school segregation or place students into tracks that could inhibit subsequent success. The fact that school districts in Florida implemented this remediation policy in dramatically different ways, with some school districts substantially increasing classroom segregation for students receiving remedial resources while others did so to a much lower extent or not at all, with consequential resulting differences in cross-subject spillovers, suggests that it is possible to offer additional remedial resources to students without segregating students within schools.

References

Abdulkadiroglu, A., J. Angrist, and P. Pathak. (2014). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. Econometrica 82(1): 137-192.

Ashenfelter, O., W. Collins, and A. Yoon. (2006). Evaluating the Role of Brown v. Board of Education in School Equalization, Desegregation, and the Income of African Americans. American Law and Economics Review 8(2):213–48.

Barg, K. (2012). The influence of students' social background and parental involvement on teachers' school track choices: Reasons and consequences. *European Sociological Review.* 29(3), 565-579.

Baye, A., A. Inns, C. Lake, and R. Slavin (2019). A Synthesis of Quantitative Research on Reading Programs for Secondary Students. *Reading Research Quarterly*, *54*(2), 133–166.

Betts, J. and J. Shkolnik. (2000). The Effects of Ability Grouping on Student Achievement and Resource Allocation in Secondary Schools. Economics of Education Review 19(1): 1-15.

Bifulco, R. and Ladd, H. F. (2007). School choice, racial segregation, and test-score gaps: Evidence from North Carolina's charter school program. Journal of Policy Analysis and Management, 26(1), 31-56.

Bui, S., S. Craig, and S. Imberman. (2014). Is Gifted Education a Bright Idea? Assessing the Impact of Gifted and Talented Programs on Students. American Economic Journal: Economic Policy 6(3): 30-62.

Calonico, S., M. Cattaneo, and M. Farrell. Coverage Error Optimal Confidence Intervals for Local Polynomial Regression. Working paper.

Calonico, S., M. Cattaneo, M. Farrell, and R. Titiunik. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, 17 (2), 372–404.

Calonico, S., M. Cattaneo, and R. Titiunik. (2015). Optimal Data-Driven Regression Discontinuity Plots. Journal of the American Statistical Association 110(512): 1753-1769.

Card, D. and L. Giuliano. (2016). Can Tracking Raise the Test Scores of High-Ability Minority Students? American Economic Review 106(1): 2783-2816.

Chetty, R., J. Friedman, and J. Rockoff. (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. American Economic Review 104(9): 2593-2632.

Chmielewski, A., H. Dumont, and U. Trautwein. (2013). Tracking Effects Depend on Tracking Type: An International Comparison of Students' Mathematical Self-Concept. American Educational Research Journal 50(5): 925-957.

Chu, J., P. Loyalka, G. Li, L. Gao, and Y. Song. (2018). Stereotype Threat and Educational Tracking: A Field Experiment in Chinese Vocational High Schools. Socius.

Clark, D. (2010). Selective Schools and Academic Achievement. BE Journal of Economic Analysis and Policy 10(1).

Clotfelter, C., Hemelt, S., Ladd, H., and Turaeva, M. (2020). School Segregation at the Classroom Level in a Southern 'New Destination' State". Brown University EdWorkingPaper 20-224.

Cortes, K. and J. Goodman. (2014). The Effect of Double-Dose Algebra on Student Achievement. American Economic Review 104(5): 400-405.

Cortes, K., J. Goodman, and T. Nomi. (2015). Intensive Math Instruction and Educational Attainment: Long-Run Impacts of Double-Dose Algebra. Journal of Human Resources 50(1): 108-158.

Conger, D. (2005). "Within-school Segregation in an Urban School District," Educational Evaluation and Policy Analysis 27(3):225–44.

Dobbie, W. and R. Fryer. (2011). Are High-Quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children's Zone. American Economic Journal: Applied Economics 3(3): 158-187.

Domina, T., A. Penner, and E. Penner. (2017). Categorical Inequality: Schools as Sorting Machines. Annual Review of Sociology 43: 311-330.

Dougherty, S. (2015). Bridging the Discontinuity in Adolescent Literacy? Mixed Evidence from a Middle Grades Intervention. Education Finance and Policy 10(2): 157-192.

Duflo, E., P. Dupas, and M. Kremer. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. American Economic Review 101(5): 1739-1774.

Epple, D., E. Newlon, and R. Romano. (2002). Ability Tracking, School Competition, and the Distribution of Educational Benefits. Journal of Public Economics 83(1): 1-48.

Eren, O., Depew, B., and Barnes, S. (2017). Test-based promotion policies, dropping out, and juvenile crime. Journal of Public Economics, 153, 9–31.

Eren, O., Lovenheim, M., and Mocan, N. (2022). The Effect of Grade Retention on Adult Crime: Evidence from a Test-Based Promotion Policy. *Journal of Labor Economics, 40*(2), 361-395.

Estrada, R. and J. Gignoux. (2017). Benefits to Elite Schools and the Expected Returns to Education: Evidence from Mexico City. European Economic Review 95(C): 168-194.

Figlio, D. and Özek, U. (2020). Cross-generational differences in educational outcomes in the second great wave of immigration. *Education Finance and Policy, 15*(4), 648–674.

Figlio, D., J. Guryan, K. Karbownik, and J. Roth. (2014). The Effects of Poor Neonatal Health on Children's Cognitive Development. *American Economic Review*, 104(12): 3921-3955.

Figlio, D. and M. Page. (2002). School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality? Journal of Urban Economics 51(3): 497-514.

Frandsen, B. (2017). Party bias in union representation elections: testing the manipulation in the regression discontinuity design when the running variable is discrete. Advances in Econometrics, in : Matias D. Cattaneo and Juan Carlos Escanciano (ed.), Regression Discontinuity Designs, volume 38, pages 281-315, Emerald Publishing Ltd.

Garcia, D. (2008). The Impact of School Choice on Racial Segregation in Charter Schools. Educational Policy, 22(6), 805-829

Gershenson, S., C. Hart, J. Hyman, C. Lindsay, and N. Papageorge. (2018). The Long-Run Impacts of Same-Race Teachers. NBER Working Paper No. 25254.

Greene, J.P. and Winters, M. A. (2012). The medium-run effects of Florida's test-based promotion policy. *Education Finance and Policy,* 7(3), 305–330.

Greene, J.P. and Winters, M. A. (2007). Revisiting grade retention: An evaluation of Florida's test-based promotion policy. *Education Finance and Policy,* 2(4), 319–340.

Grissom, J. and Redding, C. (2016). Discretion and Disproportionality: Explaining the Underrepresentation of High-Achieving Students of Color in Gifted Programs. AERA Open 2 (January-March), 1-25.

Guryan, J. (2004). Desegregation and Black Dropout Rates. The American Economic Review 94(4):919–43.

Johnson, R. (2019). Children of the Dream: Why School Integration Works. New York: Basic Books.

Kalogrides, D. and S. Loeb. (2013). Different Teachers, Different Peers: the Magnitude of Student Sorting Within Schools. Educational Researcher 42: 304-316.

Kolesár, M. and C. Rothe. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, 108, 2277–2304.

Krause, A. and Schüller, S. (2014). Evidence and persistence of education inequality in an early-tracking system: The German case, in "Scuola democratica, Learning for Democracy" doi: 10.12828/77684.

Larsen, M. F. and Valant, J. (2018). Socially Promoted, Academically Retained: The Effects of Assorted Grade Retention Policies on High School and College Persistence. Unpublished Manuscript.

Lee, D. and Card, D. (2008). Regression discontinuity inference with specification error. Journal of Econometrics, 142(2), 655-674.

LiCalsi, C., U. Ozek, and D. Figlio. (2019). The uneven implementation of universal school policies: Maternal education and Florida's mandatory grade retention policy. Education Finance and Policy 14(3), 383-413.

Lucas, A. and I. Mbiti. (2014). Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya. American Economic Journal: Applied Economics 6(3): 234-263.

Marcotte, D. and K. Dalane. (2019). Socioeconomic Segregation and School Choice in American Public Schools. Educational Researcher, 48(8), 493-503.

Mariano, L. T., Martorell P., and Berglund, T. (2018). The Effects of Grade Retention on High School Outcomes: Evidence from New York City Schools. Santa Monica, CA: RAND Corporation.

McEachin, A., T. Domina, and A. Penner. (2020). Heterogeneous Effects of Early Algebra Across California Middle Schools. Journal of Policy Analysis and Management 39(3): 772-800.

Mickelson, R. (2015). "The Cumulative Disadvantages of First-and Second-Generation Segregation for Middle School Achievement," American Educational Research Journal 52, issue 4, pp. 657-692.

Nomi, T. and E. Allensworth. (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.

Orfield, G., J. Ee, E. Frankenberg, and G. Siegel-Hawley. (2016). " Brown" at 62: School Segregation by Race, Poverty and State. Civil Rights Project-Proyecto Derechos Civiles.

Özek, U. (2021). The Effects of Middle School Remediation on Postsecondary Success: Regression Discontinuity Evidence from Florida. Journal of Public Economics.

Ozier, O. (2018). The Impact of Secondary Schooling in Kenya: A Regression Discontinuity Analysis. Journal of Human Resources 53(1): 157-188.

Pop-Eleches, C. and M. Urquiola. (2013). Going to a Better School: Effects and Behavioral Responses. American Economic Review 103(4): 1289-1324.

Reardon, S., E. Weathers, E. Fahle, H. Jang, and D. Kalogrides. (2019). Is Separate Still Unequal? New Evidence on School Segregation and Racial Academic Achievement Gaps (CEPA Working Paper No.19-06). Retrieved from Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp19-06.

Reber, S. (2010). School Desegregation and Educational Attainment for Blacks. The Journal of Human Resources 45(4):893–914.

Schwerdt, G., West, M. R., and Winters, M. A. (2017). The Effects of Test-Based Retention on Student Outcomes over Time: Regression Discontinuity Evidence from Florida. Journal of Public Economics, 152, 154–169.

Taylor, E. (2014). Spending more of the school day in math class: Evidence from a regression discontinuity in middle school. *Journal of Public Economics*, 117, 162-181.

Weiher, G. and K. Tedin. (2001). Does choice lead to racially distinctive schools? Charter schools and household preferences. Journal of Policy Analysis and Management, 21(1), 79-92.

Wouters, S.. H. De Fraine, J. Colpin, J. Van Damme, and K. Verschueren. (2012). The Effect of Track Changes on the Development of Academic Self-Concept in High School: A Dynamic Test of the Big-Fish-Little-Pond Effect. Journal of Educational Psychology 104(3): 793-806.

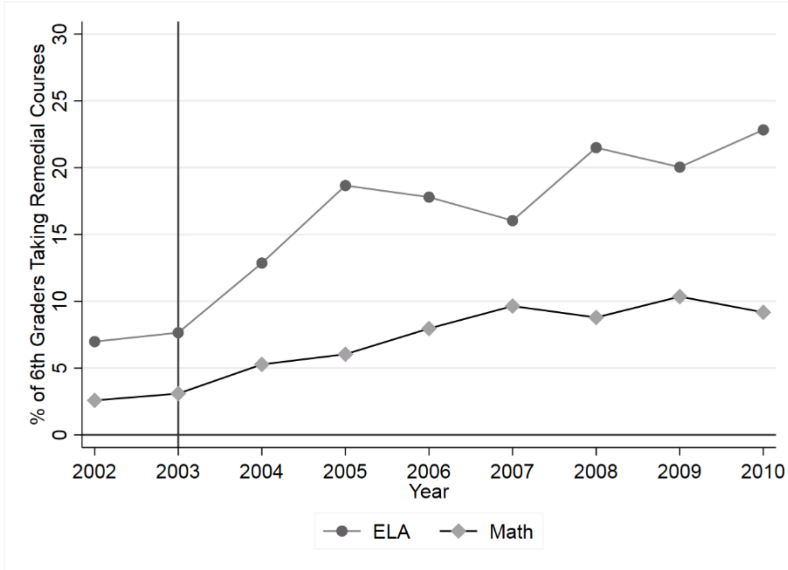
Table 1. Student Characteristics by Prior Scores, 6th Graders between 2004-05 and 2010-11

	Reading		Math	
	Below cutoff	Above cutoff	Below cutoff	Above cutoff
Prior year reading score	-1.033	0.516	-0.651	0.467
	(0.673)	(0.638)	(0.839)	(0.776)
Prior year math score	-0.709	0.422	-0.828	0.641
	(0.902)	(0.742)	(0.723)	(0.559)
Subsidized meal eligible	0.733	0.456	0.706	0.439
	(0.443)	(0.498)	(0.456)	(0.496)
Prior year disciplinary incident	0.115	0.042	0.107	0.038
	(0.320)	(0.200)	(0.309)	(0.191)
Prior year % absent days	0.049	0.041	0.050	0.038
	(0.052)	(0.043)	(0.054)	(0.041)
White	0.242	0.453	0.267	0.463
	(0.428)	(0.498)	(0.442)	(0.499)
Black	0.405	0.199	0.396	0.179
	(0.491)	(0.399)	(0.489)	(0.383)
Hispanic	0.307	0.273	0.291	0.279
	(0.461)	(0.445)	(0.454)	(0.449)
Male	0.547	0.482	0.482	0.515
	(0.498)	(0.500)	(0.500)	(0.500)
Prior year - special education	0.235	0.071	0.203	0.071
	(0.424)	(0.257)	(0.403)	(0.257)
Prior year – English learner	0.126	0.024	0.098	0.029
	(0.332)	(0.153)	(0.298)	(0.167)
Foreign born	0.139	0.106	0.123	0.112
	(0.346)	(0.308)	(0.328)	(0.316)
English non-native	0.351	0.295	0.324	0.305
	(0.477)	(0.456)	(0.468)	(0.460)
Number of unique students	211,897	463,620	263,627	411,804
Maternal characteristics -				
Less than HS diploma	0.347	0.167	0.328	0.154
	(0.476)	(0.373)	(0.470)	(0.361)
College degree or higher	0.060	0.214	0.064	0.231
	(0.237)	(0.410)	(0.245)	(0.422)
Married at birth	0.443	0.678	0.458	0.700
	(0.497)	(0.467)	(0.498)	(0.458)
Teenage pregnancy	0.163	0.0888	0.162	0.080
	(0.370)	(0.284)	(0.368)	(0.272)
Number of unique students	103,975	252,605	132,979	223,601

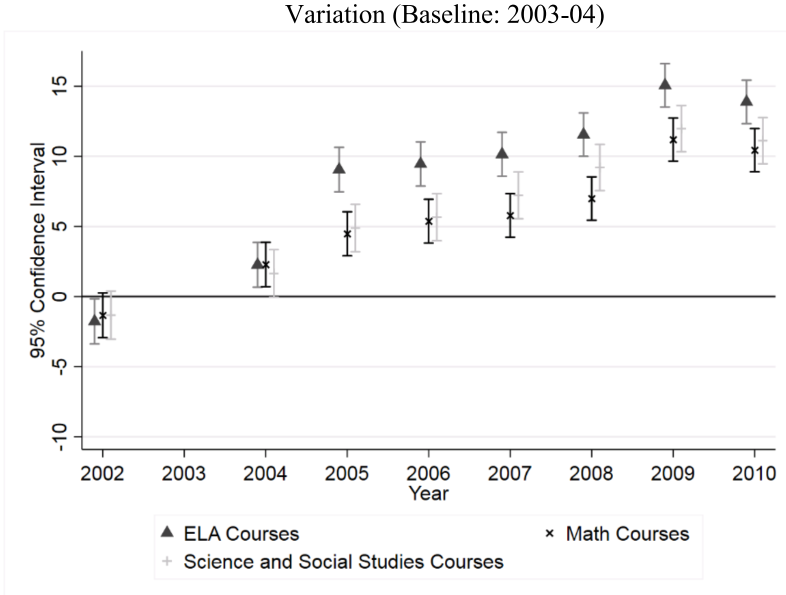
Notes: Standard deviations are given in parentheses. Reading and math scores are standardized at the grade-year level to zero mean and unit variance.

Figure 1. Share of 6th Graders Taking Remedial ELA and Math Courses and Across-Classroom Variation in Prior Year Achievement Over Time

(A) Share of 6th Graders Taking Remedial ELA and Math Courses

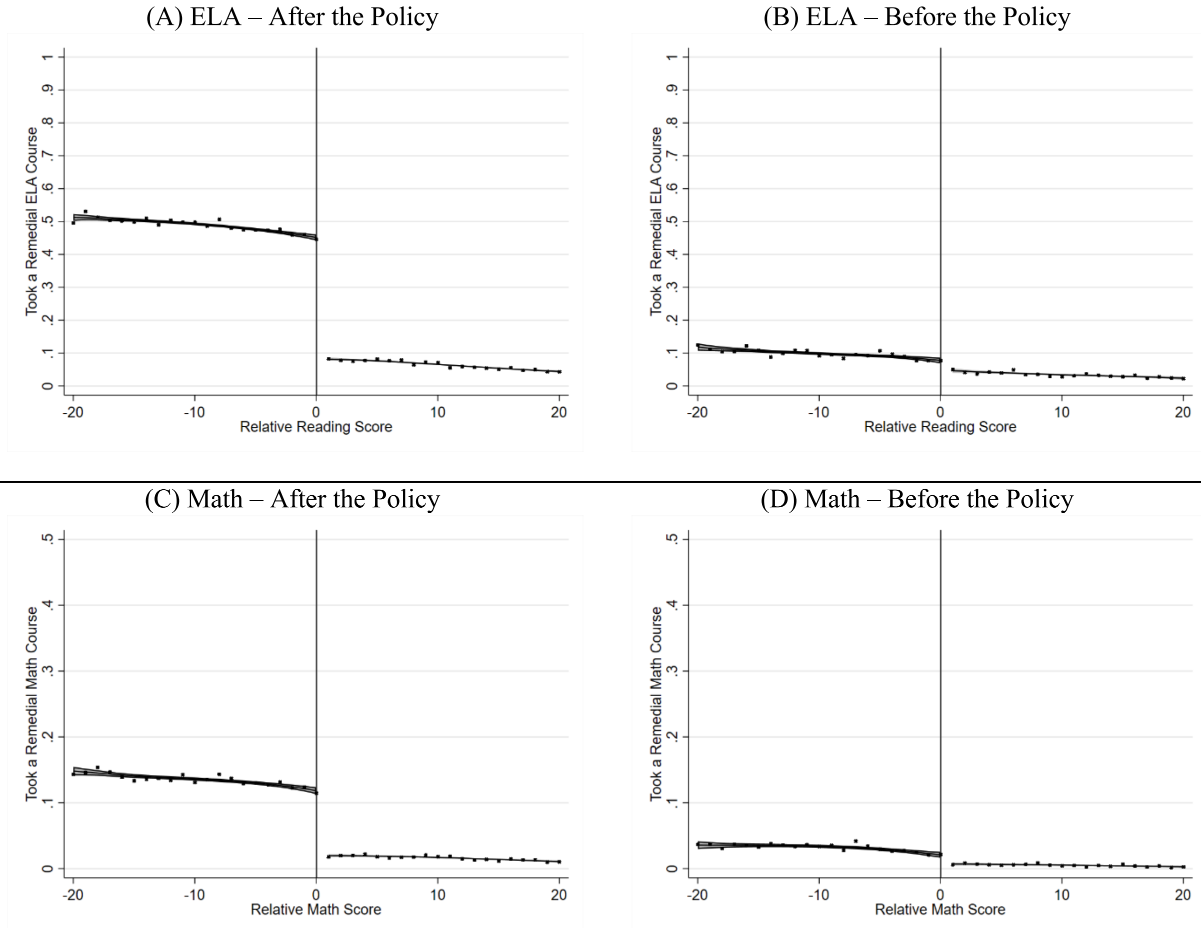


(B) Share of Within-School Variation in Prior Year Achievement Explained by Across-Classroom Variation (Baseline: 2003-04)



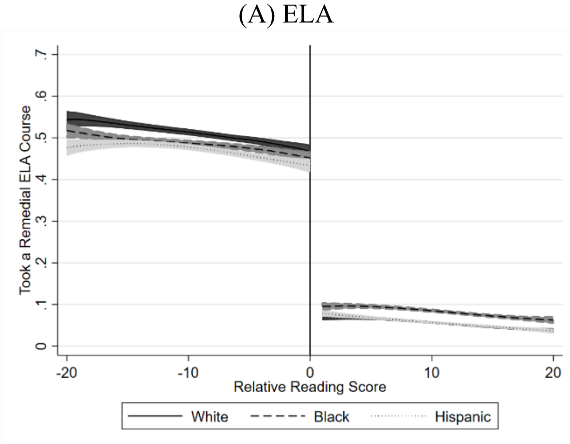
Notes: The upper panel presents the share of 6th graders taking a remedial course in ELA or math between 2002-03 and 2010-11. The bottom panel presents the share of within-school variation in student prior achievement that is explained by across-classroom variation over the same time frame. In this exercise, for each school-year, we first obtain the R^2 from regressions where we regress the averaged prior year test scores of 6th graders on classroom fixed effects in ELA, math, and science/social studies classrooms. We then regress these R^2s on year indicators -- with the year before the policy (2003-04 school year) serving as the baseline -- to examine how classroom achievement segregation changed over time by subject. Panel (B) plots the estimated coefficients on these year indicators (along with their 95% confidence intervals obtained using standard errors clustered at the school level).

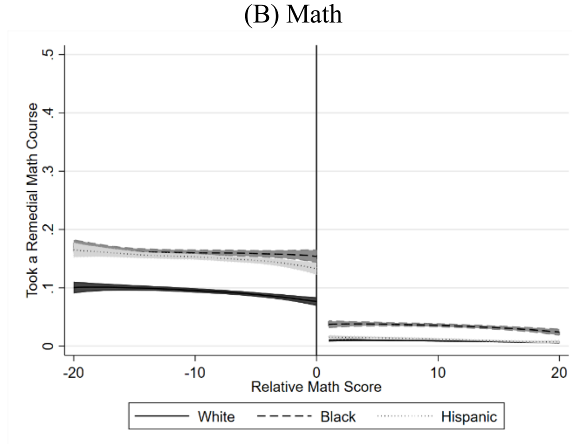
Figure 2. Remedial Course-Taking Around the Cutoff in ELA and Math, Before and After the Policy



Notes: The figures present the local linear smoothing of remedial course-taking indicators in the corresponding subject in 6th grade on relative prior year test score of the student separately for the left of the corresponding cutoff and the right. The triangle kernel and a bandwidth of 10 points are used in the estimation. The solid circles represent raw cell means and the shaded areas represent 95 percent confidence intervals.

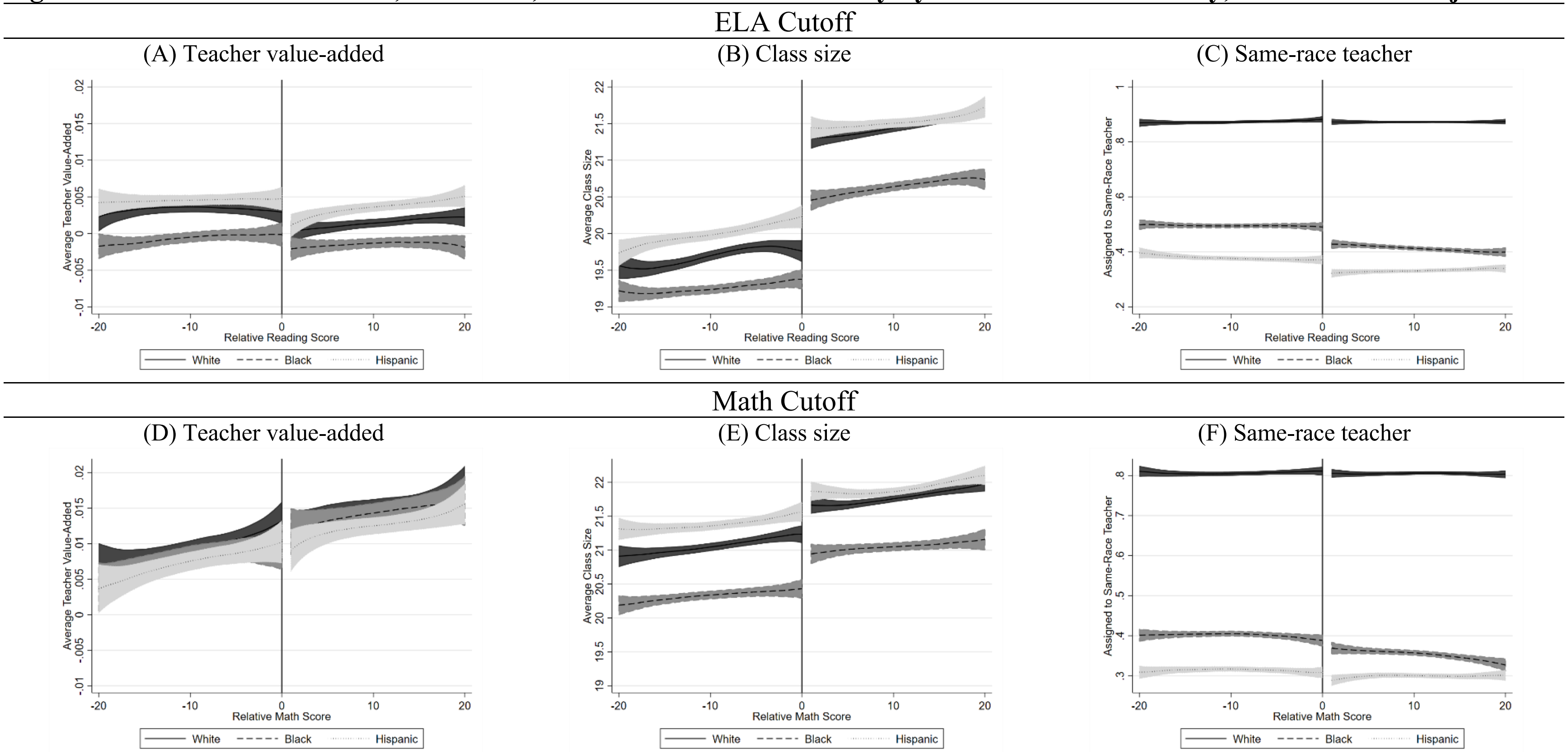
Figure 3 - Remedial Course-Taking Around the Cutoff in ELA and Math, by Race/Ethnicity





Notes: The figures present the local linear smoothing of remedial course-taking indicators in the corresponding subject in 6th grade on relative prior year test score of the student separately for the left of the corresponding cutoff and the right, broken down by student race/ethnicity. The triangle kernel and a bandwidth of 10 points are used in the estimation. The shaded areas represent 95 percent confidence intervals.

Figure 4. Teacher Value-Added, Class Size, and Teacher Race/Ethnicity by Student Race/Ethnicity, Remediation Subject

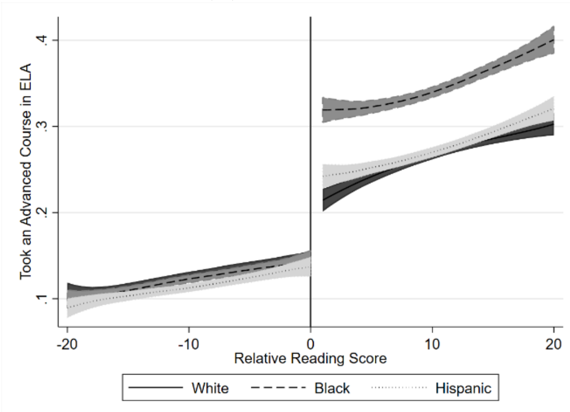


Notes: The figures present the local linear smoothing of average teacher value-added scores, average class size, and the likelihood of being assigned to at least one same-race/ethnicity teacher on relative prior year test score of the student separately for the left of the corresponding cutoff and the right, broken down by student race/ethnicity in the remediation subject. The triangle kernel and a bandwidth of 10 points are used in the estimation. The shaded areas represent 95 percent confidence intervals.

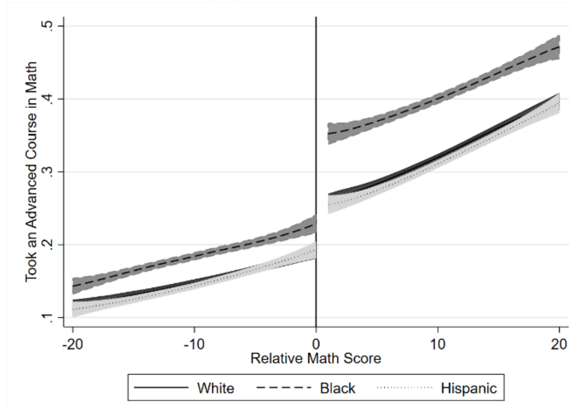
Figure 5. Advanced Course-Taking in 6th Grade Around the Remediation Cutoffs, by Race/Ethnicity and Course Subject

Remediation subject

(A) ELA cutoff

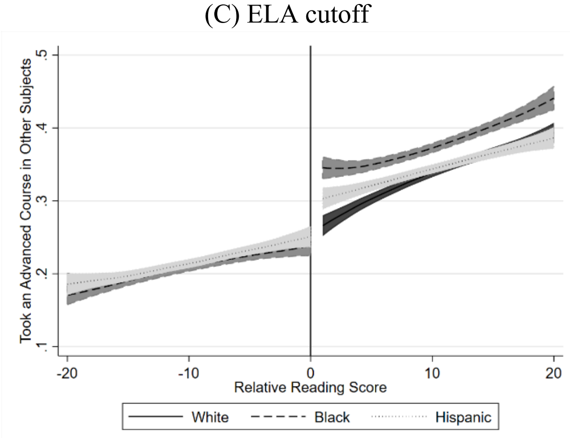


(B) Math cutoff

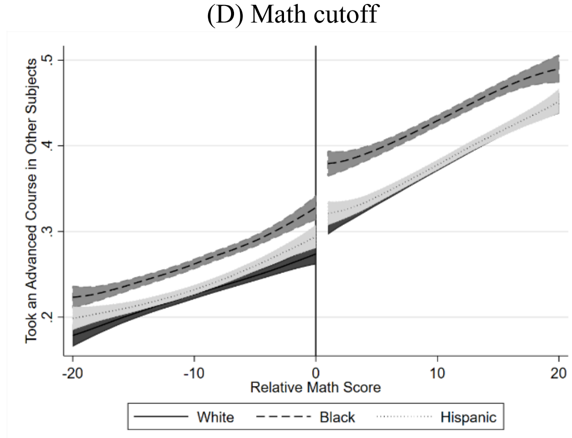


Other core subjects

(C) ELA cutoff



(D) Math cutoff

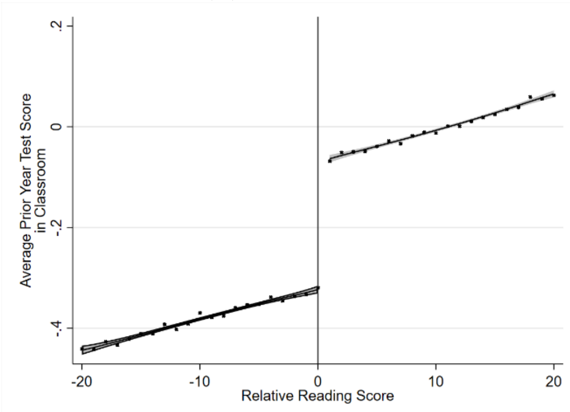


Notes: The figures present the local linear smoothing of advanced course-taking indicator in the remediation subject and other core subjects in 6th grade on relative prior year reading (first column) and math score (second column) broken down by student race/ethnicity. The triangle kernel and a bandwidth of 10 points are used in the estimation. The solid circles represent raw cell means and the shaded areas represent 95 percent confidence intervals.

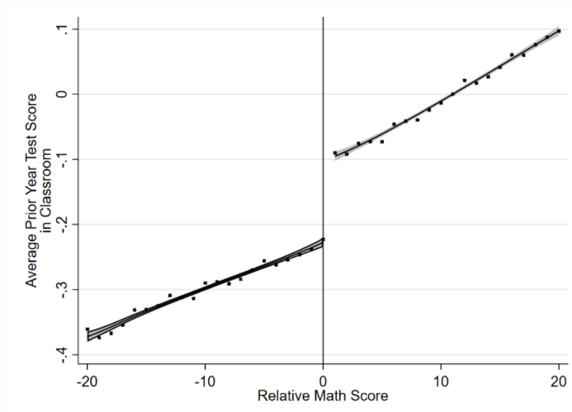
Figure 6. Average 6th Grade Classroom-Peer Prior Achievement Around the Remediation Cutoffs in ELA and Math in Remediation and Other Core Subjects, After the Policy

Remediation subject

(A) ELA cutoff

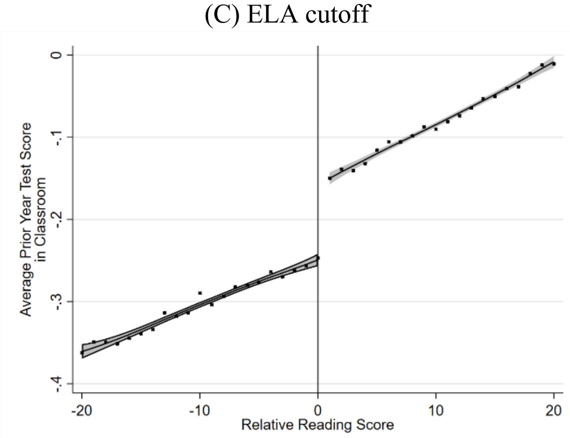


(B) Math cutoff

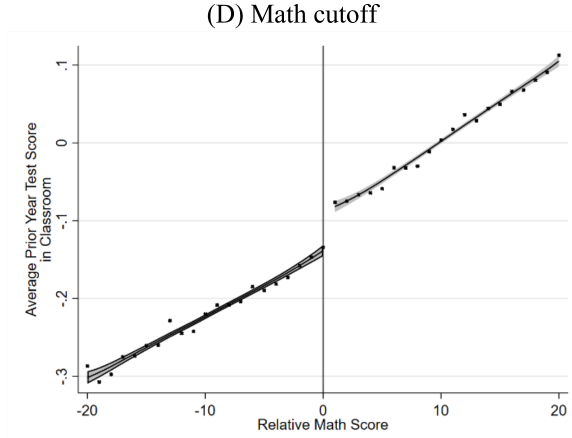


Other core subjects

(C) ELA cutoff



(D) Math cutoff

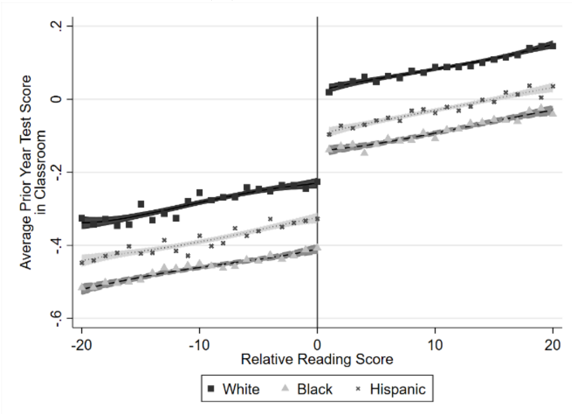


Notes: The figures present the local linear smoothing of average classroom-peer prior achievement in the remediation subject and other core subjects in 6th grade on relative prior year reading (first column) and math score (second column) of the student after the policy separately for the left of the corresponding cutoff and the right. The triangle kernel and a bandwidth of 10 points are used in the estimation. The solid circles represent raw cell means and the shaded areas represent 95 percent confidence intervals.

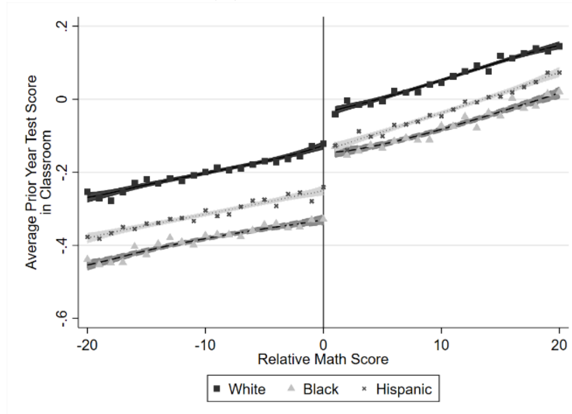
Figure 7. Average 6th Grade Classroom-Peer Prior Achievement Around the Remediation Cutoffs in ELA and Math, by Race/Ethnicity and Course Subject

Remediation subject

(A) ELA cutoff

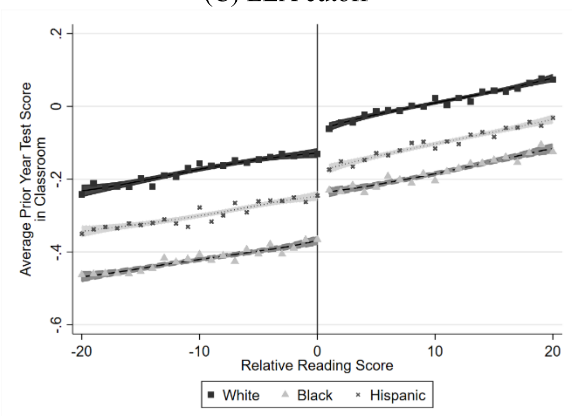


(B) Math cutoff

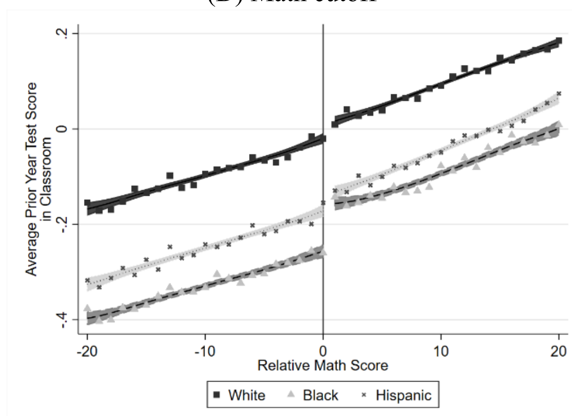


Other core subjects

(C) ELA cutoff



(D) Math cutoff

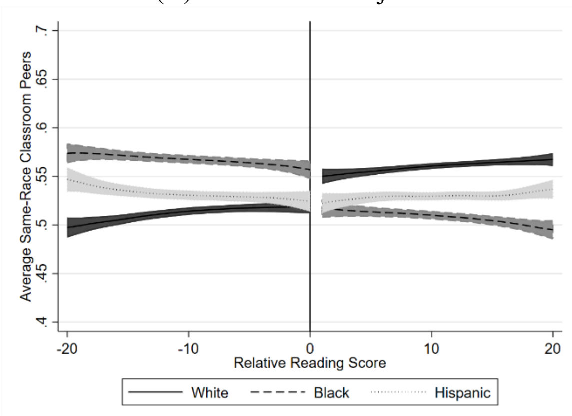


Notes: The figures present the local linear smoothing of average classroom-peer prior achievement in the remediation subject and other core subjects in 6th grade on relative prior year reading (first column) and math score (second column) of the student, by student race/ethnicity. The triangle kernel and a bandwidth of 10 points are used in the estimation. The solid circles represent raw cell means and the shaded areas represent 95 percent confidence intervals.

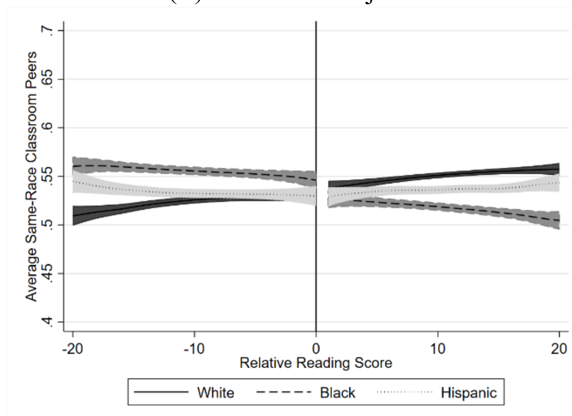
Figure 8. Average 6th Grade Same-Race/Ethnicity Classroom-Peers in Subject of Remediation and Other Core Subjects Around the Remediation Cutoffs in ELA and Math, by Race/Ethnicity

ELA cutoff

(A) Remediation subject

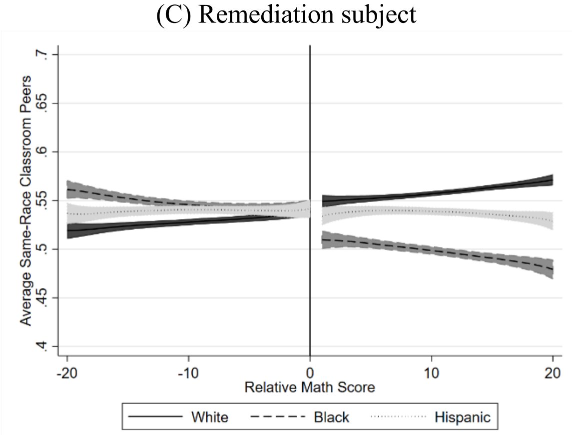


(B) Other core subjects

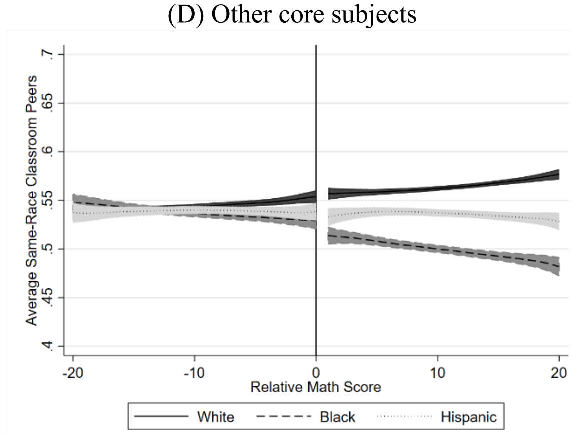


Math cutoff

(C) Remediation subject



(D) Other core subjects

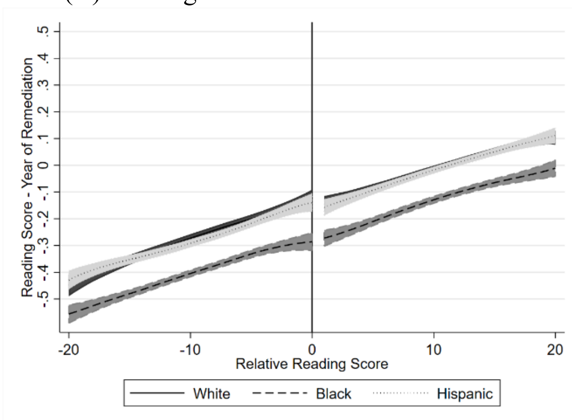


Notes: The figures present the local linear smoothing of average same-race/ethnicity classroom-peers in the subject of remediation in panels (A) and (C) and other core subjects in panels (B) and (D) in 6th grade on relative prior year reading (first column) and math score (second column) of the broken down by own race/ethnicity. The triangle kernel and a bandwidth of 10 points are used in the estimation. The solid circles represent raw cell means and the shaded areas represent 95 percent confidence intervals.

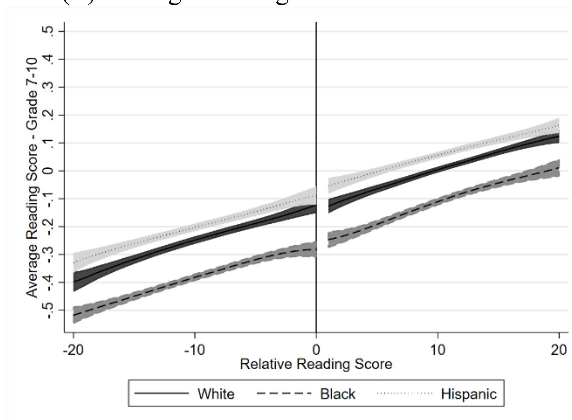
Figure 9 – Test Scores Around the Remediation Cutoffs in ELA and Math, by Race/Ethnicity

ELA Cutoff

(A) Reading Scores: Year of Remediation

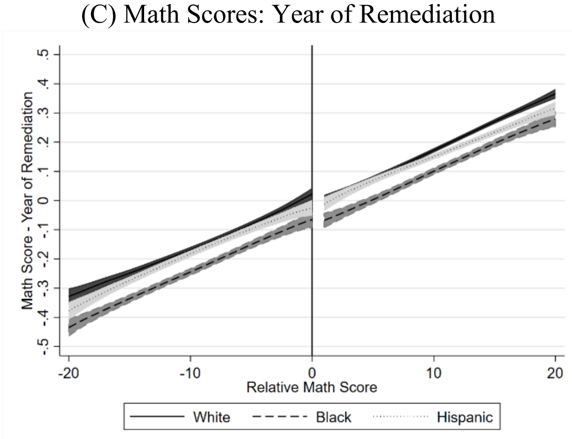


(B) Average Reading Scores: Grades 7 - 10

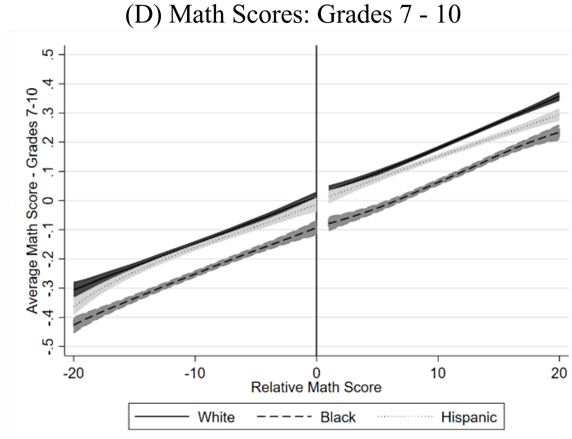


Math Cutoff

(C) Math Scores: Year of Remediation



(D) Math Scores: Grades 7 - 10

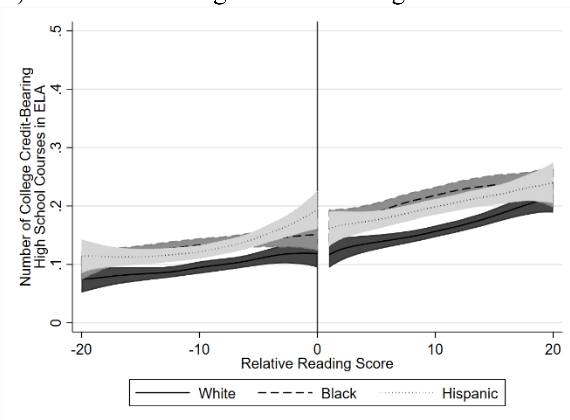


Notes: The figures present the local linear smoothing of standardized test scores in the corresponding subject in the year of remediation and averaged over grades 7 through 10 later on relative prior year test score of the student separately for the left of the corresponding cutoff and the right, broken down by student race/ethnicity. The triangle kernel and a bandwidth of 10 points are used in the estimation. The shaded areas represent 95 percent confidence intervals.

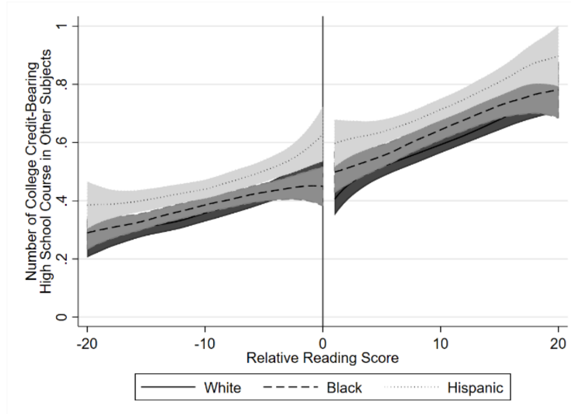
Figure 10 – Number of College-Credit Bearing Courses in High School Around the Remediation Cutoffs in ELA and Math, by Race/Ethnicity

ELA Cutoff

(A) Number of College-Credit Bearing Courses in ELA

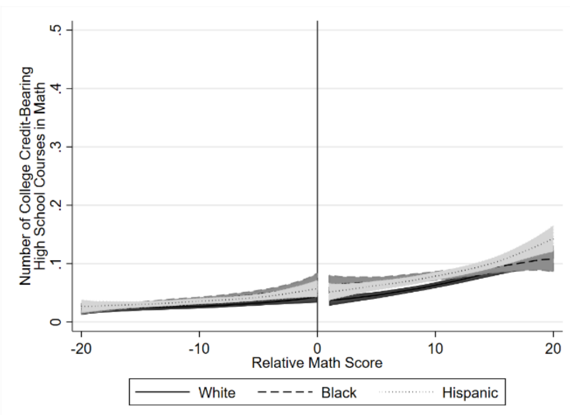


(B) Number of College-Credit Bearing Courses in Other Subjects

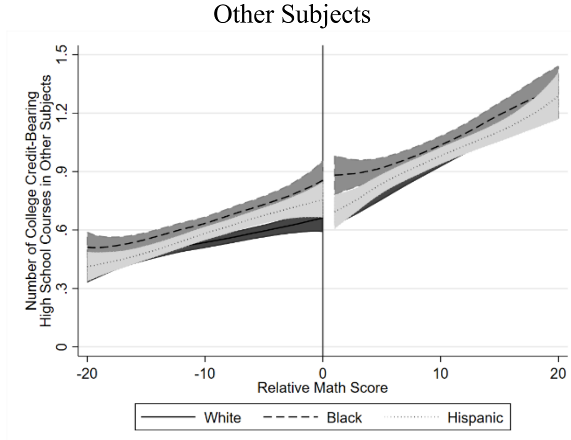


Math Cutoff

(C) Number of College-Credit Bearing Courses in Math



(D) Number of College-Credit Bearing Courses in Other Subjects

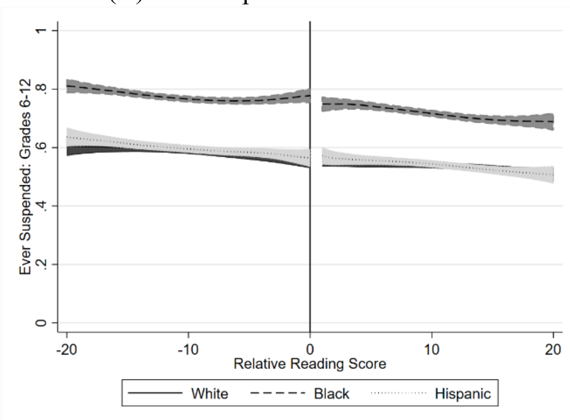


Notes: The figures present the local linear smoothing of college-credit bearing course-taking indicators in the corresponding subject on relative prior year test score of the student separately for the left of the corresponding cutoff and the right, broken down by student race/ethnicity. The triangle kernel and a bandwidth of 10 points are used in the estimation. The shaded areas represent 95 percent confidence intervals.

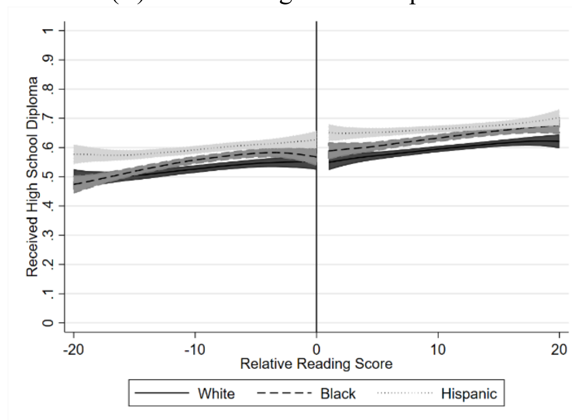
Figure 11 – Student Mobility and High School Graduation Around the Remediation Cutoffs in ELA and Math, by Race/Ethnicity

ELA Cutoff

(A) Ever suspended: Grades 6-12

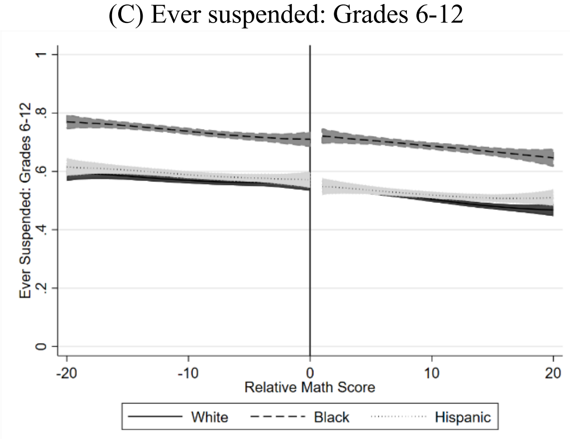


(B) Received high school diploma

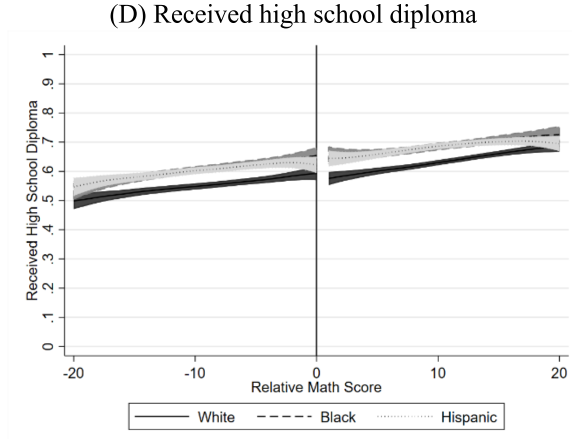


Math Cutoff

(C) Ever suspended: Grades 6-12



(D) Received high school diploma



Notes: The figures present the local linear smoothing of student mobility indicators in panels (A) and (C) that equal 1 if the student ever received a suspension in grades 6-12 and the likelihood of receiving a high school diploma on relative prior year test score of the student separately for the left of the corresponding cutoff and the right, broken down by student race/ethnicity. The triangle kernel and a bandwidth of 10 points are used in the estimation. The shaded areas represent 95 percent confidence intervals.

The Unintended Consequences of Test-Based Remediation
David Figlio and Umut Ozek
Online Appendix

Online Appendix Table 1. Baseline Equivalency of Student Characteristics Around the Remediation Cutoff

	Reading Cutoff	Math Cutoff
Prior year score - other subject	-0.011*	0.000
	(0.006)	(0.007)
Subsidized meal eligible	0.005	-0.004
	(0.004)	(0.004)
Prior year disciplinary incident	0.005	0.000
	(0.003)	(0.002)
Prior year % absent days	0.000	-0.000
	(0.000)	(0.001)
White	-0.010*	0.009*
	(0.005)	(0.005)
Black	0.010	-0.002
	(0.006)	(0.004)
Hispanic	-0.001	-0.001
	(0.006)	(0.004)
Male	0.006	0.006
	(0.006)	(0.005)
Prior year - special education	0.011**	0.001
	(0.004)	(0.004)
Prior year – English learner	-0.003	-0.002
	(0.005)	(0.003)
Foreign born	-0.003	-0.003
	(0.004)	(0.003)
English non-native	-0.002	-0.001
	(0.006)	(0.006)
N	90,414	115,092
Maternal characteristics -		
Less than HS diploma	-0.002	0.008
	(0.006)	(0.007)
College degree or higher	0.001	-0.005
	(0.003)	(0.004)
Married at birth	-0.006	-0.005
	(0.011)	(0.008)
Teenage pregnancy	-0.006	0.003
	(0.006)	(0.005)
N	48,382	61,705
Joint test of significance		
F-stat	1.06	0.87
p-value	0.38	0.60

Notes: Robust standard errors, clustered at the prior test score level, are given in parentheses. The estimates represent the discontinuities in student characteristics at the remediation cutoff, obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively.

Online Appendix Table 2. Estimated Effects on Remedial Course-Taking in 6th Grade, by Race/Ethnicity and SES Quartile

	Failed prior year reading test			
	Overall	White	Black	Hispanic
Bottom quartile	0.420***	0.563***	0.415***	0.396***
	(0.012)	(0.050)	(0.015)	(0.023)
N	19,183	1,923	10,931	5,785
Second quartile	0.426***	0.523***	0.367***	0.423***
	(0.011)	(0.027)	(0.018)	(0.020)
N	19,137	4,079	7,369	6,783
Third quartile	0.415***	0.462***	0.377***	0.392***
	(0.011)	(0.021)	(0.028)	(0.021)
N	19,150	6,430	4,935	6,563
Top quartile	0.396***	0.414***	0.389***	0.350***
	(0.011)	(0.016)	(0.031)	(0.023)
N	19,174	10,164	2,463	4,455
Overall	0.373***	0.403***	0.358***	0.357***
	(0.006)	(0.009)	(0.009)	(0.010)
N	89,905	27,697	29,776	27,391
	Failed prior year math test			
	Overall	White	Black	Hispanic
Bottom quartile	0.131***	0.078***	0.127***	0.160***
	(0.007)	(0.018)	(0.010)	(0.014)
N	24,543	2,818	13,003	7,934
Second quartile	0.108***	0.062***	0.116***	0.128***
	(0.006)	(0.009)	(0.012)	(0.011)
N	24,461	6,369	7,880	8,861
Third quartile	0.104***	0.068***	0.129***	0.133***
	(0.006)	(0.007)	(0.015)	(0.012)
N	24,474	10,087	5,158	7,530
Top quartile	0.093***	0.083***	0.095***	0.108***
	(0.006)	(0.007)	(0.019)	(0.013)
N	24,505	14,554	2,836	5,212
Overall	0.100***	0.068***	0.116***	0.123***
	(0.003)	(0.004)	(0.006)	(0.006)
N	114,475	40,427	33,403	34,024

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year test score level, are given in parentheses. The estimates represent the effect of failing the 5th grade test in reading (top panel) or math (bottom panel) on the likelihood of taking an additional remedial course in that subject in 6th grade, broken down by race/ethnicity and SES quartile, obtained using linear polynomial specification and a bandwidth of 10 points. The breakdown by SES quartiles use the subset of students who were also observed in birth records. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively.

Online Appendix Table 3. Estimated Effects on Teacher Experience, Class Size, and Teacher Race/Ethnicity by Course Subject, Student Race/Ethnicity, and SES Quartile

	Failed prior year reading test							
	Average Teacher VA Score		Assigned to a Teacher with 10+ Years of Experience		Class size		Assigned to a Same Race/Ethnicity Teacher	
	ELA	Math	ELA	Other Subjects	ELA	Other Subjects	ELA	Other Subjects
Overall	0.003***	0.002	0.039***	-0.007	-1.177***	-0.172***	0.036***	-0.003
	(0.000)	(0.001)	(0.006)	(0.005)	(0.048)	(0.036)	(0.005)	(0.005)
			[0.474]	[0.687]	[21.021]	[21.547]	[0.514]	[0.652]
Race/Ethnicity								
White	0.003***	0.004	0.013	0.002	-1.370***	-0.176***	0.000	-0.011***
	(0.001)	(0.003)	(0.010)	(0.009)	(0.087)	(0.067)	(0.006)	(0.003)
			[0.542]	[0.742]	[21.263]	[21.68]	[0.878]	[0.963]
Black	0.001	0.001	0.053***	-0.002	-1.087***	-0.199***	0.057***	0.008
	(0.001)	(0.003)	(0.009)	(0.009)	(0.083)	(0.065)	(0.010)	(0.008)
			[0.413]	[0.627]	[20.468]	[20.966]	[0.432]	[0.629]
Hispanic	0.005***	0.001	0.059***	-0.010	-1.039***	-0.114*	0.044***	-0.009
	(0.001)	(0.003)	(0.010)	(0.009)	(0.091)	(0.067)	(0.008)	(0.008)
			[0.467]	[0.690]	[21.377]	[22.042]	[0.331]	[0.457]
SES Quartile								
Bottom	0.002**	0.006	0.058***	-0.007	-1.009***	-0.137*	0.048***	-0.011
	(0.001)	(0.004)	(0.012)	(0.011)	(0.107)	(0.077)	(0.012)	(0.010)
			[0.422]	[0.640]	[19.853]	[20.551]	[0.486]	[0.664]
Second	0.001	-0.001	0.047***	0.001	-1.359***	-0.123	0.028**	-0.022**
	(0.001)	(0.004)	(0.012)	(0.011)	(0.106)	(0.082)	(0.012)	(0.010)
			[0.464]	[0.663]	[20.448]	[20.936]	[0.475]	[0.613]
Third	0.005***	0.003	0.044***	-0.004	-1.352***	-0.225***	0.032***	-0.011
	(0.001)	(0.004)	(0.012)	(0.011)	(0.101)	(0.079)	(0.012)	(0.011)
			[0.486]	[0.700]	[20.896]	[21.242]	[0.450]	[0.633]
Top	0.004***	0.007**	0.028**	-0.005	-1.389***	-0.227***	0.039***	-0.015*
	(0.001)	(0.003)	(0.013)	(0.011)	(0.100)	(0.078)	(0.010)	(0.009)
			[0.532]	[0.737]	[21.254]	[21.652]	[0.587]	[0.705]
	Failed prior year math test							
	Average Teacher VA Score		Assigned to a Teacher with 10+ Years of Experience		Class size		Assigned to a Same Race/Ethnicity Teacher	
	Math	Other subjects	Math	Other Subjects	Math	Other Subjects	Math	Other Subjects
Overall	0.000	-0.000	0.013***	0.008*	-0.362***	-0.143***	0.017***	0.002

	(0.001)	(0.001)	(0.005)	(0.004)	(0.044)	(0.031)	(0.005)	(0.004)
			[0.382]	[0.727]	[21.478]	[21.395]	[0.478]	[0.684]
Race/Ethnicity								
White	0.002	0.000	0.017**	0.004	-0.341***	-0.089*	0.009	0.004
	(0.001)	(0.001)	(0.008)	(0.007)	(0.076)	(0.051)	(0.006)	(0.003)
			[0.434]	[0.784]	[21.613]	[21.559]	[0.804]	[0.968]
Black	0.000	0.001	0.005	0.030***	-0.345***	-0.201***	0.018**	-0.005
	(0.002)	(0.001)	(0.008)	(0.009)	(0.082)	(0.062)	(0.008)	(0.008)
			[0.330]	[0.654]	[20.888]	[20.842]	[0.373]	[0.652]
Hispanic	-0.001	-0.001	0.010	-0.007	-0.456***	-0.177***	0.015**	0.004
	(0.001)	(0.001)	(0.008)	(0.008)	(0.085)	(0.059)	(0.007)	(0.007)
			[0.372]	[0.729]	[21.913]	[21.776]	[0.297]	[0.513]
SES Quartile								
Bottom	-0.002	-0.000	0.026***	0.023**	-0.189*	-0.126*	0.012	-0.002
	(0.002)	(0.001)	(0.010)	(0.009)	(0.097)	(0.071)	(0.011)	(0.009)
			[0.328]	[0.674]	[20.399]	[20.218]	[0.425]	[0.680]
Second	-0.001	-0.002	-0.007	-0.000	-0.345***	-0.094	0.028***	0.020**
	(0.002)	(0.001)	(0.010)	(0.009)	(0.095)	(0.064)	(0.010)	(0.009)
			[0.379]	[0.719]	[21.031]	[20.773]	[0.405]	[0.619]
Third	0.004**	0.001	0.013	0.004	-0.310***	-0.100	0.019*	-0.000
	(0.002)	(0.001)	(0.010)	(0.009)	(0.097)	(0.068)	(0.010)	(0.009)
			[0.396]	[0.747]	[21.200]	[21.154]	[0.480]	[0.672]
Top	-0.001	0.000	0.019*	0.003	-0.521***	-0.089	0.004	0.008
	(0.002)	(0.001)	(0.010)	(0.009)	(0.095)	(0.068)	(0.009)	(0.007)
			[0.431]	[0.777]	[21.572]	[21.533]	[0.579]	[0.726]

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year reading score level, are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively. Sample sizes are provided in Online Appendix Table 2.

Online Appendix Table 4. Estimated Effects on Advanced Course-Taking in 6th Grade, by Course Subject, Race/Ethnicity, and SES Quartile

	Failed prior year reading test			
	Took an advanced course in…			
	ELA	Math	Science	Social Studies
Overall	-0.114***	-0.033***	-0.053***	-0.058***
	(0.004)	(0.004)	(0.004)	(0.003)
	[0.264]	[0.246]	[0.221]	[0.138]
Race/Ethnicity				
White	-0.074***	-0.015*	-0.028***	-0.032***
	(0.007)	(0.008)	(0.007)	(0.005)
	[0.225]	[0.233]	[0.195]	[0.090]
Black	-0.172***	-0.055***	-0.089***	-0.100***
	(0.008)	(0.008)	(0.008)	(0.007)
	[0.320]	[0.255]	[0.249]	[0.193]
Hispanic	-0.085***	-0.022***	-0.036***	-0.040***
	(0.008)	(0.008)	(0.007)	(0.006)
	[0.236]	[0.235]	[0.213]	[0.134]
SES Quartile				
Bottom	-0.199***	-0.048***	-0.090***	-0.101***
	(0.011)	(0.010)	(0.010)	(0.009)
	[0.341]	[0.257]	[0.269]	[0.175]
Second	-0.118***	-0.039***	-0.053***	-0.062***
	(0.010)	(0.009)	(0.009)	(0.007)
	[0.273]	[0.241]	[0.227]	[0.141]
Third	-0.105***	-0.038***	-0.048***	-0.052***
	(0.009)	(0.010)	(0.009)	(0.007)
	[0.250]	[0.235]	[0.206]	[0.130]
Top	-0.088***	-0.022**	-0.046***	-0.048***
	(0.010)	(0.010)	(0.009)	(0.007)
	[0.269]	[0.271]	[0.234]	[0.138]
	Failed prior year math test			
	Took an advanced course in…			
	ELA	Math	Science	Social Studies
Overall	-0.024***	-0.076***	-0.030***	-0.023***
	(0.004)	(0.004)	(0.004)	(0.003)
	[0.280]	[0.284]	[0.243]	[0.156]
Race/Ethnicity				
White	-0.011*	-0.052***	-0.014**	-0.014***
	(0.006)	(0.006)	(0.006)	(0.004)
	[0.259]	[0.252]	[0.217]	[0.108]
Black	-0.057***	-0.128***	-0.061***	-0.050***
	(0.008)	(0.008)	(0.008)	(0.007)
	[0.332]	[0.352]	[0.289]	[0.222]
Hispanic	-0.009	-0.059***	-0.025***	-0.017***
	(0.007)	(0.007)	(0.007)	(0.006)
	[0.252]	[0.253]	[0.231]	[0.159]
SES Quartile				
Bottom	-0.042***	-0.121***	-0.051***	-0.034***
	(0.009)	(0.010)	(0.009)	(0.008)

	[0.312]	[0.339]	[0.283]	[0.180]
Second	-0.026***	-0.089***	-0.042***	-0.026***
	(0.009)	(0.009)	(0.008)	(0.007)
	[0.274]	[0.276]	[0.244]	[0.150]
Third	-0.018**	-0.063***	-0.017**	-0.022***
	(0.008)	(0.008)	(0.008)	(0.007)
	[0.286]	[0.273]	[0.242]	[0.158]
Top	-0.015*	-0.061***	-0.014*	-0.014**
	(0.009)	(0.009)	(0.008)	(0.007)
	[0.322]	[0.286]	[0.257]	[0.167]

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year reading score level, are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively. Sample sizes are provided in Online Appendix Table 2.

Online Appendix Table 5. Estimated Effects on Advanced Course-Taking in 6th Grade in the Subject of Remediation, by Race/Ethnicity and SES Quartile

	Failed prior year reading test		
	White	Black	Hispanic
Bottom quartile	-0.114***	-0.253***	-0.129***
	(0.043)	(0.015)	(0.020)
	[0.262]	[0.374]	[0.312]
Second quartile	-0.086***	-0.157***	-0.098***
	(0.018)	(0.018)	(0.017)
	[0.219]	[0.317]	[0.252]
Third quartile	-0.094***	-0.116***	-0.077***
	(0.015)	(0.022)	(0.018)
	[0.236]	[0.298]	[0.207]
Top quartile	-0.074***	-0.142***	-0.043*
	(0.013)	(0.038)	(0.023)
	[0.266]	[0.341]	[0.195]
	Failed prior year math test		
	White	Black	Hispanic
Bottom quartile	-0.082**	-0.157***	-0.088***
	(0.032)	(0.014)	(0.016)
	[0.273]	[0.377]	[0.305]
Second quartile	-0.052***	-0.146***	-0.072***
	(0.016)	(0.017)	(0.016)
	[0.219]	[0.330]	[0.263]
Third quartile	-0.059***	-0.133***	-0.027
	(0.012)	(0.023)	(0.017)
	[0.258]	[0.394]	[0.220]
Top quartile	-0.071***	-0.045	-0.044**
	(0.012)	(0.033)	(0.020)
	[0.289]	[0.337]	[0.237]

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year reading score level, are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively. Sample sizes are provided in Online Appendix Table 2.

Online Appendix Table 6. Estimated Effects on Advanced Course-Taking in 6th Grade in the Other Subjects, by Race/Ethnicity and SES Quartile

	Failed prior year reading test		
	White	Black	Hispanic
Bottom quartile	-0.058	-0.111***	-0.070***
	(0.042)	(0.015)	(0.020)
	[0.277]	[0.371]	[0.358]
Second quartile	-0.044*	-0.102***	-0.019
	(0.024)	(0.017)	(0.019)
	[0.276]	[0.329]	[0.315]
Third quartile	-0.036**	-0.090***	-0.026
	(0.018)	(0.025)	(0.021)
	[0.273]	[0.339]	[0.268]
Top quartile	-0.028*	-0.118***	-0.002
	(0.014)	(0.041)	(0.023)
	[0.321]	[0.394]	[0.266]
	Failed prior year math test		
	White	Black	Hispanic
Bottom quartile	-0.076**	-0.060***	-0.028
	(0.031)	(0.013)	(0.018)
	[0.313]	[0.393]	[0.361]
Second quartile	-0.034*	-0.077***	-0.019
	(0.020)	(0.019)	(0.016)
	[0.269]	[0.373]	[0.331]
Third quartile	-0.011	-0.077***	0.005
	(0.014)	(0.022)	(0.018)
	[0.310]	[0.442]	[0.300]
Top quartile	-0.020*	0.002	-0.008
	(0.011)	(0.030)	(0.021)
	[0.359]	[0.396]	[0.325]

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year reading score level, are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively. Sample sizes are provided in Online Appendix Table 2.

Online Appendix Table 7. Estimated Effects on Advanced Course-Taking in 6th Grade, by Race/Ethnicity, SES Quartile, and Prior Achievement Quartile

		ELA Cutoff			
		Prior Achievement Quartile			
		Bottom	Second	Third	Top
Race/Ethnicity					
	White	-0.070***	-0.060***	-0.038**	-0.043**
		(0.017)	(0.016)	(0.017)	(0.019)
		[0.170]	[0.226]	[0.285]	[0.470]
	N	4,676	6,184	7,591	9,246
	Black	-0.147***	-0.133***	-0.103***	-0.084***
		(0.015)	(0.019)	(0.021)	(0.030)
		[0.296]	[0.356]	[0.417]	[0.542]
	N	9,908	8,035	6,774	5,059
	Hispanic	-0.047***	-0.062***	-0.058***	-0.028
		(0.016)	(0.017)	(0.020)	(0.025)
		[0.215]	[0.263]	[0.350]	[0.510]
	N	6,926	7,074	6,734	6,657
SES Quartile					
	Bottom	-0.158***	-0.125***	-0.130***	-0.097**
		(0.018)	(0.022)	(0.029)	(0.047)
		[0.325]	[0.375]	[0.464]	[0.579]
	N	7,292	5,243	3,995	2,653
	Second	-0.075***	-0.114***	-0.055**	-0.099**
		(0.017)	(0.020)	(0.026)	(0.039)
		[0.236]	[0.327]	[0.407]	[0.545]
	N	5,912	5,385	4,543	3,297
	Third	-0.073***	-0.063***	-0.092***	-0.085**
		(0.020)	(0.022)	(0.025)	(0.035)
		[0.194]	[0.255]	[0.347]	[0.532]
	N	4,816	5,021	4,898	4,415
	Top	-0.081***	-0.049**	-0.074***	-0.054**
		(0.025)	(0.022)	(0.024)	(0.026)
		[0.185]	[0.239]	[0.346]	[0.556]
	N	2,941	4,342	5,382	6,509
		Math Cutoff			
Race/Ethnicity					
	White	-0.042***	-0.059***	-0.028**	-0.020
		(0.016)	(0.014)	(0.014)	(0.014)
		[0.200]	[0.276]	[0.353]	[0.468]
	N	6,636	9,079	11,074	13,638
	Black	-0.094***	-0.091***	-0.113***	-0.052**
		(0.015)	(0.016)	(0.018)	(0.025)
		[0.300]	[0.413]	[0.499]	[0.613]
	N	10,676	9,049	7,695	5,983
	Hispanic	-0.053***	-0.029*	-0.028	0.009
		(0.014)	(0.015)	(0.019)	(0.022)
		[0.233]	[0.304]	[0.401]	[0.501]
	N	9,905	8,827	8,153	7,138
SES Quartile					
	Bottom	-0.095***	-0.073***	-0.092***	-0.063
		(0.017)	(0.020)	(0.027)	(0.046)
		[0.322]	[0.430]	[0.524]	[0.622]
	N	9,925	6,636	4,942	3,040
	Second	-0.060***	-0.075***	-0.061***	-0.049

	(0.017)	(0.019)	(0.022)	(0.030)
	[0.234]	[0.345]	[0.426]	[0.539]
N	7,110	6,849	5,884	4,618
Third	-0.036*	-0.057***	-0.040*	-0.041
	(0.020)	(0.019)	(0.021)	(0.025)
	[0.219]	[0.326]	[0.405]	[0.543]
N	5,447	6,314	6,496	6,217
Top	-0.024	-0.022	-0.014	-0.022
	(0.021)	(0.018)	(0.018)	(0.019)
	[0.192]	[0.293]	[0.403]	[0.555]
N	3,582	5,653	6,997	8,273

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year reading score level, are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively.

Online Appendix Table 8. Estimated Effects on Advanced Course-Taking in 6th Grade, by Race/Ethnicity, SES Quartile, and School Prior Achievement Quartile

		ELA Cutoff			
		School Quartile Based on Averaged Student Prior Achievement			
		Bottom	Second	Third	Top
Race/Ethnicity					
	White	-0.094***	-0.053***	-0.063***	-0.015
		(0.021)	(0.018)	(0.015)	(0.015)
		[0.335]	[0.325]	[0.309]	[0.300]
	N	4,577	6,196	7,923	9,001
	Black	-0.201***	-0.109***	-0.078***	-0.041*
		(0.015)	(0.017)	(0.017)	(0.022)
		[0.455]	[0.379]	[0.333]	[0.321]
	N	10,242	7,499	6,825	5,210
	Hispanic	-0.093***	-0.057***	-0.059***	-0.017
		(0.017)	(0.016)	(0.018)	(0.018)
		[0.394]	[0.336]	[0.311]	[0.306]
	N	6,570	7,523	6,357	6,941
SES Quartile					
	Bottom	-0.183***	-0.120***	-0.069***	-0.045
		(0.016)	(0.020)	(0.026)	(0.050)
		[0.444]	[0.406]	[0.348]	[0.342]
	N	9,061	5,842	3,186	1,094
	Second	-0.128***	-0.065***	-0.059***	-0.084**
		(0.018)	(0.017)	(0.021)	(0.038)
		[0.401]	[0.341]	[0.326]	[0.345]
	N	6,479	6,358	4,645	1,655
	Third	-0.090***	-0.060***	-0.095***	-0.053**
		(0.021)	(0.019)	(0.018)	(0.026)
		[0.354]	[0.305]	[0.333]	[0.334]
	N	4,438	5,834	5,760	3,118
	Top	-0.138***	-0.082***	-0.054***	-0.045**
		(0.034)	(0.022)	(0.017)	(0.020)
		[0.409]	[0.372]	[0.325]	[0.398]
	N	2,356	4,177	6,860	5,781
		Math Cutoff			
Race/Ethnicity					
	White	-0.035	-0.056***	-0.030**	-0.020*
		(0.021)	(0.015)	(0.012)	(0.011)
		[0.391]	[0.366]	[0.351]	[0.309]
	N	4,624	8,750	11,915	15,138
	Black	-0.124***	-0.086***	-0.038**	-0.049**
		(0.013)	(0.016)	(0.017)	(0.020)
		[0.476]	[0.440]	[0.363]	[0.425]
	N	12,541	8,729	6,806	5,327
	Hispanic	-0.066***	-0.033**	0.006	0.014
		(0.014)	(0.015)	(0.016)	(0.018)
		[0.396]	[0.337]	[0.313]	[0.337]
	N	10,184	9,389	8,086	6,364
SES Quartile					
	Bottom	-0.099***	-0.058***	-0.052*	0.013
		(0.013)	(0.018)	(0.028)	(0.040)
		[0.470]	[0.405]	[0.342]	[0.292]
	N	13,376	6,622	3,226	1,319
	Second	-0.080***	-0.061***	-0.053***	-0.004

	(0.016)	(0.017)	(0.021)	(0.025)
	[0.404]	[0.360]	[0.363]	[0.309]
N	7,840	7,980	5,703	2,938
Third	-0.086***	-0.040**	-0.013	-0.034*
	(0.022)	(0.017)	(0.016)	(0.020)
	[0.416]	[0.372]	[0.354]	[0.365]
N	4,889	7,105	7,329	5,151
Top	0.017	-0.079***	0.011	-0.041***
	(0.037)	(0.021)	(0.015)	(0.015)
	[0.336]	[0.404]	[0.356]	[0.431]
N	2,043	5,369	8,412	8,681

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year reading score level, are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively.

Online Appendix Table 9. Estimated Effects on Course-Taking in 7th and 8th Grade, by Race/Ethnicity and SES Quartile

	Failed prior year reading test					
	7th Grade			8th Grade		
	Remedial: ELA	Advanced: ELA	Advanced: Other Subjects	Remedial: ELA	Advanced: ELA	Advanced: Other Subjects
Overall	0.013**	-0.025***	-0.018***	-0.003	-0.009	-0.006
	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
	[0.327]	[0.262]	[0.350]	[0.293]	[0.292]	[0.372]
Race/Ethnicity						
White	0.002	-0.011	-0.002	-0.010	-0.005	-0.012
	(0.011)	(0.008)	(0.009)	(0.013)	(0.010)	(0.011)
	[0.307]	[0.247]	[0.329]	[0.284]	[0.277]	[0.360]
Black	0.039***	-0.047***	-0.040***	0.011	-0.029***	-0.016
	(0.011)	(0.009)	(0.010)	(0.012)	(0.011)	(0.012)
	[0.369]	[0.276]	[0.349]	[0.331]	[0.315]	[0.380]
Hispanic	0.000	-0.015*	-0.005	-0.002	-0.002	0.005
	(0.011)	(0.009)	(0.010)	(0.012)	(0.011)	(0.012)
	[0.317]	[0.250]	[0.356]	[0.266]	[0.278]	[0.364]
SES Quartile						
Bottom	0.033**	-0.049***	-0.045***	-0.002	-0.016	-0.001
	(0.016)	(0.011)	(0.013)	(0.018)	(0.015)	(0.016)
	[0.403]	[0.279]	[0.364]	[0.389]	[0.291]	[0.366]
Second	0.023*	-0.023**	-0.018	-0.011	-0.012	0.004
	(0.014)	(0.011)	(0.011)	(0.017)	(0.014)	(0.015)
	[0.358]	[0.255]	[0.335]	[0.346]	[0.270]	[0.345]
Third	0.001	-0.027**	-0.019	0.008	0.009	0.004
	(0.014)	(0.011)	(0.012)	(0.017)	(0.014)	(0.013)
	[0.325]	[0.251]	[0.342]	[0.299]	[0.262]	[0.346]
Top	0.004	-0.025**	-0.016	-0.018	-0.011	-0.008
	(0.013)	(0.011)	(0.012)	(0.014)	(0.013)	(0.013)
	[0.260]	[0.301]	[0.390]	[0.249]	[0.325]	[0.409]

	Failed prior year math test					
	7th Grade			8th Grade		
	Remedial: Math	Advanced: Math	Advanced: Other Subjects	Remedial: Math	Advanced: Math	Advanced: Other Subjects
Overall	-0.001	-0.022***	-0.010**	0.001	-0.005	-0.014**
	(0.003)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)
	[0.117]	[0.254]	[0.380]	[0.082]	[0.124]	[0.428]
Race/Ethnicity						

White	-0.005	-0.011	-0.010	0.003	-0.005	-0.013
	(0.005)	(0.007)	(0.008)	(0.005)	(0.007)	(0.009)
	[0.090]	[0.239]	[0.355]	[0.058]	[0.141]	[0.399]
Black	-0.001	-0.028***	-0.015	0.003	-0.019**	-0.025**
	(0.007)	(0.009)	(0.009)	(0.007)	(0.007)	(0.011)
	[0.141]	[0.267]	[0.400]	[0.105]	[0.130]	[0.458]
Hispanic	-0.000	-0.025***	-0.008	-0.001	-0.000	-0.002
	(0.007)	(0.008)	(0.009)	(0.007)	(0.007)	(0.011)
	[0.131]	[0.248]	[0.377]	[0.090]	[0.100]	[0.422]
SES Quartile						
Bottom	0.000	-0.025**	-0.014	-0.010	-0.018*	-0.022
	(0.009)	(0.010)	(0.011)	(0.010)	(0.009)	(0.014)
	[0.162]	[0.263]	[0.386]	[0.120]	[0.112]	[0.434]
Second	-0.004	-0.018**	-0.012	0.002	-0.003	-0.013
	(0.008)	(0.009)	(0.010)	(0.008)	(0.009)	(0.013)
	[0.128]	[0.243]	[0.366]	[0.089]	[0.114]	[0.414]
Third	0.003	-0.026***	-0.007	0.001	-0.005	-0.016
	(0.008)	(0.010)	(0.010)	(0.008)	(0.009)	(0.012)
	[0.110]	[0.251]	[0.385]	[0.073]	[0.127]	[0.431]
Top	0.002	-0.015	-0.011	0.010	0.013	-0.001
	(0.007)	(0.009)	(0.010)	(0.007)	(0.009)	(0.012)
	[0.094]	[0.254]	[0.420]	[0.058]	[0.136]	[0.456]

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year reading score level, are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively. Sample sizes are provided in Online Appendix Table 2.

Online Appendix Table 10. Estimated Effects of Taking a Remedial Course in 6th Grade on Advanced Course-Taking in Middle School, by Race/Ethnicity and SES Quartile, 2SLS Estimates

	Failed prior year reading test					
	6th grade		7th grade		8th grade	
	Advanced: ELA	Advanced: Other Subjects	Advanced: ELA	Advanced: Other Subjects	Advanced: ELA	Advanced: Other Subjects
Overall	-0.307***	-0.141***	-0.066***	-0.043***	-0.039**	-0.029
	(0.012)	(0.010)	(0.009)	(0.014)	(0.016)	(0.022)
	[0.262]	[0.309]	[0.262]	[0.348]	[0.294]	[0.371]
Race/Ethnicity						
Black	-0.490***	-0.285***	-0.149***	-0.122***	-0.120***	-0.092***
	(0.022)	(0.025)	(0.024)	(0.023)	(0.027)	(0.030)
	[0.321]	[0.343]	[0.280]	[0.350]	[0.324]	[0.388]
White or Hispanic	-0.214***	-0.066***	-0.024	-0.009	-0.010	-0.008
	(0.014)	(0.014)	(0.017)	(0.021)	(0.017)	(0.023)
	[0.229]	[0.285]	[0.245]	[0.340]	[0.275]	[0.358]
SES Quartile						
Bottom	-0.482***	-0.264***	-0.143***	-0.145***	-0.081*	-0.086
	(0.021)	(0.020)	(0.022)	(0.021)	(0.046)	(0.055)
	[0.344]	[0.363]	[0.287]	[0.373]	[0.299]	[0.383]
Top	-0.233***	-0.119***	-0.073**	-0.047**	-0.051	-0.063*
	(0.021)	(0.020)	(0.030)	(0.020)	(0.044)	(0.033)
	[0.268]	[0.329]	[0.301]	[0.388]	[0.327]	[0.409]
	Failed prior year math test					
	Advanced: Math	Advanced: Other Subjects	Advanced: Math	Advanced: Other Subjects	Advanced: Math	Advanced: Other Subjects
Overall	-0.785***	-0.309***	-0.232***	-0.103*	-0.070*	-0.129**
	(0.033)	(0.061)	(0.055)	(0.053)	(0.039)	(0.056)
	[0.286]	[0.331]	[0.255]	[0.380]	[0.124]	[0.428]
Race/Ethnicity						
Black	-0.999***	-0.372***	-0.219***	-0.056	-0.141*	-0.118
	(0.084)	(0.113)	(0.080)	(0.076)	(0.078)	(0.121)
	[0.354]	[0.376]	[0.266]	[0.397]	[0.126]	[0.453]
White or Hispanic	-0.671***	-0.273***	-0.232***	-0.148***	-0.039	-0.127**
	(0.035)	(0.060)	(0.067)	(0.057)	(0.049)	(0.064)
	[0.254]	[0.310]	[0.245]	[0.368]	[0.123]	[0.412]
SES Quartile						
Bottom	-0.853***	-0.299*	-0.242***	-0.063	-0.227***	-0.078

	(0.116)	(0.157)	(0.085)	(0.106)	(0.073)	(0.147)
	[0.340]	[0.363]	[0.265]	[0.381]	[0.119]	[0.425]
Top	-0.761***	-0.242***	-0.171*	-0.169**	0.047	-0.169
	(0.095)	(0.062)	(0.092)	(0.081)	(0.071)	(0.112)
	[0.296]	[0.365]	[0.262]	[0.430]	[0.141]	[0.467]

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year reading score level, are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively. Sample sizes are provided in Online Appendix Table 2.

Online Appendix Table 11. Estimated Effects on Average Classroom-Peer 5th Grade Achievement in Middle School, by Course Subject, Race/Ethnicity, and SES Quartile

	Failed prior year reading test					
	6th grade		7th grade		8th grade	
	ELA	Other subjects	ELA	Other subjects	ELA	Other subjects
Overall	-0.249***	-0.090***	-0.031***	-0.018***	0.001	0.002
	(0.005)	(0.004)	(0.006)	(0.005)	(0.007)	(0.006)
Race/Ethnicity						
White	-0.248***	-0.064***	-0.031***	-0.016*	0.011	-0.002
	(0.009)	(0.008)	(0.011)	(0.009)	(0.013)	(0.011)
Black	-0.268***	-0.133***	-0.053***	-0.034***	-0.015	-0.009
	(0.009)	(0.008)	(0.012)	(0.010)	(0.013)	(0.012)
Hispanic	-0.216***	-0.061***	-0.003	-0.001	-0.000	0.013
	(0.009)	(0.008)	(0.011)	(0.010)	(0.013)	(0.011)
SES Quartile						
Bottom	-0.303***	-0.130***	-0.039**	-0.033**	-0.023	-0.019
	(0.011)	(0.010)	(0.016)	(0.014)	(0.019)	(0.017)
Second	-0.273***	-0.103***	-0.037***	-0.012	-0.018	0.011
	(0.011)	(0.010)	(0.014)	(0.012)	(0.017)	(0.015)
Third	-0.276***	-0.085***	-0.037***	-0.024**	0.030*	0.004
	(0.011)	(0.009)	(0.014)	(0.011)	(0.016)	(0.015)
Top	-0.258***	-0.075***	-0.042***	-0.023**	0.018	0.005
	(0.010)	(0.009)	(0.014)	(0.011)	(0.015)	(0.013)
	Failed prior year math test					
	6th grade		7th grade		8th grade	
	Math	Other subjects	Math	Other subjects	Math	Other subjects
Overall	-0.122***	-0.046***	-0.018***	-0.008	-0.007	-0.012**
	(0.005)	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)
Race/Ethnicity						
White	-0.090***	-0.025***	-0.018**	0.002	-0.020**	-0.017**
	(0.007)	(0.006)	(0.009)	(0.008)	(0.010)	(0.008)
Black	-0.176***	-0.088***	-0.027**	-0.016*	0.001	-0.018
	(0.009)	(0.008)	(0.011)	(0.010)	(0.014)	(0.012)
Hispanic	-0.115***	-0.037***	-0.007	-0.011	-0.005	0.006
	(0.009)	(0.007)	(0.010)	(0.009)	(0.012)	(0.010)
SES Quartile						

Bottom	-0.171***	-0.060***	-0.018	-0.025**	0.002	-0.008
	(0.012)	(0.009)	(0.014)	(0.012)	(0.017)	(0.015)
Second	-0.142***	-0.059***	-0.012	-0.014	-0.007	-0.029**
	(0.010)	(0.008)	(0.012)	(0.011)	(0.015)	(0.013)
Third	-0.104***	-0.041***	-0.018	0.009	-0.003	-0.006
	(0.009)	(0.008)	(0.011)	(0.011)	(0.014)	(0.012)
Top	-0.117***	-0.027***	-0.028**	-0.014	0.001	-0.007
	(0.009)	(0.008)	(0.011)	(0.010)	(0.014)	(0.011)

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year reading score level, are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively. Sample sizes are provided in Online Appendix Table 2.

Online Appendix Table 12. Estimated Effects on Average Classroom-Peer SES in Middle School, by Course Subject and SES Quartile

	Failed prior year reading test					
	6th grade		7th grade		8th grade	
	ELA	Other subjects	ELA	Other subjects	ELA	Other subjects
Overall	-0.073***	-0.028***	-0.012***	-0.008*	-0.005	-0.001
	(0.003)	(0.003)	(0.005)	(0.004)	(0.006)	(0.006)
SES Quartile						
Bottom	-0.059***	-0.031***	0.004	0.004	0.012	0.012
	(0.006)	(0.005)	(0.009)	(0.008)	(0.013)	(0.012)
Second	-0.067***	-0.030***	-0.004	0.002	0.002	0.010
	(0.006)	(0.005)	(0.010)	(0.009)	(0.012)	(0.012)
Third	-0.071***	-0.024***	-0.018*	-0.015*	-0.010	-0.009
	(0.006)	(0.006)	(0.010)	(0.009)	(0.013)	(0.012)
Top	-0.092***	-0.026***	-0.031***	-0.021**	-0.011	-0.008
	(0.007)	(0.006)	(0.009)	(0.008)	(0.011)	(0.010)
	Failed prior year math test					
	6th grade		7th grade		8th grade	
	Math	Other subjects	Math	Other subjects	Math	Other subjects
Overall	-0.038***	-0.013***	-0.015***	-0.010***	-0.007	-0.006
	(0.003)	(0.002)	(0.004)	(0.004)	(0.006)	(0.005)
SES Quartile						
Bottom	-0.033***	-0.011**	-0.012	-0.011	0.000	-0.001
	(0.006)	(0.005)	(0.009)	(0.008)	(0.013)	(0.011)
Second	-0.039***	-0.020***	-0.014	-0.011	-0.006	-0.011
	(0.006)	(0.005)	(0.009)	(0.008)	(0.012)	(0.010)
Third	-0.034***	-0.012**	-0.019**	-0.012	-0.013	-0.011
	(0.006)	(0.005)	(0.009)	(0.008)	(0.012)	(0.010)
Top	-0.041***	-0.007	-0.010	-0.002	0.002	0.006
	(0.006)	(0.005)	(0.008)	(0.007)	(0.011)	(0.008)

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year reading score level, are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively. Sample sizes are provided in Online Appendix Table 2.

Online Appendix Table 13. Estimated Effects on Average Classroom-Peer Race/Ethnicity in Middle School, by Course Subject, Race/Ethnicity, and SES Quartile

	Failed prior year reading test					
	6th grade		7th grade		8th grade	
	ELA	Other subjects	ELA	Other subjects	ELA	Other subjects
% same race						
White	-0.032***	-0.009***	-0.001	-0.005	-0.000	-0.004
	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)
	[0.551]	[0.540]	[0.531]	[0.536]	[0.536]	[0.537]
Black	0.037***	0.017***	0.012**	0.012***	0.009	0.005
	(0.003)	(0.002)	(0.005)	(0.004)	(0.005)	(0.005)
	[0.520]	[0.530]	[0.541]	[0.533]	[0.533]	[0.528]
Hispanic	-0.003	-0.002	-0.003	-0.000	-0.001	0.004
	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)
	[0.528]	[0.533]	[0.532]	[0.538]	[0.537]	[0.542]
	Failed prior year math test					
	6th grade		7th grade		8th grade	
	Math	Other subjects	Math	Other subjects	Math	Other subjects
% same race						
White	-0.012***	-0.004**	-0.005	-0.000	-0.009**	-0.001
	(0.002)	(0.002)	(0.003)	(0.002)	(0.004)	(0.003)
	[0.548]	[0.555]	[0.542]	[0.550]	[0.547]	[0.555]
Black	0.029***	0.012***	0.009**	0.009**	0.005	0.005
	(0.003)	(0.002)	(0.004)	(0.004)	(0.006)	(0.005)
	[0.509]	[0.513]	[0.521]	[0.517]	[0.540]	[0.511]
Hispanic	0.003	0.002	0.003	-0.000	0.002	-0.000
	(0.002)	(0.002)	(0.003)	(0.002)	(0.004)	(0.003)
	[0.539]	[0.537]	[0.545]	[0.542]	[0.582]	[0.545]

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year reading score level, are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively. Sample sizes are provided in Online Appendix Table 2.

Online Appendix Table 14. Estimated Effects on Middle and High School Outcomes, Overall and by Race/Ethnicity

	Failed prior year reading test			
	Overall	White	Black	Hispanic
Middle school: Reading scores				
6th grade	0.031***	0.026	0.021	0.059**
	(0.012)	(0.021)	(0.023)	(0.023)
7th grade	0.002	0.009	0.030	0.001
	(0.012)	(0.023)	(0.023)	(0.023)
8th grade	0.004	0.006	0.012	0.013
	(0.013)	(0.024)	(0.025)	(0.025)
9th grade	0.003	0.017	-0.008	0.004
	(0.014)	(0.026)	(0.024)	(0.028)
10th grade	0.002	0.017	-0.026	0.016
	(0.015)	(0.028)	(0.029)	(0.030)
Changed schools by 7th grade	0.001	-0.011	0.029	-0.028*
	(0.009)	(0.015)	(0.019)	(0.017)
	[0.225]	[0.211]	[0.254]	[0.221]
Grades 6-12: ever suspended	0.006	0.007	0.022	-0.019
	(0.011)	(0.020)	(0.017)	(0.021)
	[0.618]	[0.553]	[0.743]	[0.574]
Grades 6-12: % absent days	0.001	0.002	-0.001	-0.000
	(0.001)	(0.002)	(0.003)	(0.003)
	[0.075]	[0.080]	[0.074]	[0.070]
High school: number of college credit-bearing courses in…				
All subjects	0.073**	0.065	-0.014	0.145*
	(0.037)	(0.056)	(0.064)	(0.077)
	[0.672]	[0.528]	[0.672]	[0.723]
ELA	0.015	-0.003	0.003	0.046**
	(0.011)	(0.017)	(0.021)	(0.022)
	[0.152]	[0.119]	[0.166]	[0.154]
Math	0.001	0.011	-0.012	-0.007
	(0.007)	(0.010)	(0.010)	(0.015)
	[0.064]	[0.043]	[0.056]	[0.080]
Science	0.020**	0.014	0.004	0.035**
	(0.008)	(0.011)	(0.013)	(0.017)
	[0.078]	[0.057]	[0.071]	[0.092]
Social studies	0.037*	0.043	-0.009	0.072
	(0.021)	(0.035)	(0.038)	(0.044)
	[0.379]	[0.310]	[0.379]	[0.397]
Received high school diploma	-0.003	-0.006	0.000	-0.007
	(0.011)	(0.019)	(0.020)	(0.021)
	[0.597]	[0.548]	[0.598]	[0.648]
N	28,066	9,547	9,023	8,010
	Failed prior year math test			
Middle school: Math scores				
6th grade	0.031***	0.045***	0.035**	0.009

	(0.009)	(0.014)	(0.017)	(0.017)
7th grade	0.001	-0.012	0.035*	-0.009
	(0.010)	(0.015)	(0.019)	(0.019)
8th grade	0.007	0.009	0.016	-0.001
	(0.010)	(0.015)	(0.019)	(0.020)
9th grade	-0.008	-0.005	-0.013	0.013
	(0.011)	(0.017)	(0.021)	(0.022)
10th grade	0.005	0.011	-0.008	0.018
	(0.010)	(0.018)	(0.020)	(0.020)
Changed schools by 7th grade	-0.010	-0.019	-0.026	0.010
	(0.008)	(0.012)	(0.016)	(0.015)
	[0.221]	[0.203]	[0.267]	[0.202]
Grades 6-12: ever suspended	0.004	-0.002	-0.009	0.017
	(0.010)	(0.016)	(0.017)	(0.019)
	[0.598]	[0.558]	[0.721]	[0.548]
Grades 6-12: % absent days	-0.001	-0.002	-0.003	0.000
	(0.001)	(0.002)	(0.002)	(0.002)
	[0.074]	[0.081]	[0.071]	[0.070]
High school: number of college credit-bearing courses in…				
All subjects	0.036	0.050	0.011	0.154**
	(0.037)	(0.054)	(0.075)	(0.070)
	[0.828]	[0.709]	[0.905]	[0.745]
ELA	0.017	0.024	0.009	0.045**
	(0.011)	(0.017)	(0.023)	(0.021)
	[0.200]	[0.166]	[0.232]	[0.173]
Math	0.003	0.010	0.005	0.013
	(0.006)	(0.007)	(0.012)	(0.011)
	[0.054]	[0.036]	[0.063]	[0.049]
Science	-0.007	-0.011	-0.005	0.003
	(0.008)	(0.011)	(0.015)	(0.015)
	[0.103]	[0.081]	[0.106]	[0.117]
Social studies	0.023	0.027	0.001	0.094**
	(0.021)	(0.033)	(0.043)	(0.041)
	[0.470]	[0.426]	[0.504]	[0.406]
Received high school diploma	0.013	0.030*	0.018	0.007
	(0.009)	(0.016)	(0.017)	(0.020)
	[0.622]	[0.576]	[0.641]	[0.643]
N	34,524	13,539	9,660	9,491

Notes: All regressions control for the baseline student characteristics listed in the upper panel of Table 1 along with school-by-year fixed effects, and robust standard errors, clustered at the prior year reading score level, are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome obtained using linear polynomial specification and a bandwidth of 10 points. The numbers in brackets represent the predicted control mean at the cutoff. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively.

Online Appendix Table 15. Estimated Effects of Failing the 5th Grade Reading Test on 6th Grade Educational Inputs, Tracking, Classroom-Peer Composition, and Student Outcomes, Non-Parametric Estimates, by Student Race/Ethnicity

	Failed prior year reading test			
	6th grade teacher characteristics and class size in ELA			
	Average Teacher VA Score	Teacher with 10+ years of experience	Class size	Same race/ethnicity teacher
White	0.003***	0.007	-1.412***	0.010*
	(0.001)	(0.009)	(0.096)	(0.006)
N	90,863	82,012	52,452	92,762
Bandwidth	40.83	30.35	19.69	34.51
Control mean at cutoff		0.536	21.21	0.877
Black	0.002*	0.039***	-1.103***	0.061***
	(0.001)	(0.009)	(0.092)	(0.008)
N	67,842	80,113	57,771	91,153
Bandwidth	31.22	29.40	20.86	34.24
Control mean at cutoff		0.419	20.46	0.419
Hispanic	0.003***	0.049***	-1.172***	0.043***
	(0.001)	(0.009)	(0.094)	(0.009)
N	62,517	89,627	55,339	69,736
Bandwidth	31.64	36.79	21.84	27.30
Control mean at cutoff		0.472	21.54	0.325
	Took an advanced course in…			
	ELA: 6th grade	Other subjects: 6th grade	ELA: 7th or 8th grade	Other subjects: 7th or 8th grade
White	-0.069***	-0.018**	-0.002	-0.018**
	(0.008)	(0.009)	(0.008)	(0.009)
N	61,789	61,789	70,101	72,798
Bandwidth	22.34	22.84	25.96	26.53
Control mean at cutoff	0.207	0.261	0.227	0.325
Black	-0.165***	-0.095***	-0.038***	-0.024***
	(0.008)	(0.009)	(0.009)	(0.009)
N	61,915	61,915	56,569	74,938
Bandwidth	21.34	21.35	19.91	26.46
Control mean at cutoff	0.318	0.352	0.296	0.366
Hispanic	-0.096***	-0.040***	-0.014	-0.003
	(0.008)	(0.008)	(0.008)	(0.009)
N	57,060	67,059	69,484	74,251
Bandwidth	21.65	25.39	26.01	28.08
Control mean at cutoff	0.245	0.306	0.249	0.350
	Classroom-Peer 5th Grade Achievement:			
	ELA: 6th grade	Other subjects: 6th grade	ELA: 7th or 8th grade	Other subjects: 7th or 8th grade
White	-0.256***	-0.069***	-0.015*	-0.017**

	(0.010)	(0.008)	(0.008)	(0.007)
N	38,899	53,029	64,570	62,685
Bandwidth	14.40	19.33	29.14	28.50
Black	-0.273***	-0.134***	-0.038***	-0.023***
	(0.010)	(0.008)	(0.009)	(0.008)
N	41,591	47,961	53,249	59,752
Bandwidth	14.43	16.80	23.40	26.77
Hispanic	-0.219***	-0.061***	-0.013	-0.003
	(0.009)	(0.007)	(0.008)	(0.008)
N	47,792	59,142	54,850	55,379
Bandwidth	18.81	22.69	26.79	26.32

	ELA scores: 6th grade	ELA scores: Grades 7-10	Number of college-credit bearing courses	High school graduation
White	0.028*	0.007	0.041	0.000
	(0.015)	(0.014)	(0.046)	(0.014)
N	33,930	29,935	23,931	31,207
Bandwidth	36.04	33.83	25.76	33.47
Control mean at cutoff			0.361	0.543
Black	0.008	-0.003	-0.039	0.012
	(0.015)	(0.013)	(0.048)	(0.014)
N	30,896	28,633	24,499	25,244
Bandwidth	37.43	35.80	28.30	29.89
Control mean at cutoff			0.646	0.580
Hispanic	0.042***	0.000	0.060	-0.001
	(0.015)	(0.014)	(0.054)	(0.012)
N	27,207	25,310	24,064	31,165
Bandwidth	37.15	35.39	32.04	44.22
Control mean at cutoff			0.680	0.682

Notes: Robust standard errors are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome, obtained non-parametrically using the optimal bandwidth calculated using the procedure described in Calonico et al. (2017) by race/ethnicity in ELA. All regressions control for the baseline student characteristics listed in the upper panel of Table 1. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively.

Online Appendix Table 16. Estimated Effects of Failing the 5th Grade Math Test on 6th Grade Educational Inputs, Tracking, Classroom-Peer Composition, and Student Outcomes, Non-Parametric Estimates, by Student Race/Ethnicity

	Failed prior year math test			
	6th grade teacher characteristics and class size in math			
	Average Teacher VA Score	Teacher with 10+ years of experience	Class size	Same race/ethnicity teacher
White	-0.001	-0.002	-0.353***	-0.005
	(0.001)	(0.007)	(0.073)	(0.005)
N	89,378	123,508	106,567	130,116
Bandwidth	29.19	33.05	28.34	35.50
Control mean at cutoff		0.434	21.61	0.798
Black	-0.003	-0.014*	-0.551***	0.009
	(0.002)	(0.008)	(0.080)	(0.008)
N	63,740	75,133	72,431	97,288
Bandwidth	27.83	24.32	23.98	33.70
Control mean at cutoff		0.340	21.01	0.363
Hispanic	-0.000	-0.002	-0.324***	0.005
	(0.002)	(0.007)	(0.085)	(0.007)
N	74,417	111,028	71,671	111,028
Bandwidth	31.74	37.73	22.77	37.50
Control mean at cutoff		0.365	21.84	0.291
	Took an advanced course in…			
	Math: 6th grade	Other subjects: 6th grade	Math: 7th or 8th grade	Other subjects: 7th or 8th grade
White	-0.065***	-0.028***	-0.012	-0.011
	(0.010)	(0.009)	(0.009)	(0.008)
N	49,957	69,420	65,608	84,674
Bandwidth	12.75	17.93	16.45	21.20
Control mean at cutoff	0.246	0.304	0.251	0.345
Black	-0.120***	-0.039***	-0.021**	-0.017*
	(0.011)	(0.009)	(0.009)	(0.009)
N	47,989	63,302	63,302	63,302
Bandwidth	14.53	19.45	19.03	19.99
Control mean at cutoff	0.370	0.399	0.250	0.415
Hispanic	-0.059***	-0.020**	-0.010	-0.000
	(0.010)	(0.009)	(0.009)	(0.008)
N	48,756	64,692	58,449	82,899
Bandwidth	14.31	19.98	17.58	25.54
Control mean at cutoff	0.263	0.335	0.247	0.380
	Classroom-Peer 5th Grade Achievement:			
	Math: 6th grade	Other subjects: 6th grade	Math: 7th or 8th grade	Other subjects: 7th or 8th grade
White	-0.092***	-0.028***	-0.023***	-0.011

	(0.008)	(0.006)	(0.008)	(0.007)
N	55,793	80,514	53,148	62,197
Bandwidth	14.33	20.11	17.22	19.39
Black	-0.183***	-0.087***	-0.031***	-0.025***
	(0.011)	(0.008)	(0.009)	(0.009)
N	37,142	50,886	54,048	48,127
Bandwidth	11.02	15.67	22.04	18.32
Hispanic	-0.114***	-0.036***	-0.017**	-0.007
	(0.009)	(0.007)	(0.009)	(0.007)
N	53,503	64,420	57,641	59,244
Bandwidth	16.45	19.64	22.48	22.64

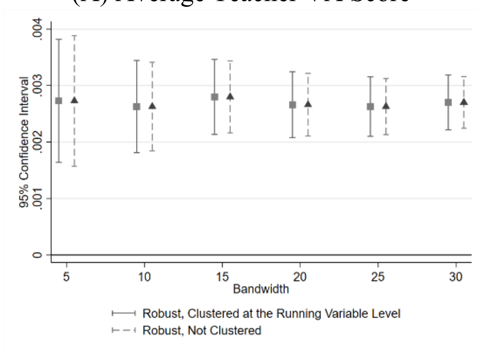
	Math scores: 6th grade	Math scores: Grades 7-10	Number of college-credit bearing courses	High school graduation
White	0.040***	0.007	0.027	0.032**
	(0.011)	(0.010)	(0.044)	(0.013)
N	31,901	34,102	30,773	30,773
Bandwidth	24.86	27.49	23.94	23.61
Control mean at cutoff			0.825	0.559
Black	0.040***	0.025**	0.060	0.011
	(0.014)	(0.012)	(0.059)	(0.014)
N	22,523	25,864	26,587	27,301
Bandwidth	24.93	29.74	29.77	30.51
Control mean at cutoff			0.976	0.659
Hispanic	0.011	-0.009	0.074	-0.006
	(0.014)	(0.013)	(0.052)	(0.013)
N	26,323	24,196	26,368	30,076
Bandwidth	29.63	27.74	30	34
Control mean at cutoff			0.819	0.671

Notes: Robust standard errors are given in parentheses. The estimates represent the treatment effect (β) on the corresponding outcome, obtained non-parametrically using the optimal bandwidth calculated using the procedure described in Calonico et al. (2017) by race/ethnicity in math. All regressions control for the baseline student characteristics listed in the upper panel of Table 1. *, **, and *** represent statistical significance at 10, 5, and 1 percent, respectively.

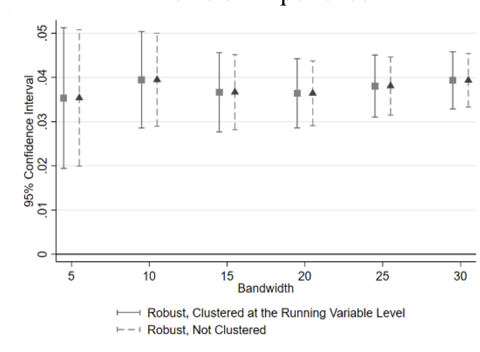
Online Appendix Figure 1. Robustness to Bandwidth Selection and Standard Error Clustering, Effects on Educational Inputs in the Subject of Remediation in 6th Grade

Failed prior year reading test

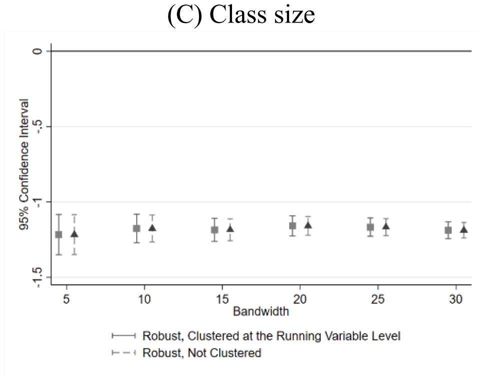
(A) Average Teacher VA Score



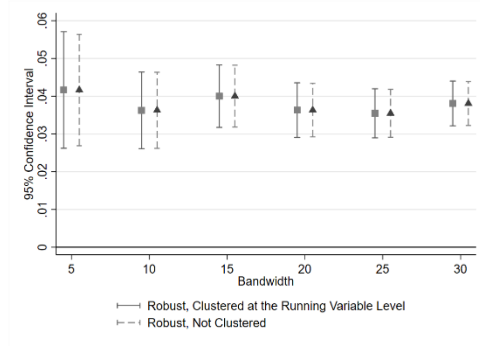
(B) Assigned to a Teacher with 10+ Years of Experience



(C) Class size

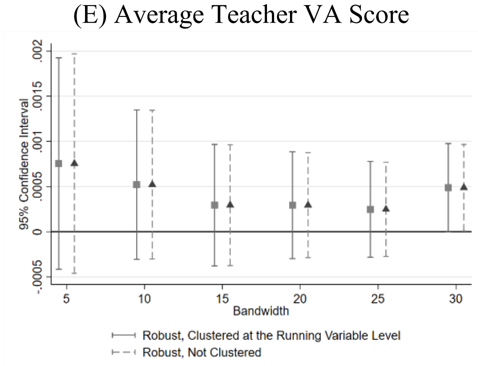


(D) Assigned to a Same Race/Ethnicity Teacher

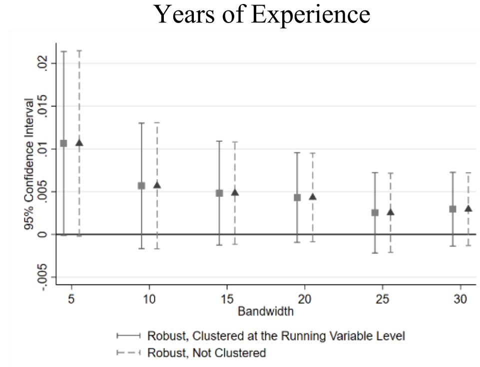


Failed prior year math test

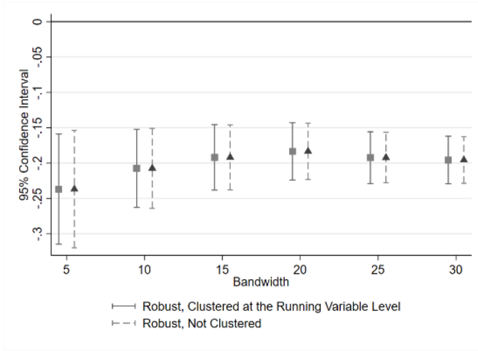
(E) Average Teacher VA Score



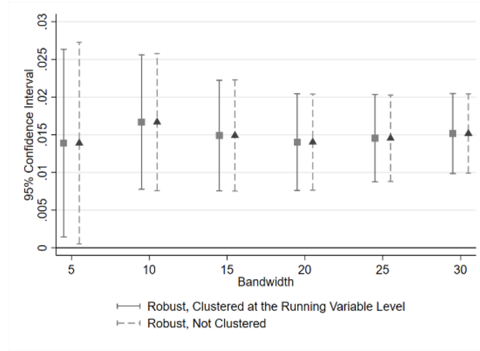
(F) Assigned to a Teacher with 10+ Years of Experience



(G) Class size



(H) Assigned to a Same Race/Ethnicity Teacher

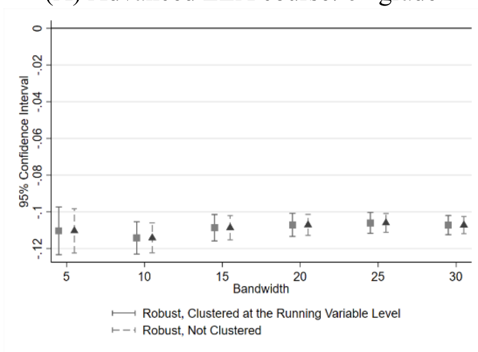


Notes: The figures present the treatment effect (β) and the 95% confidence interval estimated using the bandwidth shown, with robust standard errors clustered at the prior year reading or math score level (solid line) and not clustered (dashed line). All regressions control for the baseline student characteristics listed in the upper panel of Table 1 and school-by-year fixed-effects.

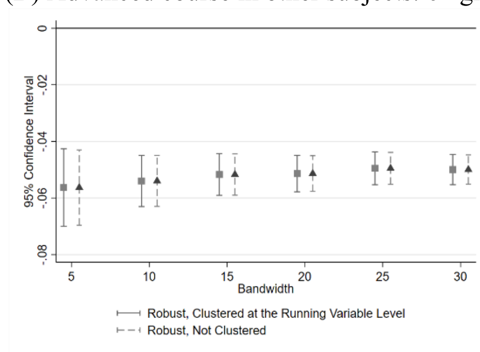
Online Appendix Figure 2. Robustness to Bandwidth Selection and Standard Error Clustering, Effects on Tracking

Failed prior year reading test

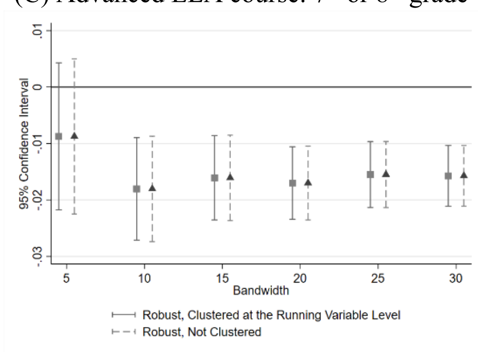
(A) Advanced ELA course: 6th grade



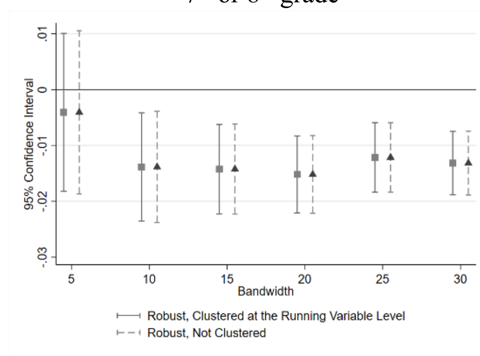
(B) Advanced course in other subjects: 6th grade



(C) Advanced ELA course: 7th or 8th grade

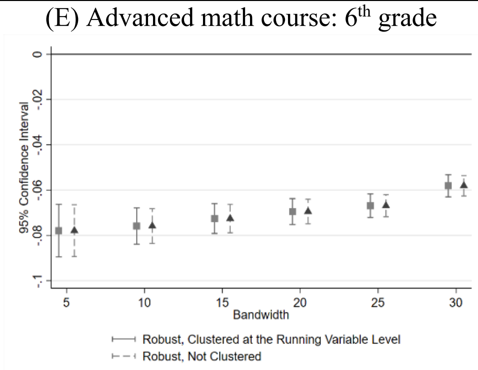


(D) Advanced course in other subjects: 7th or 8th grade

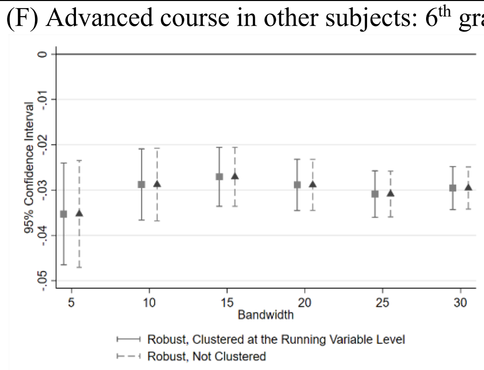


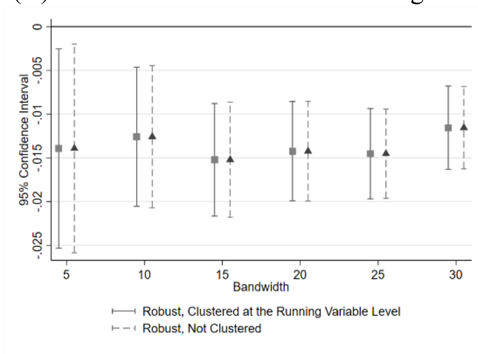
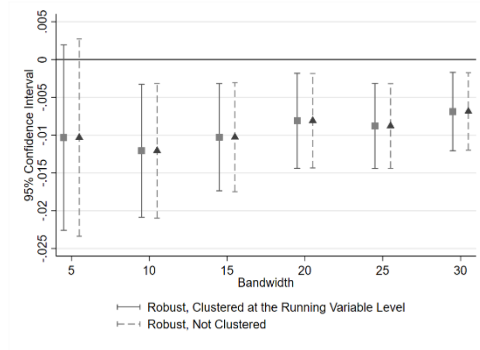
Failed prior year math test

(E) Advanced math course: 6th grade



(F) Advanced course in other subjects: 6th grade



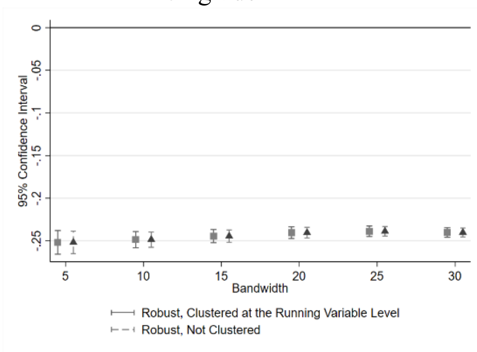
(G) Advanced math course: 7th or 8th grade	(H) Advanced course in other subjects: 7th or 8th grade
	

Notes: The figures present the treatment effect (β) and the 95% confidence interval estimated using the bandwidth shown, with robust standard errors clustered at the prior year reading or math score level (solid line) and not clustered (dashed line). All regressions control for the baseline student characteristics listed in the upper panel of Table 1 and school-by-year fixed-effects.

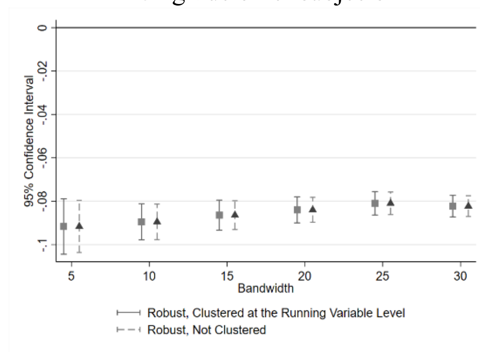
Online Appendix Figure 3. Robustness to Bandwidth Selection and Standard Error Clustering, Effects on Classroom Peers

Failed prior year reading test

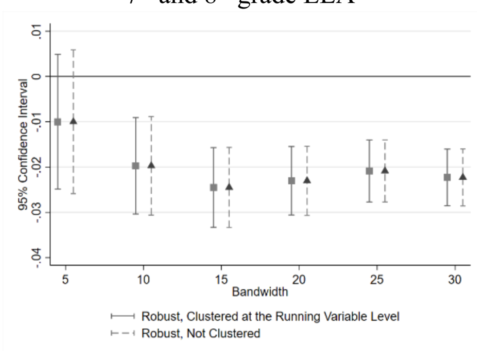
(A) Classroom-Peer 5th Grade Achievement: 6th grade ELA



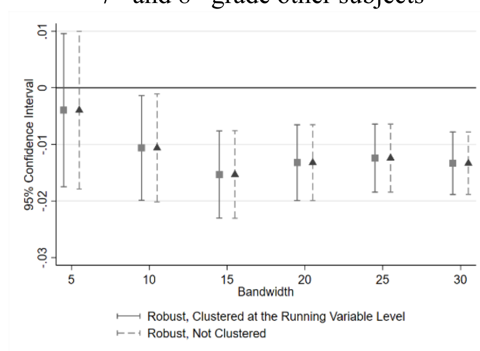
(B) Classroom-Peer 5th Grade Achievement: 6th grade other subjects



(C) Classroom-Peer 5th Grade Achievement: 7th and 8th grade ELA

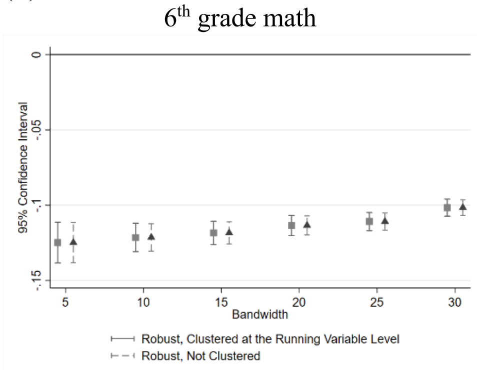


(D) Classroom-Peer 5th Grade Achievement: 7th and 8th grade other subjects

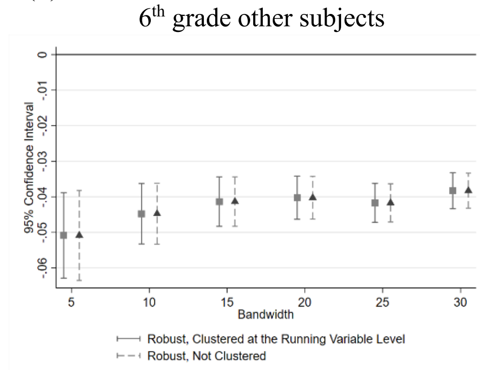


Failed prior year math test

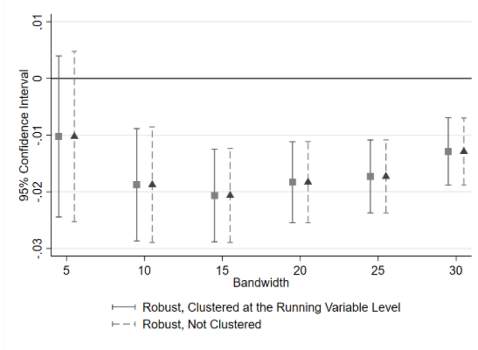
(E) Classroom-Peer 5th Grade Achievement: 6th grade math



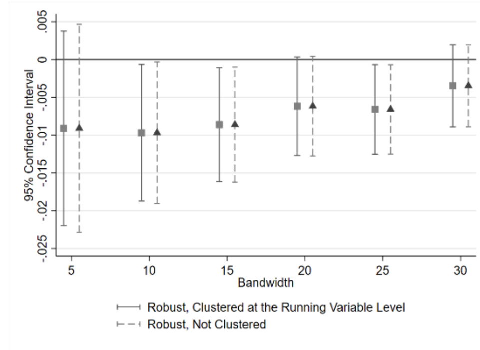
(F) Classroom-Peer 5th Grade Achievement: 6th grade other subjects



(G) Classroom-Peer 5th Grade Achievement: 7th and 8th grade math



(H) Classroom-Peer 5th Grade Achievement: 7th and 8th grade other subjects

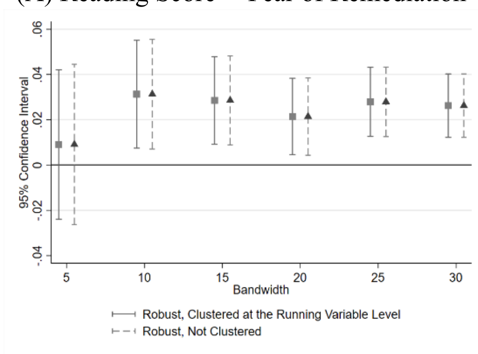


Notes: The figures present the treatment effect (β) and the 95% confidence interval estimated using the bandwidth shown, with robust standard errors clustered at the prior year reading or math score level (solid line) and not clustered (dashed line). All regressions control for the baseline student characteristics listed in the upper panel of Table 1 and school-by-year fixed-effects.

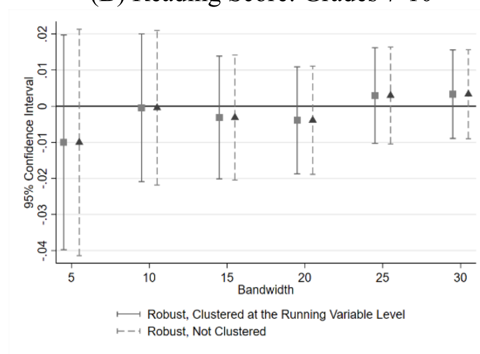
Online Appendix Figure 4. Robustness to Bandwidth Selection and Standard Error Clustering, Effects on Student Outcomes

Failed prior year reading test

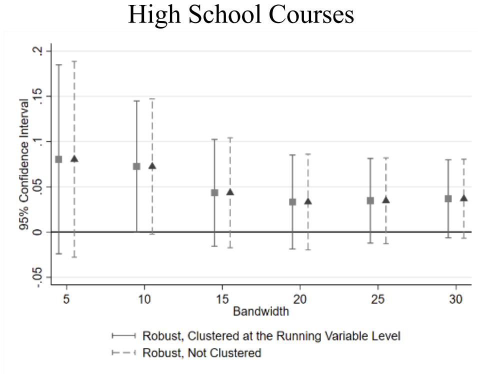
(A) Reading Score – Year of Remediation



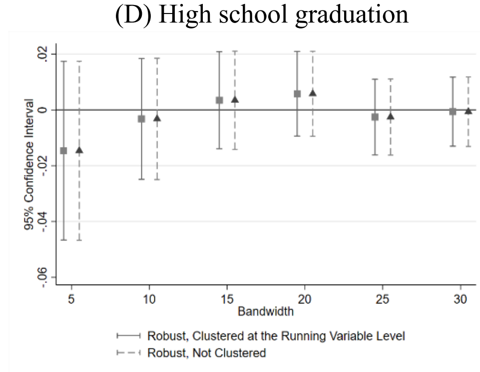
(B) Reading Score: Grades 7-10



(C) Number of College-Credit Bearing High School Courses

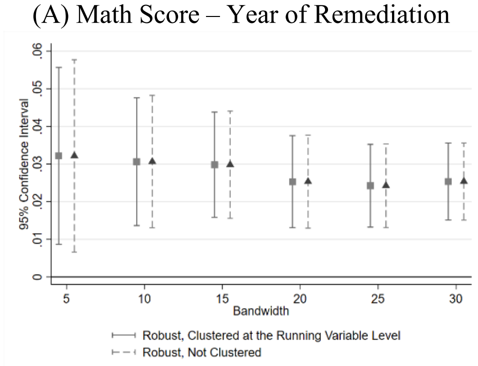


(D) High school graduation

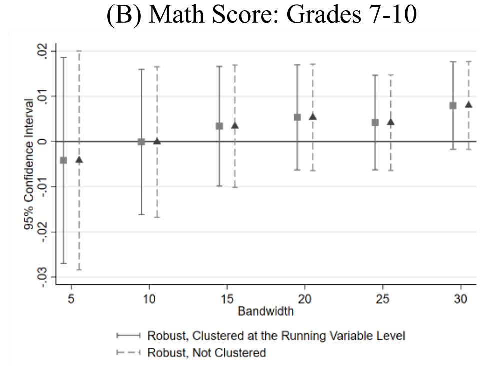


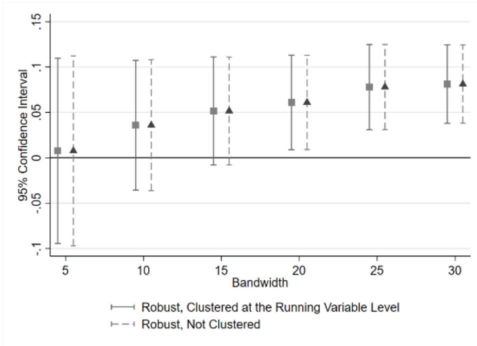
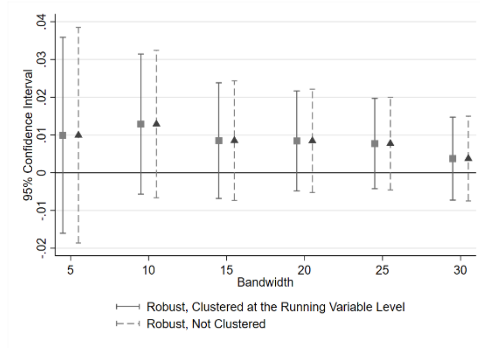
Failed prior year math test

(A) Math Score – Year of Remediation



(B) Math Score: Grades 7-10

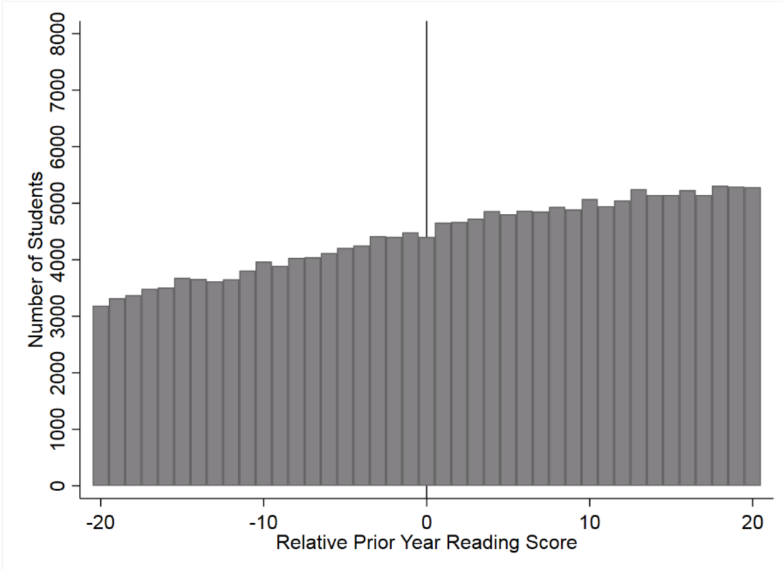


(C) Number of College-Credit Bearing High School Courses	(D) High school graduation
	

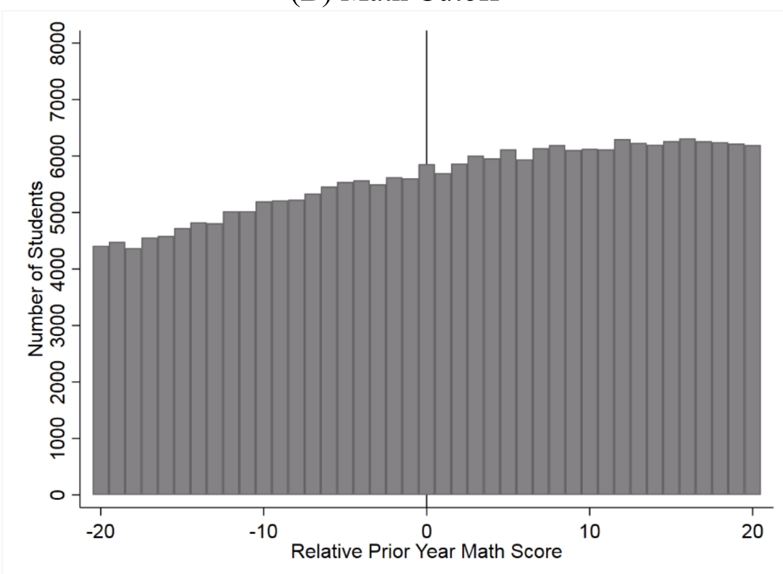
Notes: The figures present the treatment effect (β) and the 95% confidence interval estimated using the bandwidth shown, with robust standard errors clustered at the prior year reading or math score level (solid line) and not clustered (dashed line). All regressions control for the baseline student characteristics listed in the upper panel of Table 1 and school-by-year fixed-effects.

Online Appendix Figure 5. Distribution of Prior Year Test Scores

(A) Reading Cutoff



(B) Math Cutoff

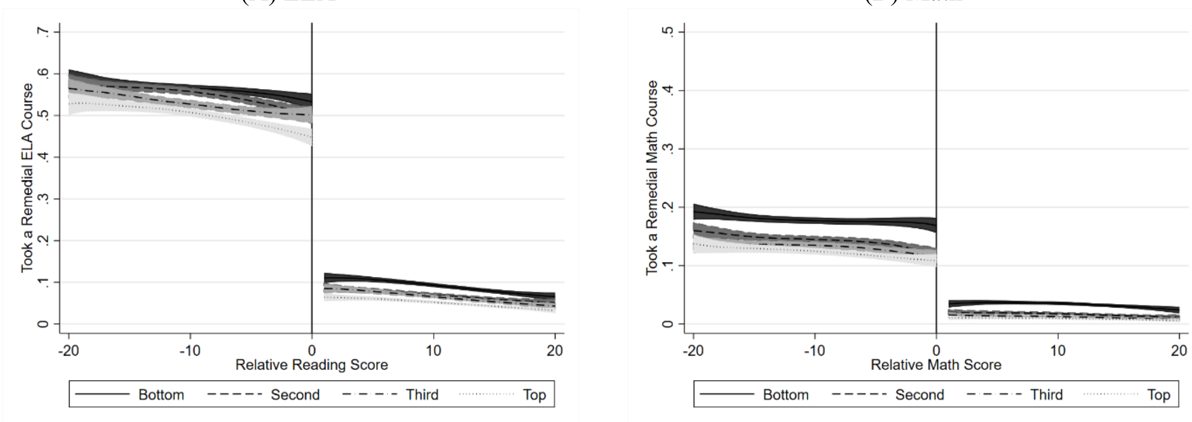


Notes: The figure presents the number of students in each test score bin between 20 points below and above the remediation cutoff in reading and math, which is shown by the vertical line.

Online Appendix Figure 6 - Remedial Course-Taking Around the Cutoff in ELA and Math, by SES Quartile

(A) ELA

(B) Math

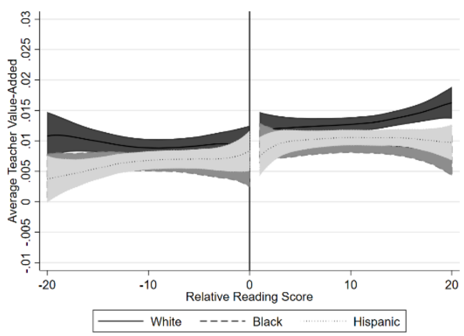


Notes: The figures present the local linear smoothing of remedial course-taking indicators in the corresponding subject in 6th grade on relative prior year test score of the student separately for the left of the corresponding cutoff and the right, broken down by student SES quartile. The triangle kernel and a bandwidth of 10 points are used in the estimation. The shaded areas represent 95 percent confidence intervals.

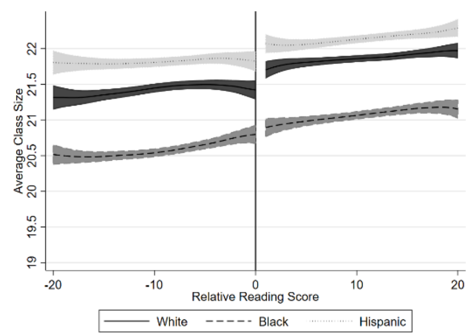
Online Appendix Figure 7. Teacher Value-Added, Class Size, and Teacher Race/Ethnicity by Course Subject and Student Race/Ethnicity, Other Core Subjects

ELA Cutoff

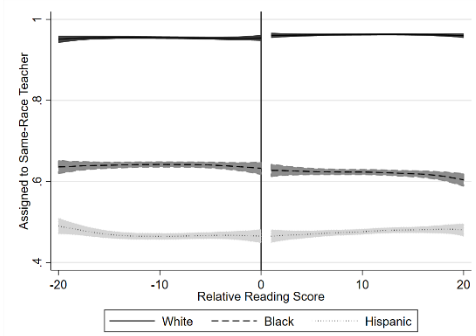
(A) Teacher value-added



(B) Class size

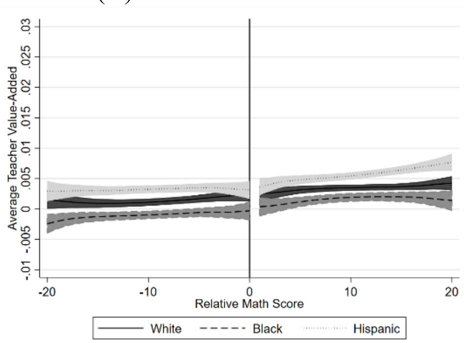


(C) Same-race teacher

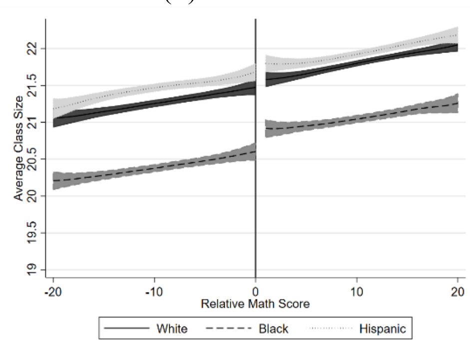


Math Cutoff

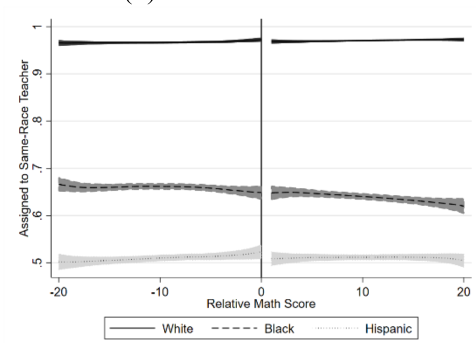
(D) Teacher value-added



(E) Class size



(F) Same-race teacher

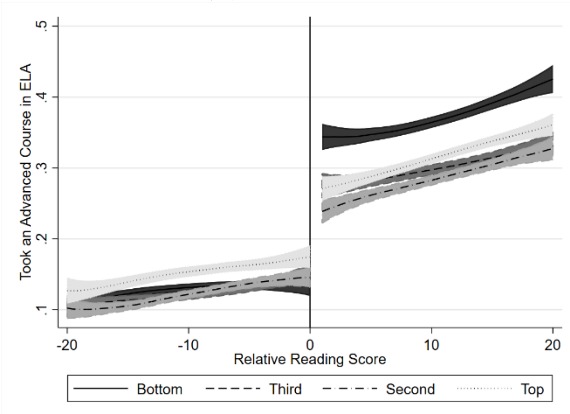


Notes: The figures present the local linear smoothing of average teacher value-added scores, average class size, and the likelihood of being assigned to at least one same-race/ethnicity teacher on relative prior year test score of the student separately for the left of the corresponding cutoff and the right, broken down by student race/ethnicity in other core subjects. The triangle kernel and a bandwidth of 10 points are used in the estimation. The shaded areas represent 95 percent confidence intervals.

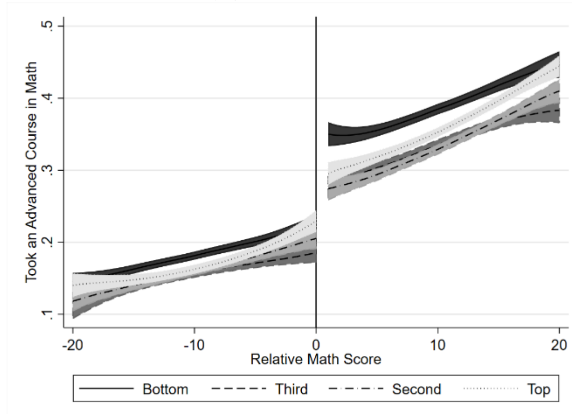
Online Appendix Figure 8. Advanced Course-Taking in 6th Grade Around the Remediation Cutoffs, by Student SES and Course Subject

Remediation subject

(A) ELA cutoff

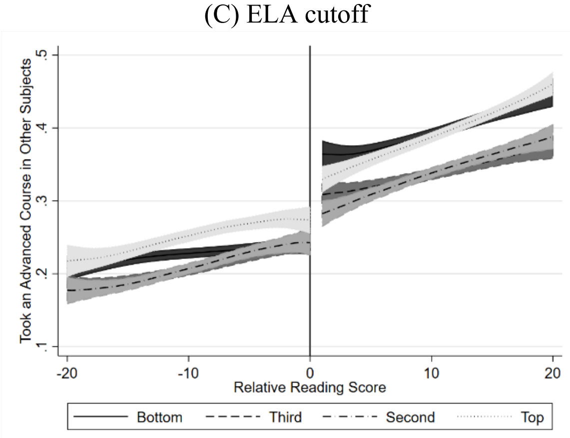


(B) Math cutoff

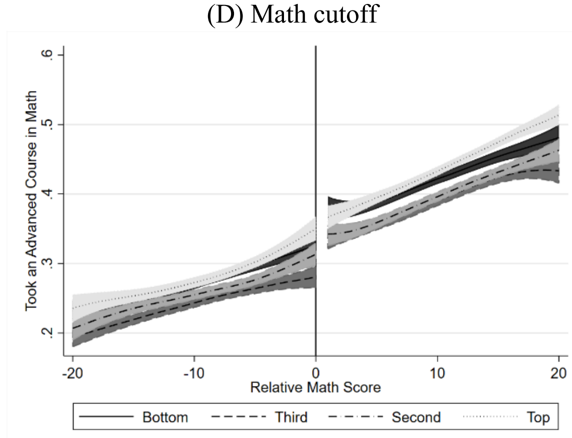


Other core subjects

(C) ELA cutoff



(D) Math cutoff

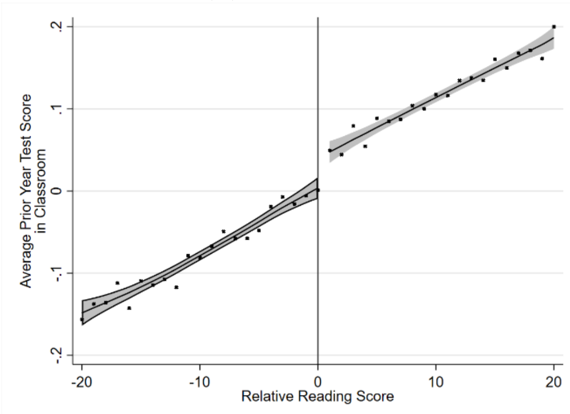


Notes: The figures present the local linear smoothing of advanced course-taking indicator in the remediation subject and other core subjects in 6th grade on relative prior year reading (first column) and math score (second column) broken down by student SES quartile. The triangle kernel and a bandwidth of 10 points are used in the estimation. The solid circles represent raw cell means and the shaded areas represent 95 percent confidence intervals.

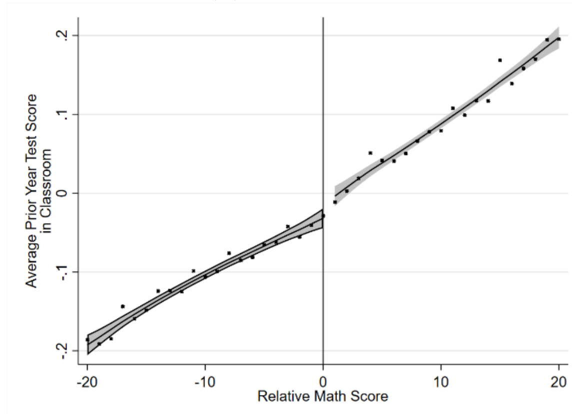
Online Appendix Figure 9. Average 6th Grade Classroom-Peer Prior Achievement Around the Remediation Cutoffs in ELA and Math in Remediation and Other Core Subjects, Before the Policy

Remediation subject

(A) ELA cutoff

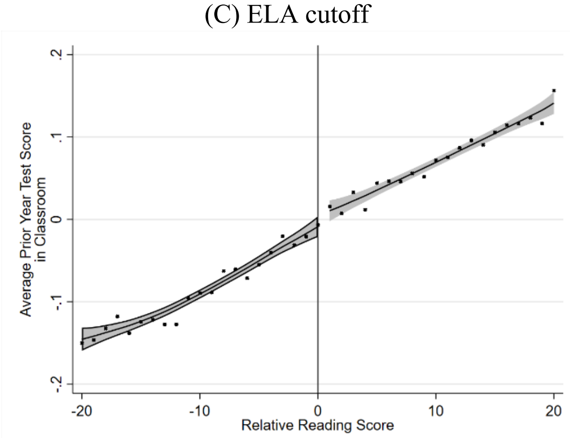


(B) Math cutoff

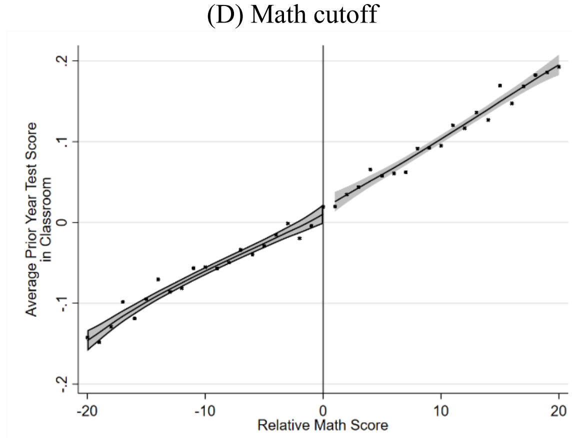


Other core subjects

(C) ELA cutoff



(D) Math cutoff

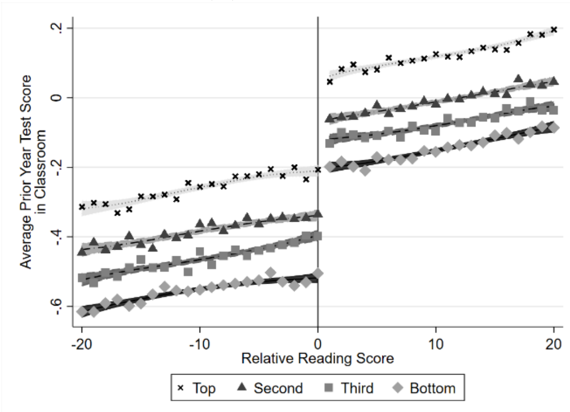


Notes: The figures present the local linear smoothing of average classroom-peer prior achievement in the remediation subject and other core subjects in 6th grade on relative prior year reading (first column) and math score (second column) of the student before the policy separately for the left of the corresponding cutoff and the right. The triangle kernel and a bandwidth of 10 points are used in the estimation. The solid circles represent raw cell means and the shaded areas represent 95 percent confidence intervals.

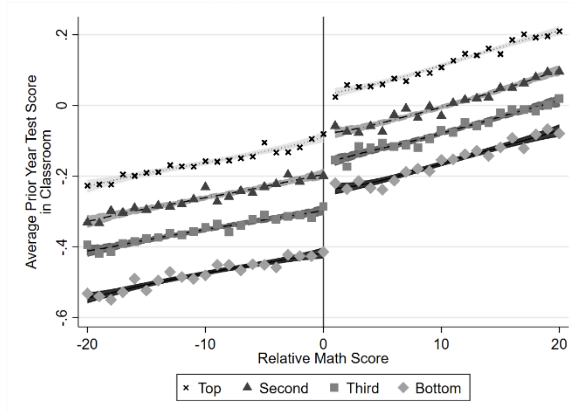
Online Appendix Figure 10. Average 6th Grade Classroom-Peer Prior Achievement Around the Remediation Cutoffs in ELA and Math, by Student SES and Course Subject

Remediation subject

(A) ELA cutoff

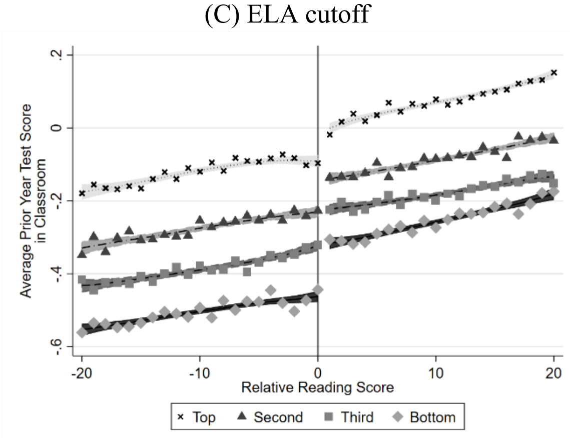


(B) Math cutoff

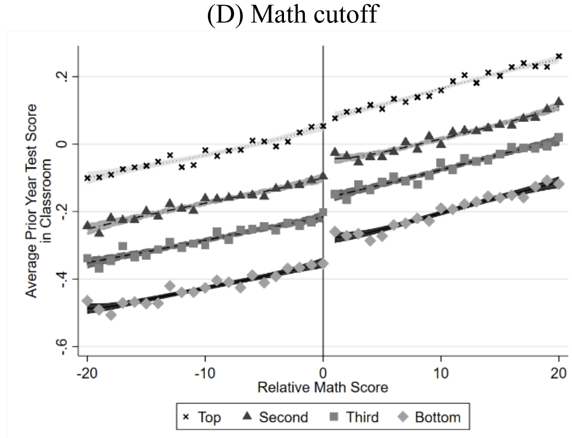


Other core subjects

(C) ELA cutoff



(D) Math cutoff

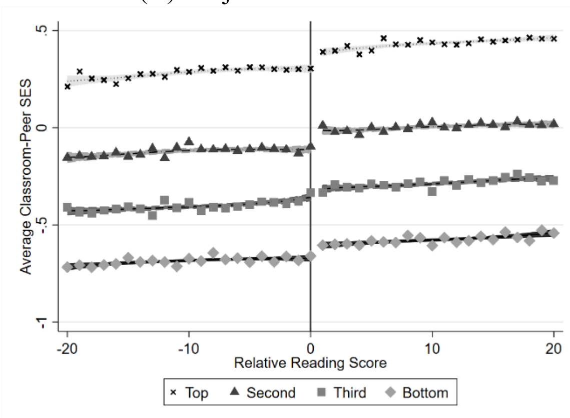


Notes: The figures present the local linear smoothing of average classroom-peer prior achievement in the remediation subject and other core subjects in 6th grade on relative prior year reading (first column) and math score (second column) of the student, by student SES quartile. The triangle kernel and a bandwidth of 10 points are used in the estimation. The solid circles represent raw cell means and the shaded areas represent 95 percent confidence intervals.

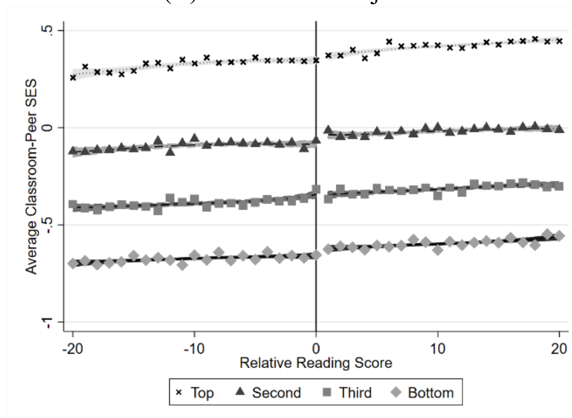
Online Appendix Figure 11. Average 6th Grade Classroom-Peer SES in Subject of Remediation and Other Core Subjects Around the Remediation Cutoffs in ELA and Math, by Own SES Quartile

ELA cutoff

(A) Subject of Remediation

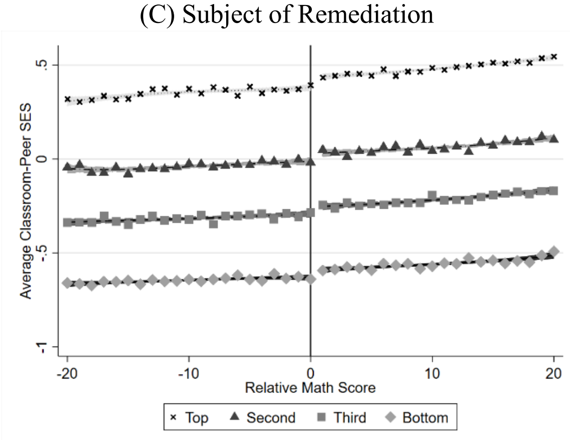


(B) Other Core Subjects

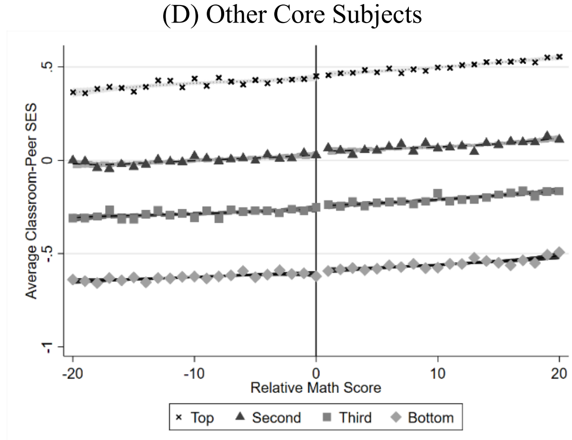


Math cutoff

(C) Subject of Remediation



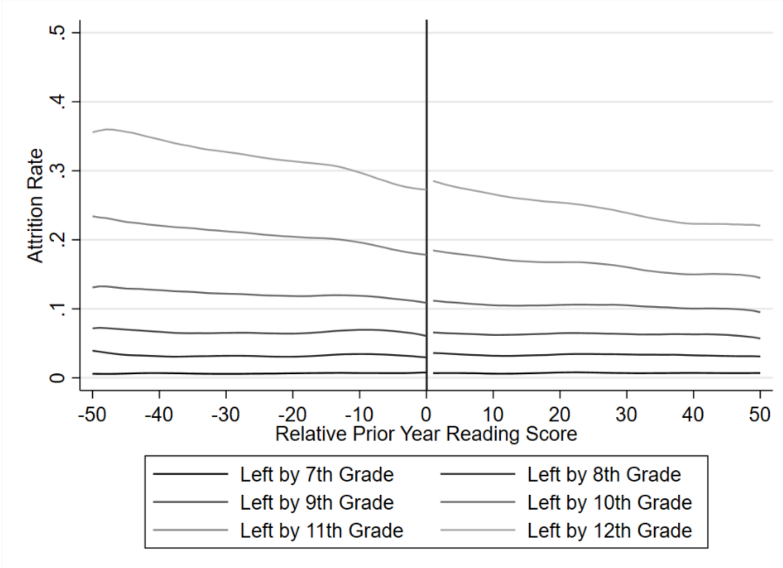
(D) Other Core Subjects



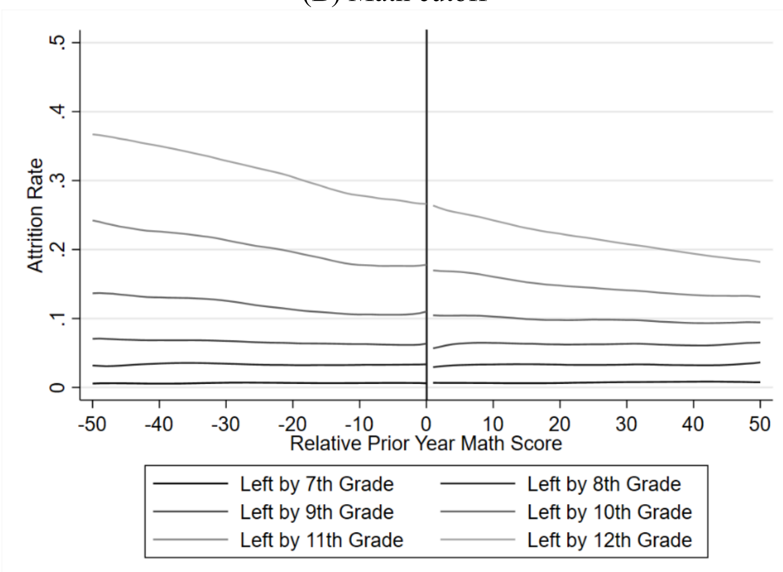
Notes: The figures present the local linear smoothing of average classroom-peer SES in the subject of remediation in panels (A) and (C) and other core subjects in panels (B) and (D) in 6th grade on relative prior year reading (first column) and math score (second column) broken down by own SES quartile. The triangle kernel and a bandwidth of 10 points are used in the estimation. The solid circles represent raw cell means and the shaded areas represent 95 percent confidence intervals.

Online Appendix Figure 12. Sample Attrition Around the Remediation Cutoffs

(A) Reading cutoff



(B) Math cutoff

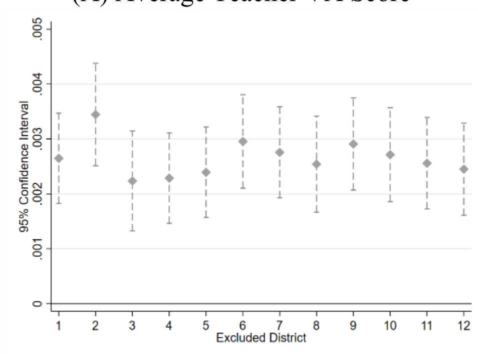


Notes: The figures present the local linear smoothing of the attrition rate in the following years separately for the left of the cutoff date and the right using a bandwidth of 10 points for reading (panel A) and math (panel B) cutoffs.

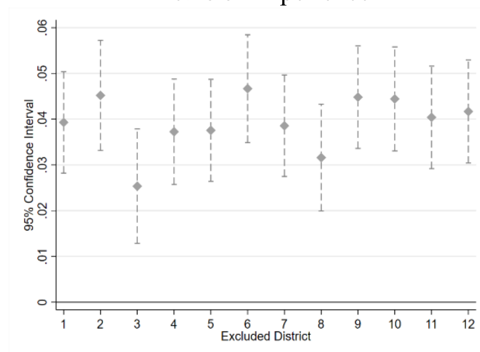
Online Appendix Figure 13. Robustness to Excluding Individual Districts, Effects on Educational Inputs in the Subject of Remediation in 6th Grade

Failed prior year reading test

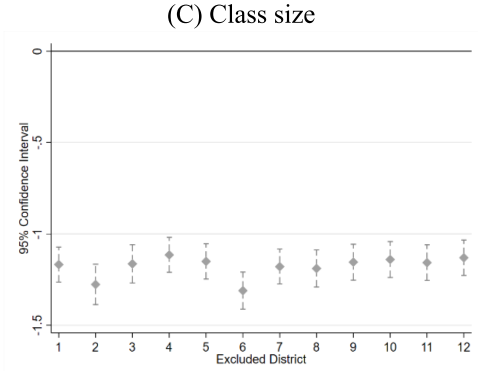
(A) Average Teacher VA Score



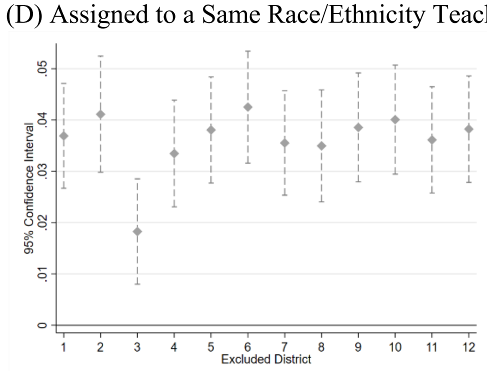
(B) Assigned to a Teacher with 10+ Years of Experience



(C) Class size

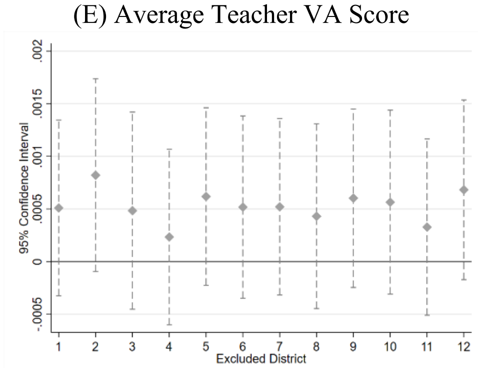


(D) Assigned to a Same Race/Ethnicity Teacher

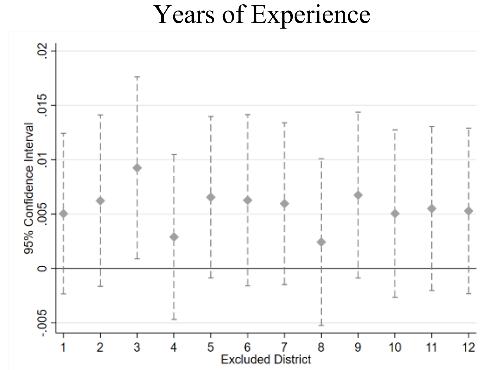


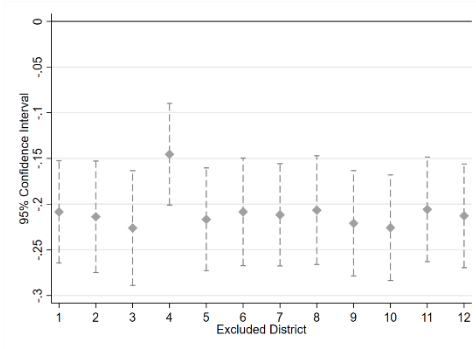
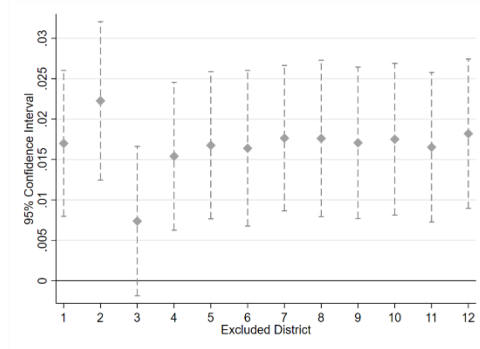
Failed prior year math test

(E) Average Teacher VA Score



(F) Assigned to a Teacher with 10+ Years of Experience



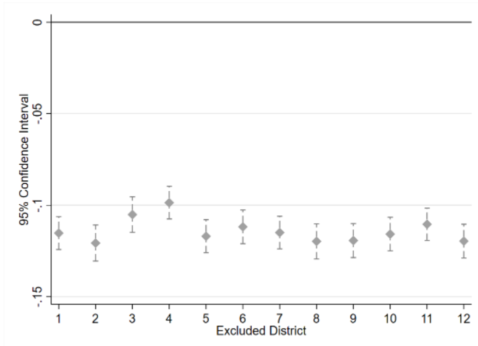
(G) Class size	(H) Assigned to a Same Race/Ethnicity Teacher
	

Notes: The figures present the treatment effect (β) and the 95% confidence interval estimated excluding the district given in the x-axis, with robust standard errors clustered at the prior year reading or math score level (solid line). All regressions control for the baseline student characteristics listed in the upper panel of Table 1 and school-by-year fixed-effects.

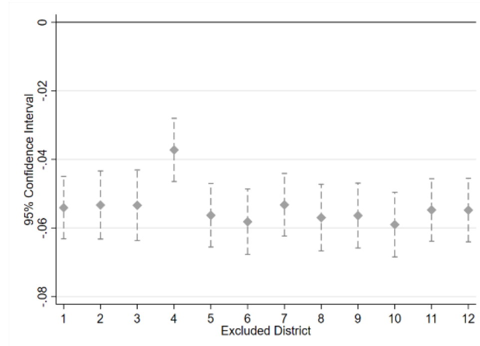
Online Appendix Figure 14. Robustness to Excluding Individual Districts, Effects on Tracking

Failed prior year reading test

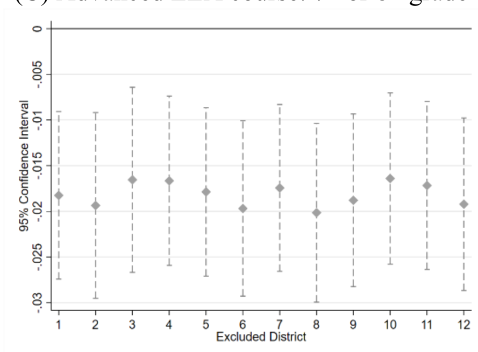
(A) Advanced ELA course: 6th grade



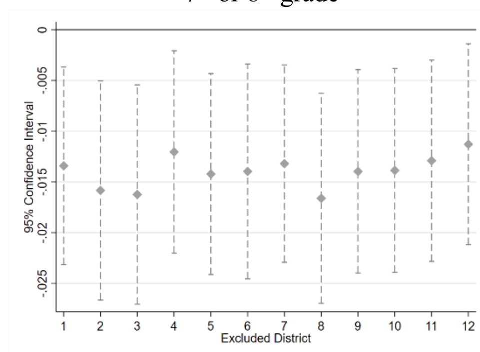
(B) Advanced course in other subjects: 6th grade



(C) Advanced ELA course: 7th or 8th grade

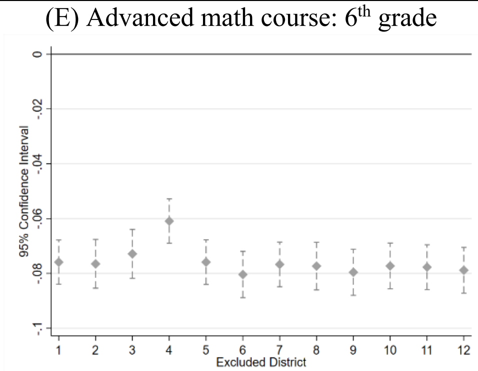


(D) Advanced course in other subjects: 7th or 8th grade

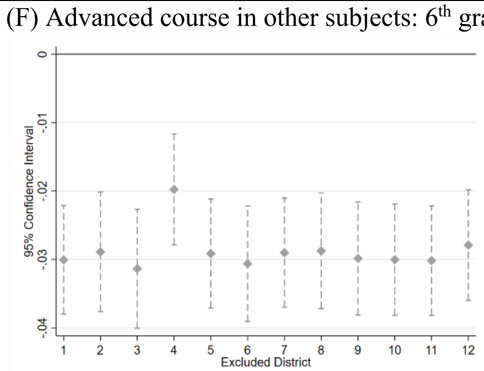


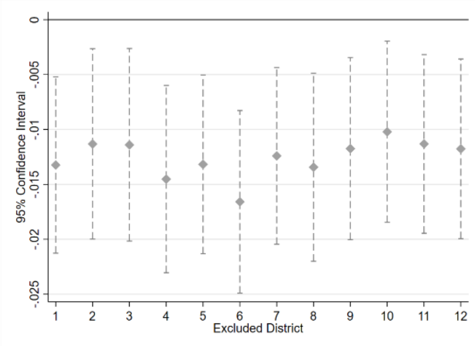
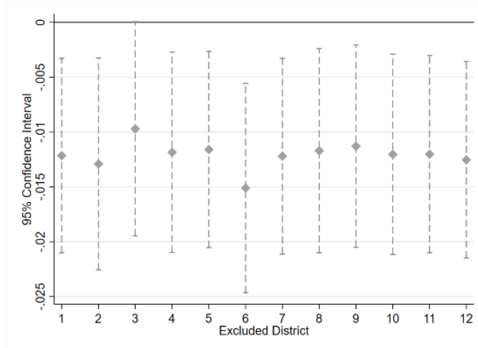
Failed prior year math test

(E) Advanced math course: 6th grade



(F) Advanced course in other subjects: 6th grade



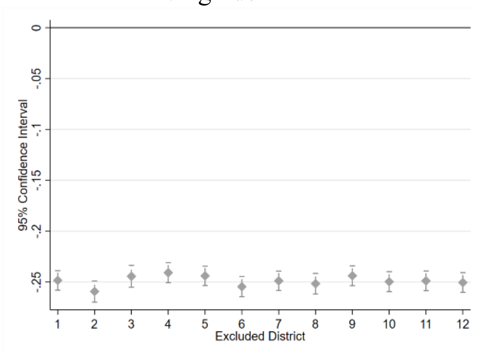
(G) Advanced math course: 7th or 8th grade	(H) Advanced course in other subjects: 7th or 8th grade
	

Notes: The figures present the treatment effect (β) and the 95% confidence interval estimated excluding the district given in the x-axis, with robust standard errors clustered at the prior year reading or math score level (solid line). All regressions control for the baseline student characteristics listed in the upper panel of Table 1 and school-by-year fixed-effects.

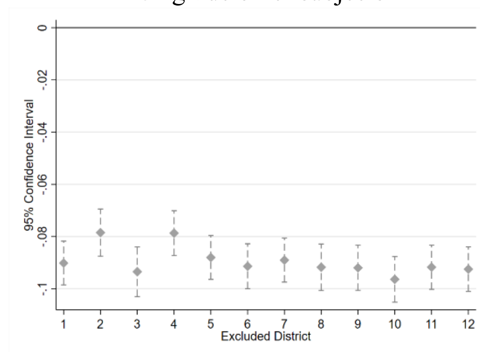
Online Appendix Figure 15. Robustness to Excluding Individual Districts, Effects on Classroom Peers

Failed prior year reading test

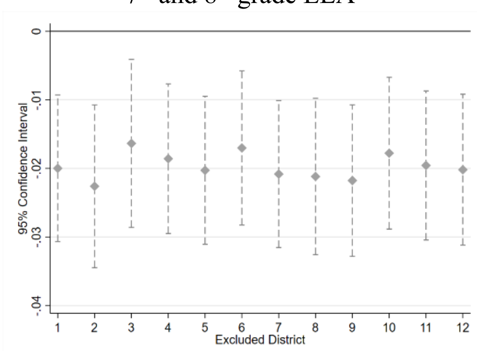
(A) Classroom-Peer 5th Grade Achievement: 6th grade ELA



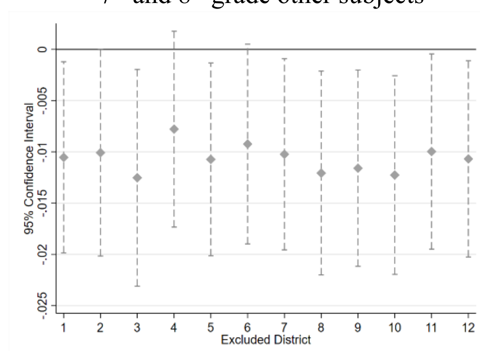
(B) Classroom-Peer 5th Grade Achievement: 6th grade other subjects



(C) Classroom-Peer 5th Grade Achievement: 7th and 8th grade ELA

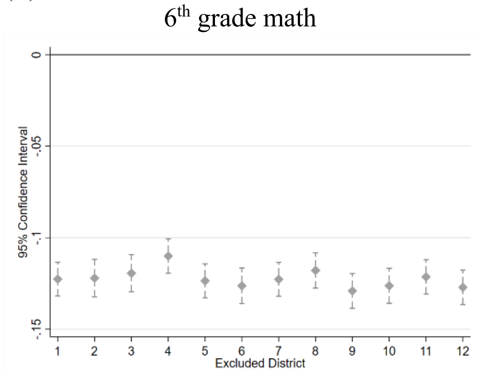


(D) Classroom-Peer 5th Grade Achievement: 7th and 8th grade other subjects

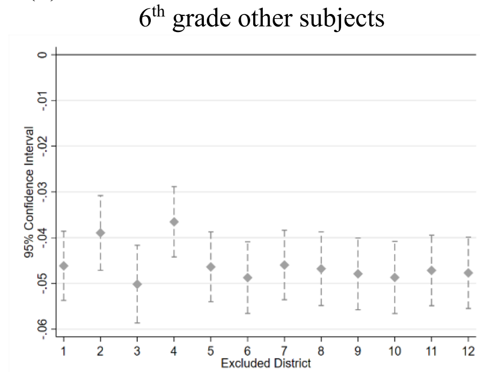


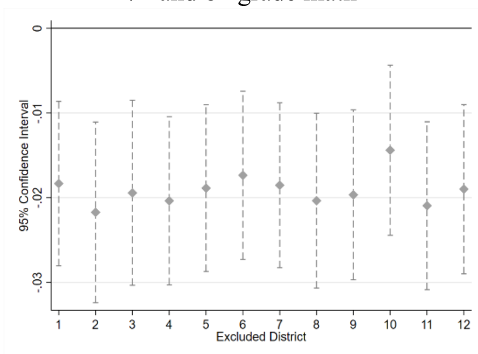
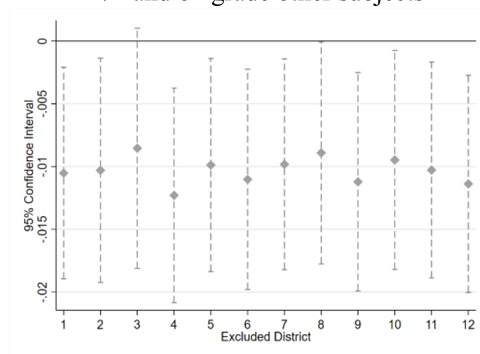
Failed prior year math test

(E) Classroom-Peer 5th Grade Achievement: 6th grade math



(F) Classroom-Peer 5th Grade Achievement: 6th grade other subjects



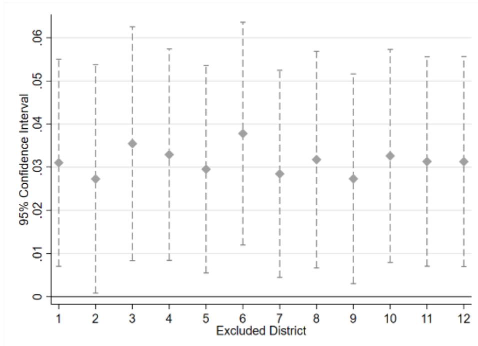
(G) Classroom-Peer 5th Grade Achievement: 7th and 8th grade math	(H) Classroom-Peer 5th Grade Achievement: 7th and 8th grade other subjects
	

Notes: The figures present the treatment effect (β) and the 95% confidence interval estimated excluding the district given in the x-axis, with robust standard errors clustered at the prior year reading or math score level (solid line). All regressions control for the baseline student characteristics listed in the upper panel of Table 1 and school-by-year fixed-effects.

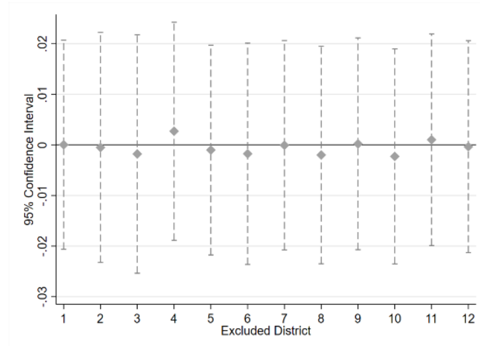
Online Appendix Figure 16. Robustness to Excluding Individual Districts, Effects on Student Outcomes

Failed prior year reading test

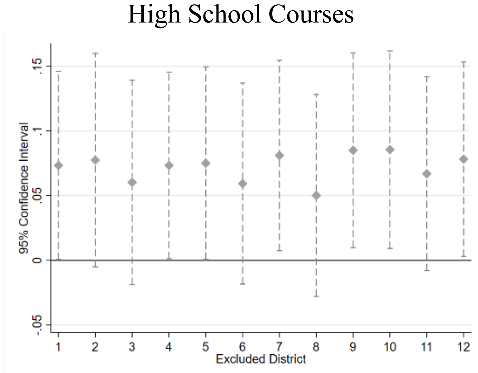
(A) Reading Score – Year of Remediation



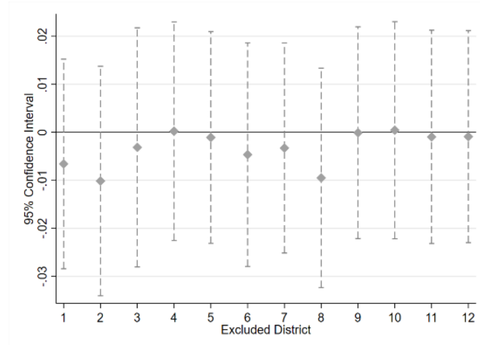
(B) Reading Score: Grades 7-10



(C) Number of College-Credit Bearing High School Courses

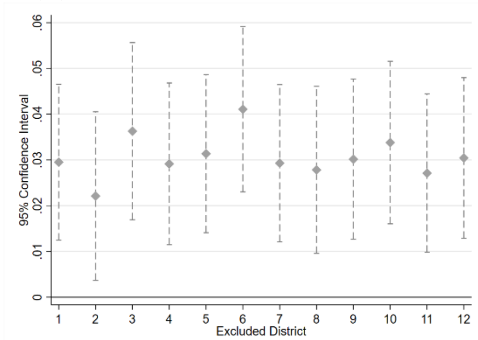


(D) High school graduation

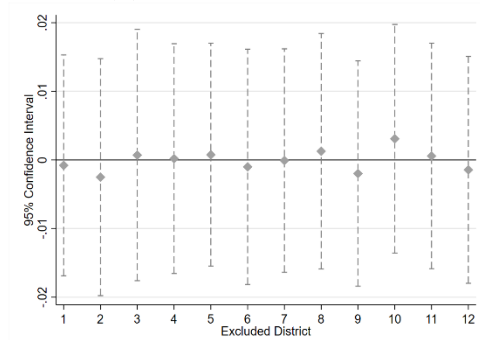


Failed prior year math test

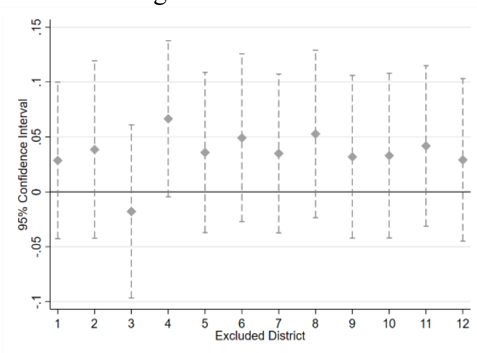
(A) Math Score – Year of Remediation



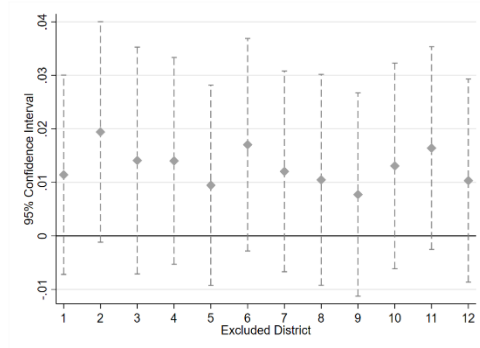
(B) Math Score: Grades 7-10



(C) Number of College-Credit Bearing High School Courses



(D) High school graduation



Notes: The figures present the treatment effect (β) and the 95% confidence interval estimated excluding the district given in the x-axis, with robust standard errors clustered at the prior year reading or math score level (solid line). All regressions control for the baseline student characteristics listed in the upper panel of Table 1 and school-by-year fixed-effects.