Doubling Up: The Long Run Impacts of Remedial Algebra on High School Graduation and College Enrollment

**** Please Do Not Cite without Author's Consent ****

Kalena E. Cortes*
The Bush School of Government and Public Service
Texas A&M University
kcortes@bushschool.tamu.edu

Joshua Goodman
John F. Kennedy School of Government
Harvard University
joshua_goodman@hks.harvard.edu

Takako Nomi
Consortium on Chicago School Research
University of Chicago
nomit@ccsr.uchicago.edu

May 2, 2012

Abstract

Success or failure in freshman algebra has long been thought to have a strong impact on subsequent high school outcomes. We study a remedial algebra policy implemented by the Chicago Public Schools (CPS) for cohorts entering high school in 2003 and 2004. Students scoring below the national median on an eighth grade exam were assigned in ninth grade to a remedial algebra course that doubled instructional time and increased tracking by ability. The assignment rule allows difference-in-difference and regression discontinuity estimates of average and local average treatments effects. Using longitudinal data that tracks students from eight grade to college enrollment, we confirm prior work showing modest or no short-run impacts on algebra passing rates and math scores. We show, however, positive and substantial long run impacts of remediation on college entrance exam scores, high school graduation rates and college enrollment rates. The bulk of this impact comes from students with below average reading skills, perhaps because the intervention focused on written expression of mathematical concepts. These facts point to the importance both of evaluating interventions beyond the short run and of targeting interventions toward appropriate students. This is the first evidence we know of to demonstrate long run impacts of remediation in an American high school setting.

JEL Classifications: I20, I21, I24, J10, J13, J15, J48
Keywords: Remedial Education, Student Outcomes, Postsecondary Enrollment

---

* We are indebted to Sue Sporte, Associate Director for Evaluation and Data Resources, the Consortium on Chicago School Research (CCSR), University of Chicago, for making the data available for this project. Special thanks for helpful comments from Richard Murnane, Bridget Terry Long, Jeffrey D. Kubik, Lori Taylor, and Nora Gordon. Also, we would like to thank the seminar participants at Harvard University's Program on Education Policy and Governance (PEPG) Education Policy Colloquia Series, State of Texas Education Research Center (ERC) at Texas A&M University; and conference participants at the Association for Education Finance and Policy (AEFP). Institutional support from Texas A&M University and Harvard University are also gratefully acknowledged. Research results, conclusions, and all errors are naturally our own.

I. Introduction

There is an increasing recognition and concern that too few high school students, especially those in urban areas, are graduating with the necessary skills needed for college and the workforce. The high school completion rate has been declining over the past decade among students at all income levels and only about half of racial minority students finish high school (NCES, 2004)[1]. Despite increases in college enrollment, there has been little increase in college graduation rates among African-American and Latino students, who often struggle on entering college (NCES, 2011). High schools, particularly urban ones, are often blamed for not graduating students with the skills and coursework they need to be successful in college.

There is now a national movement calling for more rigorous high school requirements that are explicitly linked to the skills that students will need for work and college. The National Governor's Association, for example, has recommended enacting high school reform through rigorous college preparatory graduation requirements, programs to encourage disadvantaged students to take Advanced Placement (AP) exams and college-preparatory classes, and the design of literacy and math support courses for students with below-grade level performance.[2] These recommendations are already being followed. A number of states are in the process of raising curricular requirements for graduation (e.g., Arkansas, Mississippi, and Illinois), and other states have enacted incentives to particularly encourage disadvantaged students to take rigorous high school course loads (e.g., Arkansas, Maine, Missouri, and Oklahoma).[3]

Calls to increase both curricular offerings and requirements are based on a large amount of evidence tying school curriculum to student outcomes. Since the late 1980s there has been

---

[1] Data from the National Center for Education Statistics (NCES) Digest of Education Statistics (2004) shows that the percentage of 17 year olds completing a high school degree has decreased from approximately 76 percent in the 1960s to about 70 percent in the late 1990s.
[2] National Governors Association Center for Best Practices (2005a).
[3] National Governors Association Center for Best Practices (2005b).

evidence that requiring students to take college preparatory classes could produce higher overall levels of achievement and reduce racial and socioeconomic achievement gaps (Bryk, Lee and Holland, 1993; Goodman, 2009). Some of this evidence comes from studies of the curricular organization of public and Catholic high schools. Of particular interest is the nature of remedial education in the PK-12 setting, which has historically been quite different in these two educational sectors. In public high schools, remedial coursework constituted an entire set of courses that could be used to satisfy graduation requirements. In contrast, in Catholic schools, students were required to enroll in additional courses to build their skills through a "double dose" of coursework. Some researchers have attributed at least part of the apparent success of Catholic schooling to such rigorous curricular requirements (Bryk, Lee and Holland, 1993; Lee, Croninger and Smith, 1997; Lee, Smith and Croninger, 1997).

Unfortunately, the vast majority of studies on the impact of remedial coursework and other curricula on student achievement have relied on cross-sectional samples analyzed with empirical strategies prone to generate substantial selection bias. The few studies on remedial education that have seriously grappled with the issue of causal inference have either looked at short-run outcomes such as test scores (Jacob and Lefgren, 2004) or longer-run outcomes in the context of U.S. colleges or high schools in other nations (Calcagno and Long, 2008; Bettinger and Long, 2009; Lavy and Schlosser, 2005). None of these studies have analyzed longer-term outcomes in the urban American high school.

However, the one study of remedial education in an urban American high school is by Nomi and Allensworth (2009) that study on which this current study builds on, the authors examined the short-term impact of a remedial math policy known as "double-dose" algebra, enacted by the Chicago Public Schools in 2003. Under this policy, students scoring below the national median on their 8th grade math exam were required to take two periods of algebra a day during the 9th grade. Students placed into these remedial classes thus received substantially more instruction time in algebra. Nomi and Allensworth (2009) analyzed the *early* high school outcomes of this policy by following students only through the 10th grade, and found positive and substantial impacts on G.P.A and standardized test scores, but no improvement in 9^{th} grade algebra course failure rates. The time frame of their initial study did not, however, allow them to explore other important outcomes beyond 10^{th} grade, such as learning in higher mathematics, high school graduation, and college attendance.

Our study examines the impact of Chicago Public Schools' remedial math policy on *longer-term outcome*s that are ultimately of more concern to students, parents and policymakers.. Specifically, we analyze advanced math coursework and performance, ACT scores, high school graduation rates and college enrollment using longitudinal transcript data from the Chicago Public Schools (CPS), which allow us to track students from 8th grade through college enrollment.

To analyze the effect of this innovative double-dose curriculum, we employ two complementary quasi-experimental techniques. The first is a difference-in-differences analysis, which compares how the differential outcomes between students above and below the remediation threshold changed upon implementation of the policy. The second is a regression discontinuity design, which compares the outcomes of students just above and just below the remediation threshold in years when the remedial math policy was in place. These empirical strategies allow us to compare students treated with the remediation policy to a similar group of untreated students who serve as the control group. Also, both of these techniques will produce reduced form estimates of the impact of the policy on all students as well as instrumental variables estimates of the impact of the policy on students induced to enroll in remedial math.

Using longitudinal data that tracks students from eight grade to college enrollment, we confirm prior work showing modest or no short-run impacts on algebra passing rates and math scores. We show, however, positive and substantial long run impacts of remediation on college entrance exam scores, high school graduation rates and college enrollment rates. The bulk of this impact comes from students with below average reading skills, perhaps because the intervention focused on written expression of mathematical concepts. These facts point to the importance both of evaluating interventions beyond the short run and of targeting interventions toward appropriate students. This is the first evidence we know of to demonstrate long run impacts of remediation in an American urban high school setting.

The double dose strategy has become an increasingly popular way to aid students struggling in mathematics. Today, nearly half of large urban districts report doubled math instruction as the most common form of support for students with lower skills (Council of Great City Schools, 2009). The central concern of urban school districts is that algebra may be a gateway for later academic success, so that early high school failure in math may have large effects on subsequent academic achievement and graduation rates. As the current policy environment calls for "algebra for all" in 9th grade or earlier grades, providing an effective and proactive intervention is particularly critical for those who lack foundational mathematical skills. A successful early intervention may have the greatest chance of having longer-term effects on students' academic outcomes.

II. Implementing Double-Dose Algebra

Since the late 1990s, Chicago Public Schools (CPS) has been at the forefront of curriculum reform designed to increase the rigor of student coursework and prepare students for college entrance. Starting with students entering high school in the fall of 1997, CPS raised its graduation requirements to align with the New Basics Curriculum.[4] CPS eliminated lower-level and remedial courses so that all first-time freshmen would enroll in algebra in 9th grade, geometry in 10th grade and algebra II or trigonometry in 11th grade. Soon after these reforms, CPS officials realized that students were unable to master the new college-prep curriculum. Passing rates in 9th grade algebra were quite low, largely because students entered high school with such poor math skills (Roderick & Camburn, 1999).

In response to these low passing rates in 9th grade algebra, CPS launched the double-dose algebra policy for all students entering high school in the fall of 2003. Instead of reinstating the traditional remedial courses from previous years, CPS required enrollment in two periods of algebra coursework for all first-time 9th graders testing below the national median on the math portion of the 8th grade Iowa Tests of Basic Skills (ITBS).[5] Such students enrolled for two math credits, a full-year regular algebra class plus a full-year algebra support class.[6] Three student cohorts, those entering high school in the fall of 2003, 2004 and 2005, were subject to the double-dose policy. Our analysis focuses on the first two cohorts because the test score-based

[4] The new basics curriculum was a minimum curriculum recommended by the National Commission of Excellence in Education in 1983, which consists of four years of English, three years of each mathematics, science, and social studies, and one-half year of computer science. The CPS requirements are actually slightly higher than the New Basics Curriculum, which includes two years of a foreign language and specific courses in mathematics (i.e., algebra, geometry, advanced algebra, and trigonometry).

[5] All CPS high schools were subject to the double-dose algebra policy, including 60 neighborhood schools, 11 magnet schools, and 6 vocational schools (Nomi and Allensworth, 2009).

[6] Double-dose algebra students received 90 minutes of math class time every day for a full academic year. The first math course (regular algebra) mostly consisted of class lectures. Whereas, the second math course (algebra with support) the teachers focused on building math skills that students lacked and covered materials in a different order than simply following the textbook. Double-dose teachers also used various instructional activities, such as working in a small group (cooperative groups), asking probing and open-ended questions, and using board work (Starkel, Martinez, and Price, 2006; Wenzel, Lawal, Conway, Fendt, and Stoelinga, 2005).

assignment rule was not followed closely for the final cohort. We will refer to these as the 2003 and 2004 cohorts.

Prior to the double-dose policy, algebra curricula had varied considerably across CPS high schools, due to the fairly decentralized nature of the district. Conversely, CPS offered teachers of double-dose algebra two specific curricula called Agile Mind and Cognitive Tutor, stand-alone lesson plans they could use, and thrice annual professional development workshops where teachers were given suggestions about how to use the extra instructional time.[7] Though it is difficult to know precisely what occurred in these extra classes, Nomi and Allensworth (2010) surveyed students to learn more about the classroom learning environment. They found that students assigned to double-dose algebra reported much more frequently: writing sentences to explain how they solved a math problem; explaining how they solved a problem to the class; writing math problems for other students to solve; discussing possible solutions with other students; and applying math to situations in life outside of school. The additional time thus focused on building verbal and analytical skills that may have conferred benefits in subjects other than math.

CPS also strongly advised schools to schedule their algebra support courses in three specific ways. First, double-dose algebra students should have the same teacher for their two periods of algebra. Second, the two algebra periods should be offered consecutively. Third, double-dose students should take their algebra support class with the same students who are in

[7] The district made the new double-dose curricula and professional development available only to teachers teaching double-dose algebra courses, but there was a possibility of spillover effects for teachers in regular algebra. However, the professional development was geared towards helping teachers structure two periods of algebra instruction. Moreover, based on CPS officials and staff members' observations of double-dose classrooms, they found that even teachers who taught both single-period and double-dose algebra tended to differentiate their instruction between the two types of classes. Specifically, teachers tended to use new practices with the double-period class, but continued to use traditional methods with the single-period class. Teachers told them that they did not feel they needed to change methods with the advanced students (i.e., non double-dose students), and that they were hesitant to try new practices that may be more time-consuming with just a single period. The double period of algebra allowed these teachers to feel like they had the time to try new practices (e.g., cooperative groups).

their regular algebra class. Most CPS schools followed these recommendations in the initial year (Nomi and Allensworth, 2009). For the 2003 cohort, 80 percent of double-dose students had the same teacher for both courses, 72 percent took the two courses consecutively, and rates of overlap between the two classes' rosters exceeded 90 percent. By 2004, schools began to object to the scheduling difficulties of assigning the same teacher to both periods so CPS removed that recommendation. For the 2004 cohort, only 54 percent of double-dose students had the same teacher for both courses and only 48 percent took the two courses consecutively. Overlap between the rosters remained, however, close to 90%. In the analysis below, we explore whether the program's impacts vary by cohort in part because of this variation in implementation.

The treatment under consideration here thus had many components. Remediation doubled the amount of instructional time and provided high quality curricula and professional development to teachers. As we will show, the recommendation that students take the two classes with the same set of peers caused tracking by skill to increase, thus reducing classroom heterogeneity. All of these factors were likely to, if anything, improve student outcomes. We will also show, however, that the increase tracking by skill placed remediated students among substantially lower skilled peers than non-remediated students. Anecdotal evidence suggests that remedial classes were taught by less experienced teachers. Both of these factors were likely to, if anything, hurt student outcomes. We will be measuring the net impact of all of these components.

II. Data and Descriptive Statistics

We use longitudinal data from CPS that tracks students from eighth grade through college enrollment. These data include demographic information, detailed high school transcripts, numerous standardized test scores, and graduation and college enrollment information. Our sample consists of all students entering ninth-grade for the first time in the fall of 2001-2004. The first two cohorts, 2001 and 2002, are untreated and the second two cohorts, 2003 and 2004, are treated. We include only students who have valid 8th grade math scores and who enroll in freshman algebra. We include only schools in which at least one classroom of students was remediated.

The main independent variable, which will provide our instrument, is each student's 8th grade score on the math portion of the Iowa Test of Basic Skills (ITBS). All CPS 8th graders are required to take this test. The 8th grade ITBS reading score will serve as an additional control, as will demographic information such as gender, race, special education status, and socioeconomic and poverty measures constructed for each student's residential block group from the 2000 Census. The latter measures are more informative than subsidized lunch measures given that over 90% of CPS students receive subsidized lunches.

The transcript data allow for detailed exploration of the treatment itself. We assign students an indicator for being remediated and continuous variables characterizing their regular algebra classes. These include the mean of each student's peers' math skills as measured by 8th grade math scores, heterogeneity as measured by the standard deviation of those scores, each student's distance to the mean peer skill, and class size. We also construct the number of full-year courses taken in subjects other than math to see the extent to which the remedial period displaces other coursework.

We focus on three sets of outcomes. First, to explore whether remediation helps students complete the math coursework CPS expects of them, we construct grades received in math coursework in freshman and later years. Second, to explore whether objective measures of math achievement improve as a result of remediation, we construct a variety of test scores standardized by cohort. These include the PLAN exam, which all CPS students take in September of both their second and third years in high school, and the ACT exam, which all CPS students take in April of their third year and is commonly used in the Midwest for college applications. Third, to explore whether remediation improves long run outcomes such as educational attainment, we construct measures of on time graduation and college enrollment. Students are coded as on time graduates if they received a regular CPS diploma within four years of starting high school. Linking CPS data with National Student Clearinghouse data allows us to construct dummies indicating whether a student enrolled in college within a year of graduating from high school and, if so, whether it was a 2-year or 4-year institution.

The summary statistics of the analytic sample are shown by cohort in Table 1. As seen in panel (A), over 90% of CPS students are black or Hispanic, with nearly 20% in special education. Panel (B) shows that the mean CPS 8th grader scored below the 50th percentile on the math portion of the ITBS. As a result, 56% of students qualified for remediation in 2003 and 2004, according to the district's rule that any student scoring below the 50th percentile should be remediated. Even though 56% qualified for remediation, only 43% were actually remediated, suggesting imperfect compliance with the rule. The primary impact of remediation appears in the increase in freshman math courses taken from 1.0 to 1.4. Tracking in freshman algebra courses also increased, with the standard deviation in math skill dropping from 17 to 15. In panel (C), only a third of students receive an A or B in freshman algebra, with nearly 20% failing that course. In panel (D), roughly 50% of students sample graduate from CPS in 4 years. 35% enroll in college within a year of graduating, evenly split between two- and four-year institutions.

III. Empirical Strategies

Comparison of the outcomes of remediated and non-remediated students might yield biased estimates of remediation's impacts given potentially large differences in unobserved characteristics between the two groups of students. To eliminate this potential bias, we exploit the fact that students scoring below the 50th percentile on the 8^{th} grade ITBS math test were supposed to enroll in double-dose algebra. This allows us to use two approaches to identifying the impact of remediation, a difference-in-difference estimate using all cohorts and a regression discontinuity design using only the treated cohorts. In both cases, we use the assignment rule as an exogenous source of variation in the probability that a given student will be remediated.

We implement the difference-in-difference (DD) approach using the regressions below:

$$Y_{it} = \alpha_0 + \alpha_1 lowscore * post + \alpha_2 lowscore + \alpha_3 post + \delta_{it} \qquad (1)\ (RF)$$

$$remediated_{it} = \gamma_0 + \gamma_1 lowscore * post + \gamma_2 lowscore + \gamma_3 post + \theta_{it} \qquad (2)\ (FS)$$

$$Y_{it} = \beta_0 + \beta_1 remediated + \beta_2 lowscore + \beta_3 post + \varepsilon_{it} \qquad (3)\ (IV)$$

Here, *lowscore* indicates an 8th grade math score below the 50th percentile, *post* indicates the 2003 and 2004 cohorts, *remediated* is a remediation indicator, and *Y* represents any outcome of interest. By controlling for differences between low- and high-scoring students in the pre-treatment cohorts and for overall differences between cohorts, the interaction coefficient (α_1) from equation (1) estimates how the difference in outcomes between low- and high-scoring students changed at the time double-dose algebra was introduced. This reduced form equation produces an intention-to-treat (ITT) estimate because compliance with the assignment rule was imperfect. We therefore use equation (2) as a first stage to predict how introduction of the policy affected the probability of a low-scoring student being remediated.

Our ultimate estimate of interest is therefore the *remediated* coefficient (β_1) from equation (3), in which remediation has been instrumented with the interaction of *lowscore* and *post* using that first stage. This IV regression produces a treatment-on-treated (TOT) estimate of the impact of remediation on students induced into remediation by the introduction of the policy. This approach estimates an average treatment effect (ATE) of double-dose algebra for all students remediated because of the policy. Here, high-scoring students serve as a control group for low-scoring students, so that these estimates will be unbiased under the assumption that no changes other than double-dose algebra differentially affected low- and high-scoring students over this time period.

If other factors are differentially changing between low- and high-scoring students over this time period, the underlying assumption critical to the DD approach may be violated. Our second empirical strategy employs a regression discontinuity (RD) that uses only the treated cohorts and thus does not depend on this assumption. Similar to the DD approach, we implement the RD approach using the regressions below:

$$Y_{it} = \alpha_0 + \alpha_1 lowscore + \alpha_2 math8 + \alpha_3 lowscore * math8 + \delta_{it} \quad \text{(4) (RF)}$$

$$remediated_{it} = \gamma_0 + \gamma_1 lowscore + \gamma_2 math8 + \gamma_3 lowscore * math8 + \theta_{it} \quad \text{(5) (FS)}$$

$$Y_{it} = \beta_0 + \beta_1 remediated + \beta_2 math8 + \beta_3 lowscore * math8 + \varepsilon_{it} \quad \text{(6) (IV)}$$

Here, the variable definitions are the same as in the DD, and *math8* is each student's 8th grade math score re-centered around the 50th percentile cutoff. The *lowscore* coefficient (α_1) from equation (4) estimates the discontinuity of interest by comparing the outcomes of students just below and just above the remediation threshold. As with the DD, this reduced form equation produces an ITT estimate because of imperfect compliance with the assignment rule. The first stage equation (5) measures the difference in remediation rates between students just below and just above the threshold.

As with the DD, our ultimate estimate of interest is the *remediated* coefficient (β_1) from equation (6), in which remediation has been instrumented with *lowscore* using that first stage. This approach estimates a local average treatment effect (LATE) of double-dose algebra for students near the 50th percentile of 8th grade math skill. Here, students just above the threshold serve as a control group for students just below the threshold, so that these estimates will be unbiased under the assumption that no other factors change discontinuously around the threshold itself.

For the RD, our preferred specification will fit straight lines on either side of the threshold using a bandwidth of 10 percentiles, and will also control for gender, race, special education status, socioeconomic and poverty measures, and 8th grade reading scores. For the DD, our preferred specification includes those controls as well as 8th grade math scores. In all specifications, heteroskedasticity robust standard errors will be clustered by each student's initial high school to account for within high school correlations in the error terms. For both the DD and RD, we will show that our central results are robust to different choices of controls and bandwidths.

IV. The Treatment

We first explore the treatment itself to learn more about how the double-dose algebra policy was changing students' freshman year experiences. Before turning to regression results, we look at visual evidence. Figure 1 plots the proportion of students remediated for each 8th grade math percentile. Panel (A) shows imperfect compliance, with remediation rates reaching a maximum of about 80% for students in the 20-40th percentiles. Students in the lowest percentiles have lower remediation rates because they are more likely to be supported through special education programs. Some students above the threshold are remediated, perhaps because schools cannot perfectly divide students into appropriately sized classes by the assignment rule. Panel (B) reveals that compliance for students just below the threshold is substantially lower in the 2004 cohort than the 2003 cohort, providing further motivation to analyze program impacts separately by cohort.

Table 2 shows the first-stage results using a low 8th grade math score indicator as an instrument for remediation. Panel (A) is the difference-in-difference specification with the treated cohorts pooled. Panel (B) replicates panel (A) but allows the estimate to vary by treated cohort. Panel (C) is the regression discontinuity specification with treated cohorts pooled. Panel (D) replicates panel (C) but allows the discontinuity to vary by treated cohort. Column (1) implements equations (2) and (5), including no additional controls. Column (2) adds demographic and test score controls. Column (3) includes those controls and high school fixed effects.

The estimates from the difference-in-difference specification imply that low-scoring students were 59 percentage points more likely to be remediated than high-scoring students. The magnitude of this impact was slightly larger in 2003 than in 2004. The regression discontinuity estimates suggest that students just below the threshold were 40 percentage points more likely to be remediated than students just above the threshold. Consistent with Figure 1, this discontinuity was much larger in 2003 (48 percentage points) than in 2004 (32 percentage points), further motivating our decision throughout the paper to explore differential impacts by cohort. Finally, the robustness of the first stage estimates to inclusion of a wide variety of controls and high school fixed effects suggest that these estimates are driven neither by changing composition of the student body over time nor by discontinuities in any variables other than remediation itself.

Table 3 explores heterogeneity in compliance rates by math and reading skill. Column (1) in both panels replicates column (2) from Table 2, allowing the first stage to vary by skill. Columns (2) and (3) use only one treated cohort at a time. Panel (A) shows what is visually apparent from Figure 1, namely that the shape of the compliance is an inverted U, with the lowest and highest mathematically skilled students to the left of the threshold complying at lower rates than those in the middle. Panel (B) shows that the discontinuity is quite similar in magnitude for students with below and above average readings skills. These patterns hold true for both cohorts.

Figures 2-4 show further visual evidence of the channels through which remediation may have affected student outcomes. Figure 2 shows that the number of algebra courses taken by freshman rose substantially when the policy was introduced and has essentially the same shape as Figure 1 because each remediated student took one additional course. Figure 3 shows that skill tracking increased substantially in the freshman algebra classes of both low- and high-scoring students, though such classroom heterogeneity decreased somewhat more for low-scoring students, particularly those near the threshold. Figure 4 shows that the increased tracking placed low-scoring students in algebra with lower skilled peers and high-scoring students with higher skilled peers.

Table 4 explores the impact of remediation on the freshman academic experience. All of the coefficients come from regressions in which remediation has been instrumented by eligibility, as in column (2) of Table 2. As such, these are TOT estimates of the impact of remediation on those actually remediated. Column (1) shows that the remediation increased the number of freshman math courses taken by one, as would be expected from the double-dose strategy. Columns (2)-(4) show that students fit this additional math course into their schedule not by replacing core academic courses (in English, social studies and science) but by replacing other courses such as fine arts and foreign languages. The result was a positive but small overall change in the number of courses taken freshman year.

Columns (5)-(8) highlight remediation impact on channels other than instructional time. The increased skill tracking meant that remediated students took algebra classes with peers whose 8th grade math scores were substantially lower than the peers of non-remediated students. The estimates in column (5) imply that remediation lowered the mean peer skill of the average remediated student by nearly 13 percentiles and of the remediated student near the threshold by over 19 percentiles. Because both low- and high-scoring students' algebra classes became more homogeneous, panel (A) of column (6) suggests little relative difference in that change between remediated and non-remediated students. Panel (C) suggests, however, that remediated students near the threshold were in more homogeneous classrooms than their non-remediated peers. Both the DD and RD estimates in column (7) implies that remediation increased the distance of students to their peers' mean skill by about 3 percentiles, which could have negative repercussions if teachers focus their energies on the mean student. Column (8) suggests that, near the threshold, remediation also increased the size of regular algebra classes by 2.4 students.

Remediation thus doubled instructional time in math by replacing other coursework and increased homogeneity of algebra classrooms, but lowered peer ability, increased distance from the class mean and increased class sizes. None of these effects varied substantially by cohort. We now turn to analysis of the overall impact of these various changes on grades, test scores and educational attainment.

V. Grades, Test Scores and Educational Attainment

Table 5 explores the impact of remediation on math coursework and grades. The DD estimates suggest that remediation increased the proportion of students earning at least a B in freshman algebra by a highly statistically significant 8.5 percentage points, a more than 35% increase from a base rate of 23.5 percentage points. Conversely, the proportion of students passing freshman algebra barely increased by a statistically insignificant 1.5 percentage points. Though remediation had little impact on passing rates for freshman algebra, remediated students were 4.3 percentage points more likely to have passed geometry by their second year and 2.9 percentage points more likely to have passed trigonometry by their third year. Mean GPA across all math courses taken after freshman year increased by a statistically insignificant 0.06 grade points.

The RD estimates imply that remediation increased the proportion of students earning at least a B in freshman algebra by 14.5 percentage points, a more than 50% increase from a base of 27.9 percentage points. Remediation had a large and marginally significant impact on passing rates for freshman algebra, raising them by 5.0 percentage points from a base of 83.4 percentage points. There is, however, little evidence of increased passing rates in subsequent math courses, though mean GPA across all math courses taken after freshman year increased by a marginally significant 0.16 grade points.

Taken as a whole, these results imply that the double-dose policy greatly improved freshman algebra grades for the upper end of the remediated distribution, but had relatively little impact on passing rates for the lower end of the distribution. This latter fact is one of the primary reasons that CPS has since moved away from this strategy. There is, however, some evidence of improved passing rates and GPA in later math courses, suggesting the possibility of longer run benefits beyond freshman year. Though coursework and grades matter for students' academic trajectories, the subjective nature of course grading motivates us to turn to standardized achievement measures as a potentially better measure of the impact of remediation on math skill.

Table 6 explores the impact of remediation on mathematics test scores as measured by the PLAN exam taken in September of each student's second year, the PLAN exam taken in the fall of each student's third year, and the ACT exam taken in April of each student's third year. The PLAN exams contain a pre-algebra/algebra section and a geometry section, which we analyze separately given that remediation focused on algebra. In column (1), the DD estimates suggest substantial and highly statistically significant impacts of remediation on the algebra section of the PLAN exam administered immediately after freshman year, with remediation increasing algebra achievement by 0.08 standard deviations. This effect comes, however, entirely from the first treatment cohort, whose algebra scores rose by 0.15 standard deviations. Column (4) shows similar but much smaller estimates for the geometry section of that exam. There is no evidence of achievement gains on either the subsequent PLAN exam or the ACT, with point estimates close to zero.

The RD estimates suggest impacts on the first PLAN exam's algebra section of 0.09 standard deviations, similar in magnitude to the DD estimates but measured less precisely. There is no evidence in column (4) of any impact on that exam's geometry section. For students near the threshold, remediation did have longer run impacts on achievement, improving algebra achievement on the second PLAN exam and overall math achievement on the ACT exam by a substantial and statistically significant 0.15 standard deviations. These effects were nearly identical across the two cohorts. Table 6 thus suggests that the average remediated student showed only modest and short-lived achievement gains but that remediated students near the threshold experienced substantial achievement gains persisting at least two years after the end of the remediation itself. Perhaps most importantly, these gains are evident on the ACT exam, which is used by many colleges as part of the admissions process.

Table 7 explores the impact of remediation on educational attainment. The DD estimates suggest that remediation increased the proportion of students graduating within four years by a statistically significant 2.9 percentage points, from a base of 40.8 percentage points, an effect largely driven by the first cohort. College enrollment rates also rose by a 2.9 percentage points, from a base of 25.4 percentage points, an effect that appears similar across the two cohorts. This overall impact in college enrollment is evenly split between increases in two- and four-year college enrollment. The RD estimates suggests that remediation increased the four year graduation rate by a substantial and statistically significant 9.1 percentage points, from a base of 50.5 percentage points. College enrollment rates rose by a similarly substantial and marginally significant 7.6 percentage points, from a base of 34.8 percentage points, driven largely by increased enrollment in two-year colleges. Both the graduation and college enrollment impacts were somewhat larger for the first cohort than the second.

Tables 5-7 imply that, for the average remediated student, remediation had little impact on algebra passing rates and test scores but did improve passing rates for later years' math classes. This increased probability of fulfilling academic expectations may explain why the average remediated student was more likely to graduate from high school and enroll in college. For students near the threshold, remediation had little impact on short run test scores and passing rates for later years' math classes but substantially improved long run test scores, graduation rates and college enrollment rates. These results highlight both that educational interventions may have important long run impacts even when short run impacts are minimal and that such interventions may operate through different channels for students of different skill levels.

VI. Robustness, Heterogeneity and Spillovers

Table 8 shows robustness checks for the central results of the previous tables. Panel (A) shows that the difference-in-difference estimates are generally robust to exclusion of the control variables and to inclusion of high school fixed effects. Panel (B) shows the same for the discontinuity specification, as well as for the use of different bandwidths.

Table 9 explores whether the impacts of remediation varied by the academic skill of the remediated student. Panel (A) interacts remediation (and its instruments) with three levels of mathematical skill as measured by eighth grade test scores. The estimates here show little consistent pattern suggesting that any one subgroup benefitted particularly from remediation. Panel (B) suggests a much clearer pattern. At the discontinuity, all students have equal math skill, so we interact remediation with a measure of eight grade reading skill, dividing the sample into those above and below the median reading ability of students at the remediation threshold. Columns (6)-(8) suggest that the bulk of the positive long run impact of remediation came through its effect on low skilled readers, perhaps because the intervention focused on reading and writing skills in the context of learning algebra.

Finally, Table 10 explores whether increased focused on mathematics and substitution away from other freshman coursework had any negative impacts on achievement in other disciplines. Panel (A) suggests that remediation actually raised science achievement while having little impact on verbal achievement and later grades in non-math classes. Panel (C) suggests that, for students near the threshold, remediation had little impact on science scores but did improve verbal scores and later grades in non-math classes. This suggests that the spillovers from remediation were, if anything, positive.

Conclusion

We show long run positive and substantial impacts of algebra remediation on college entrance exam scores, high school graduation rates and college enrollment rates. This is the first evidence we know of to demonstrate long run impacts of remediation in an American urban high school setting. Given the number of school districts that struggle with low-performing and at-risk students, the possibility that the double-dose math program improved high school graduation and college enrollment rates is extraordinarily promising.

References

Aaronson, D., Barrow, L., and Sander, W. (2002). "Teachers and student achievement in the Chicago public high schools." Working Paper Series WP-02-28, Federal Reserve Bank of Chicago.

Bryk, A. S., Lee, V. E., and Holland, P. B. (1993). *Catholic schools and the common good*. Cambridge, MA: Harvard University Press.

Bettinger, E. and Long, B.T. (2009). "Addressing the Needs of Under-Prepared College Students: Does College Remediation Work?" Journal of Human Resources. 44(3): 736-771

Calcagno, C. and Long, B.T. (2008). "The Impact of Postsecondary Remediation Using a Regression Discontinuity Approach: Addressing Endogenous Sorting and Noncompliance." NCPR Working Paper.

Cullen, J., Jacob, B. and Levitt, S. (2005). "The Impact of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools." Journal of Public Economics. 89(5-6): 729-760.

Cullen, J., Jacob, B. and Levitt, S. (2006). "The Effect of School Choice on Student Outcomes: Evidence from Randomized Lotteries." Econometrica. 74(5): 1191-1230.

Goodman, J. (2009). "The Labor of Division: The Labor Market Returns to Compulsory Math Coursework." Unpublished manuscript, Harvard Kennedy School.

Jacob, B. (2005). "Accountability, Incentives and Behavior: Evidence from School Reform in Chicago." Journal of Public Economics. 89(5-6): 761-796.

Jacob, B. and Lefgren, L. (2003). "Are Idle Hands the Devil's Workshop? Incapacitation, Concentration and Juvenile Crime." NBER Working Papers No. 9653.

Jacob, B. and Lefgren, L. (2004). "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis." Review of Economics and Statistics. LXXXVI (1): 226-244.

Jacob, B. and Levitt, S. (2003). "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." Quarterly Journal of Economics. 118(3): 843-877.

Lavy, V. and Schlosser, A. (2005). "Targeted Remedial Education for Underperforming Teenagers: Costs and Benefits." Journal of Labor Economics. 23(4): 839-874.

Lee, V. E., Croninger, R.G. and Smith, J.B. (1997). "Course taking, equity, and mathematics learning: Testing the constrained curriculum hypothesis in U.S. secondary schools." Educational Evaluation and Policy Analysis. 19(2): 99-121.

Lee, V. E., Smith, J.B, and Croninger, R.G. (1997). "How high school organization influences the equitable distribution of learning in mathematics and science." Sociology of Education. 70(2): 129-152.

National Commission on Excellence in Education. (1983). A Nation at risk: The imperative for education reform. Washington, DC: U.S. Government Printing Office.

National Governors Association Center for Best Practices. (2005a). Getting it done: Ten steps for a state action agenda. (Retrieved from: http://www.nga.org/Files/pdf/05warnerguide.pdf.)

National Governors Association Center for Best Practices. (2005b). A profile of state action to improve America's high schools.

Nomi, T., & Allensworth, E. (2009). "Double-dose" algebra as an alternative strategy to remediation: Effects on students' academic outcomes." Journal of Research on Educational Effectiveness. 2: 111-148.

Roderick, M., & Camburn, E. (1999). "Risk and recovery from course failure in the early years of high school." American Educational Research Journal. 36: 303-343.

Starkel, R., Martinez, J., and Price, K. (2006). "Two-period Algebra in the 05-06 school year: Implementation Report."

U.S. Department of Education, National Center for Education Statistics (2004). The Condition of Education 2004 (NCES 2004-077). Washington, DC: U.S. Government Printing Office.

U.S. Department of Education, National Center for Education Statistics. (2011). *Digest of Education Statistics, 2010* (NCES 2011-015), Table 235.

Wenzel, S., Lawal, K., Conway, B., Fendt, C., and Stoelinga, S. (2005). "*Algebra problem solving: Teachers talk about their experiences, December 2004*." Retrieved May 3, 2012 from http://www.prairiegroup.org/images/Algebra_Problem_Solving,_Jan._04.pdf

Figure 1: Remediation rates

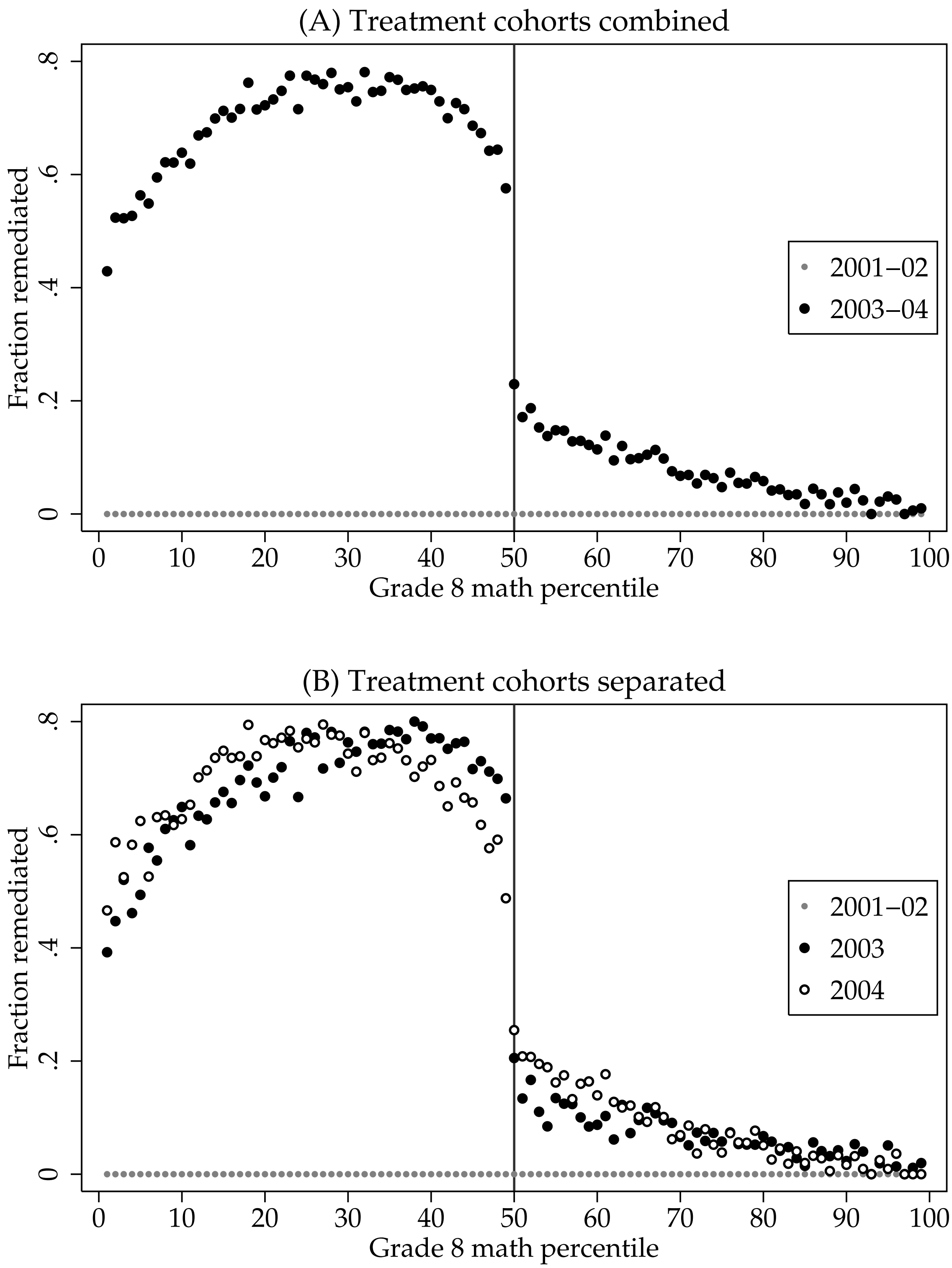
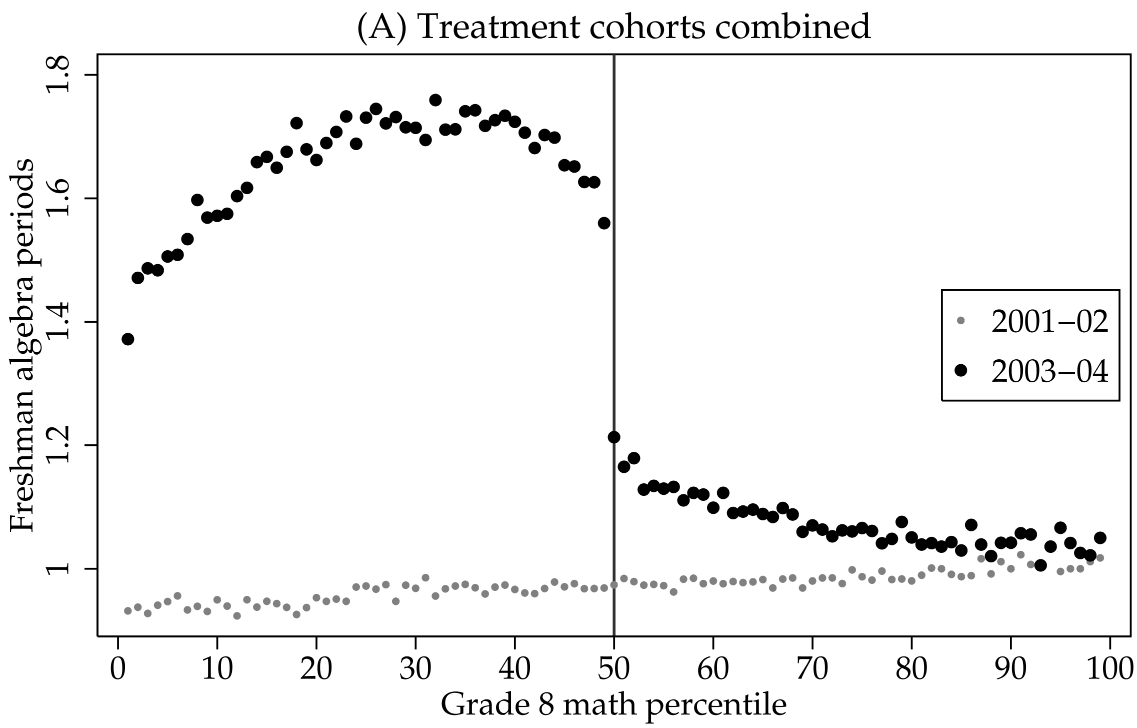


Figure 2: Freshman algebra periods



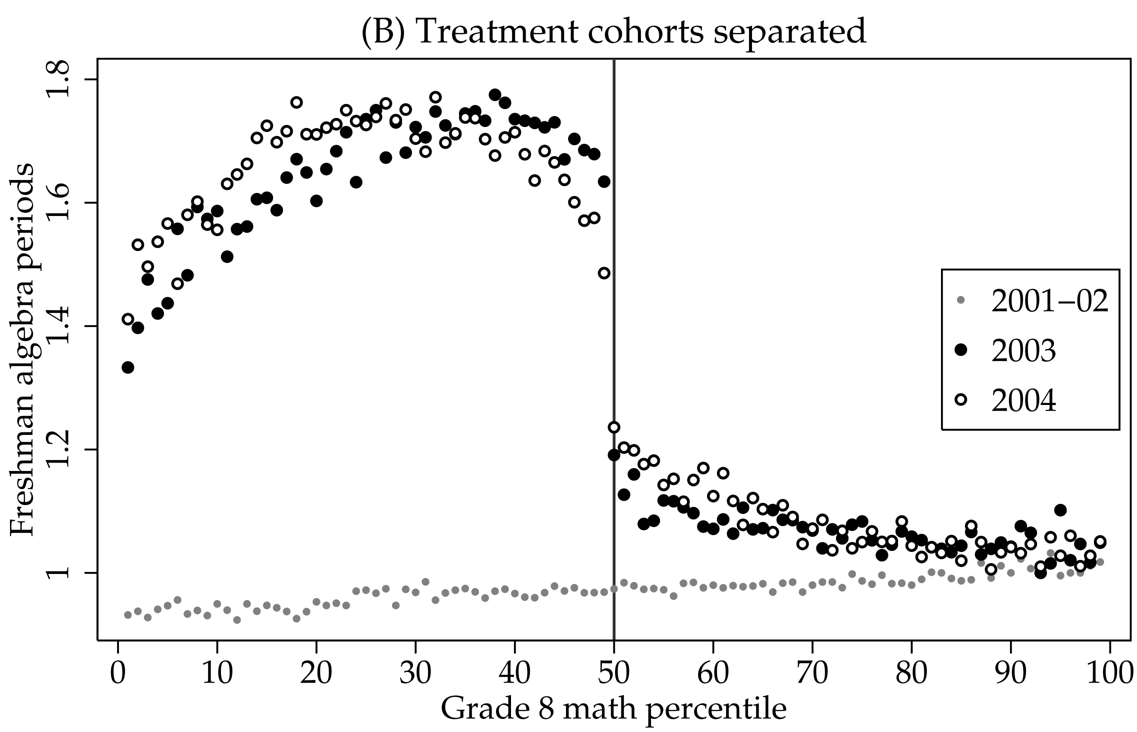


Figure 3: SD of peer math skill

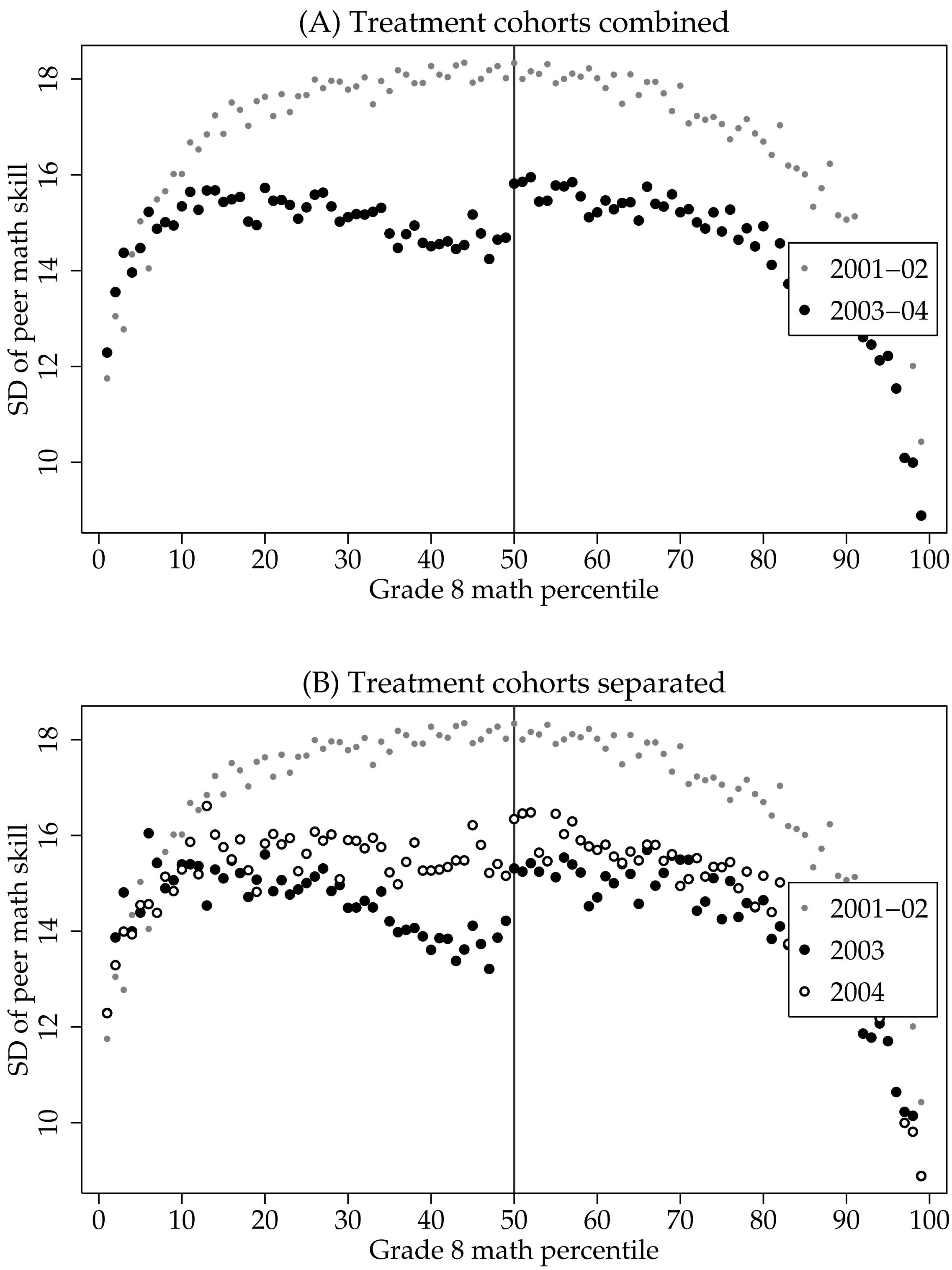
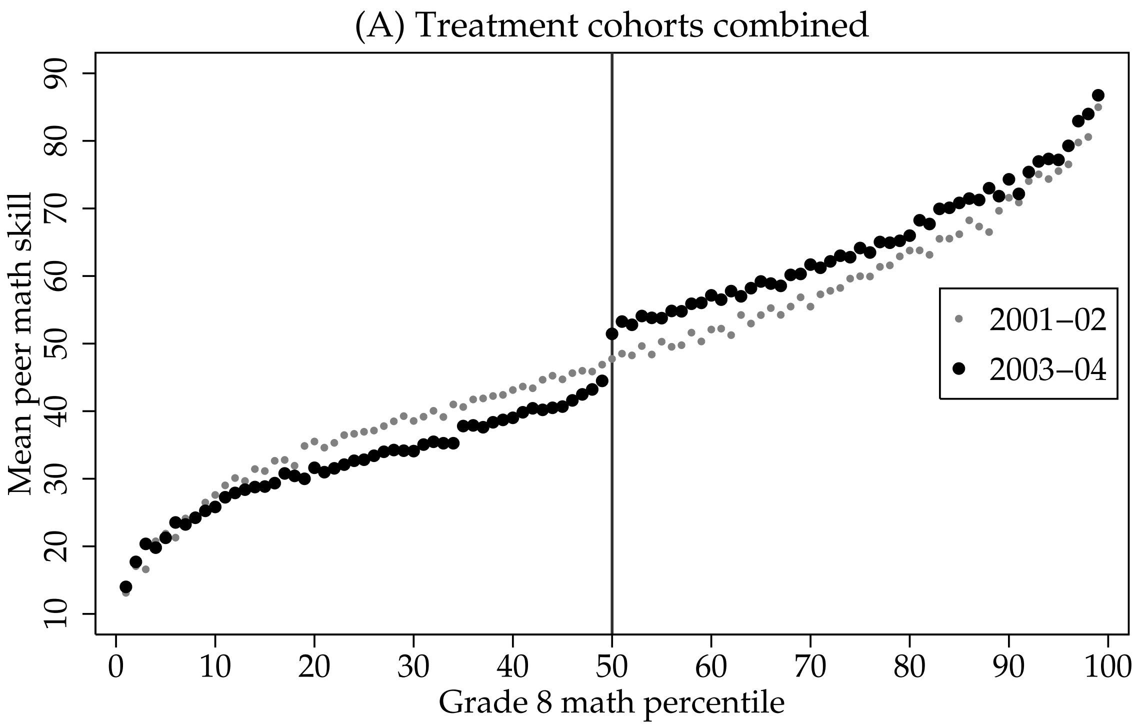


Figure 4: Mean peer math skill



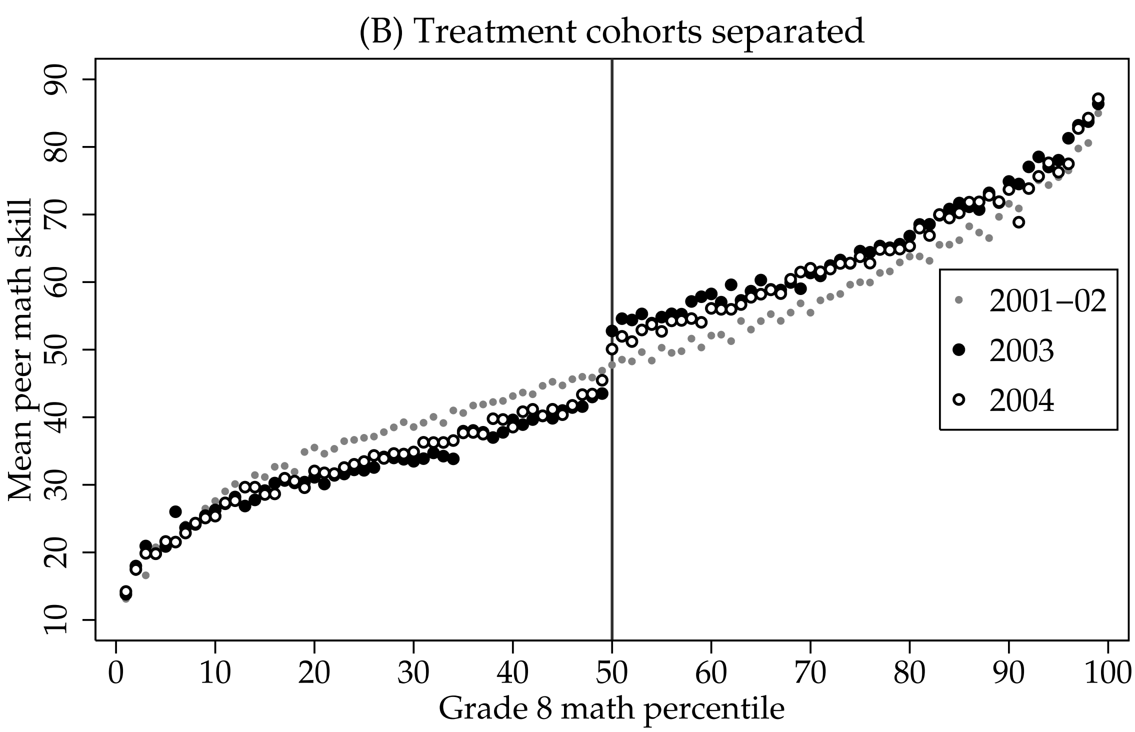


Figure 5: Earned A or B in freshman algebra

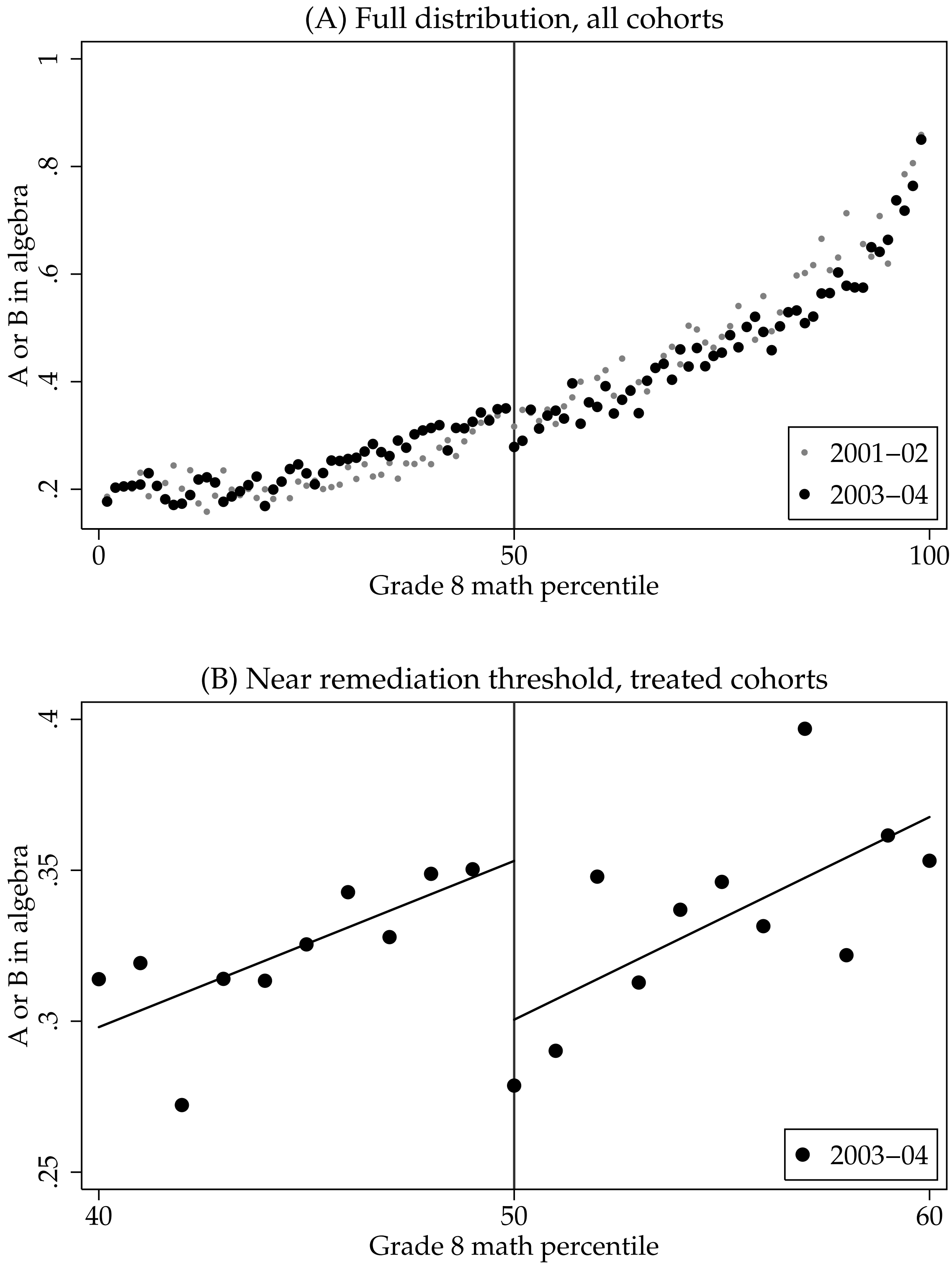
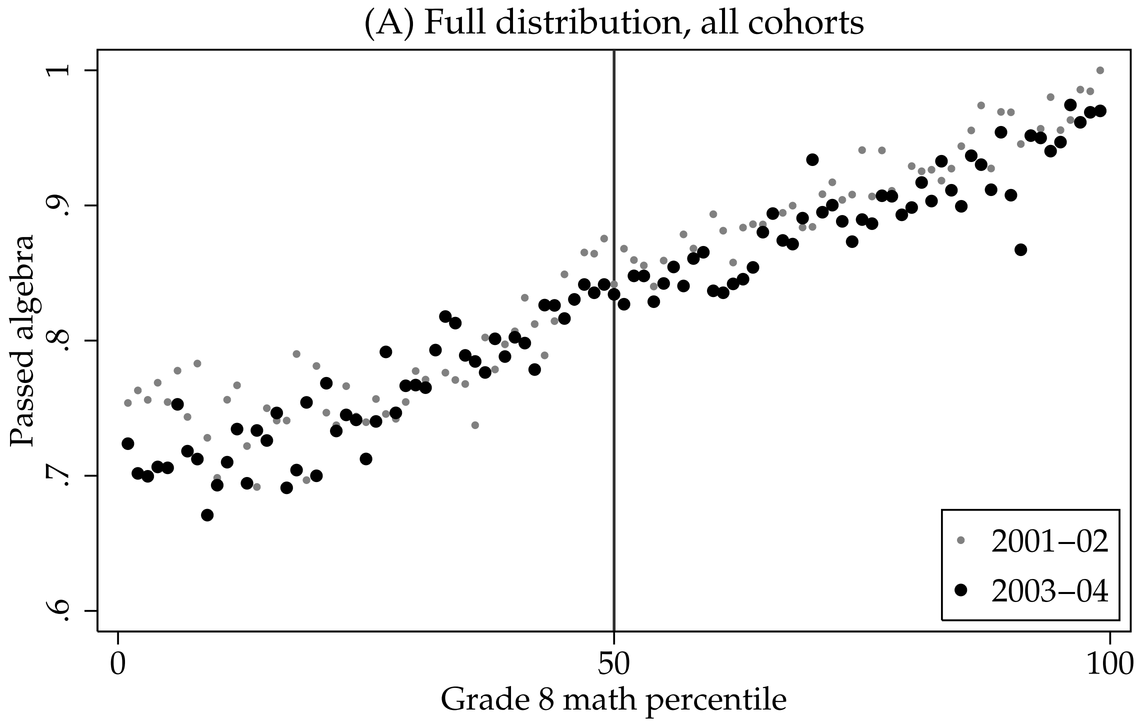


Figure 6: Passed freshman algebra



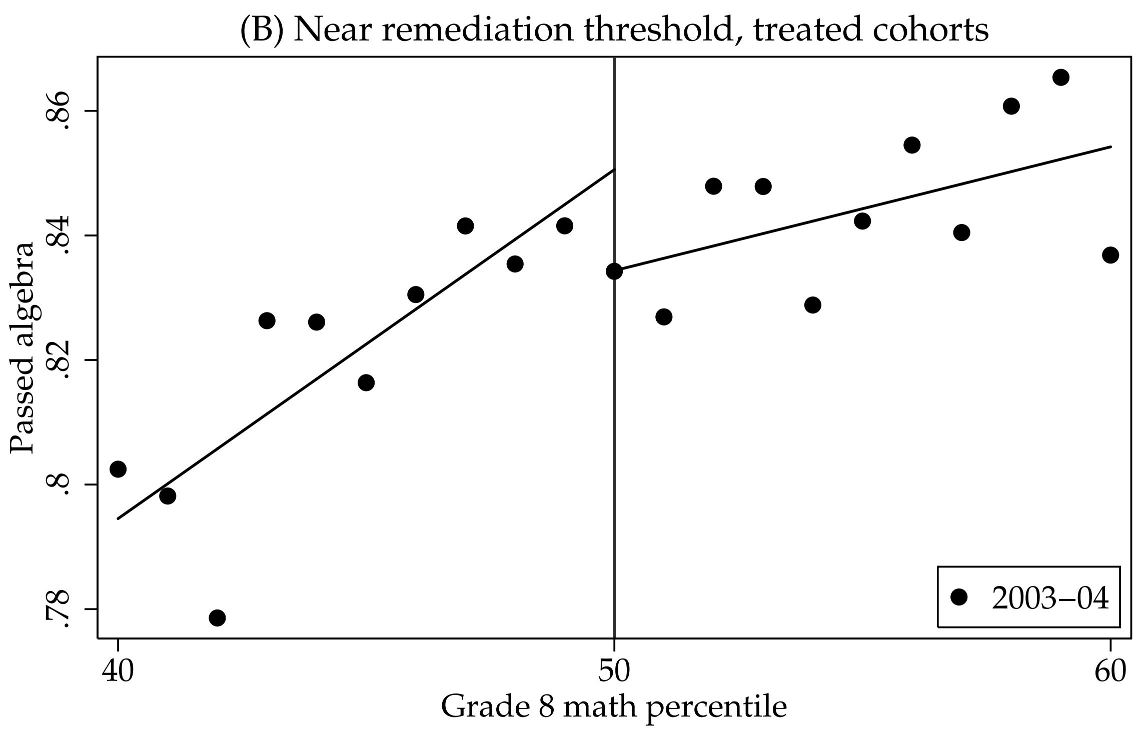


Figure 7: ACT math score

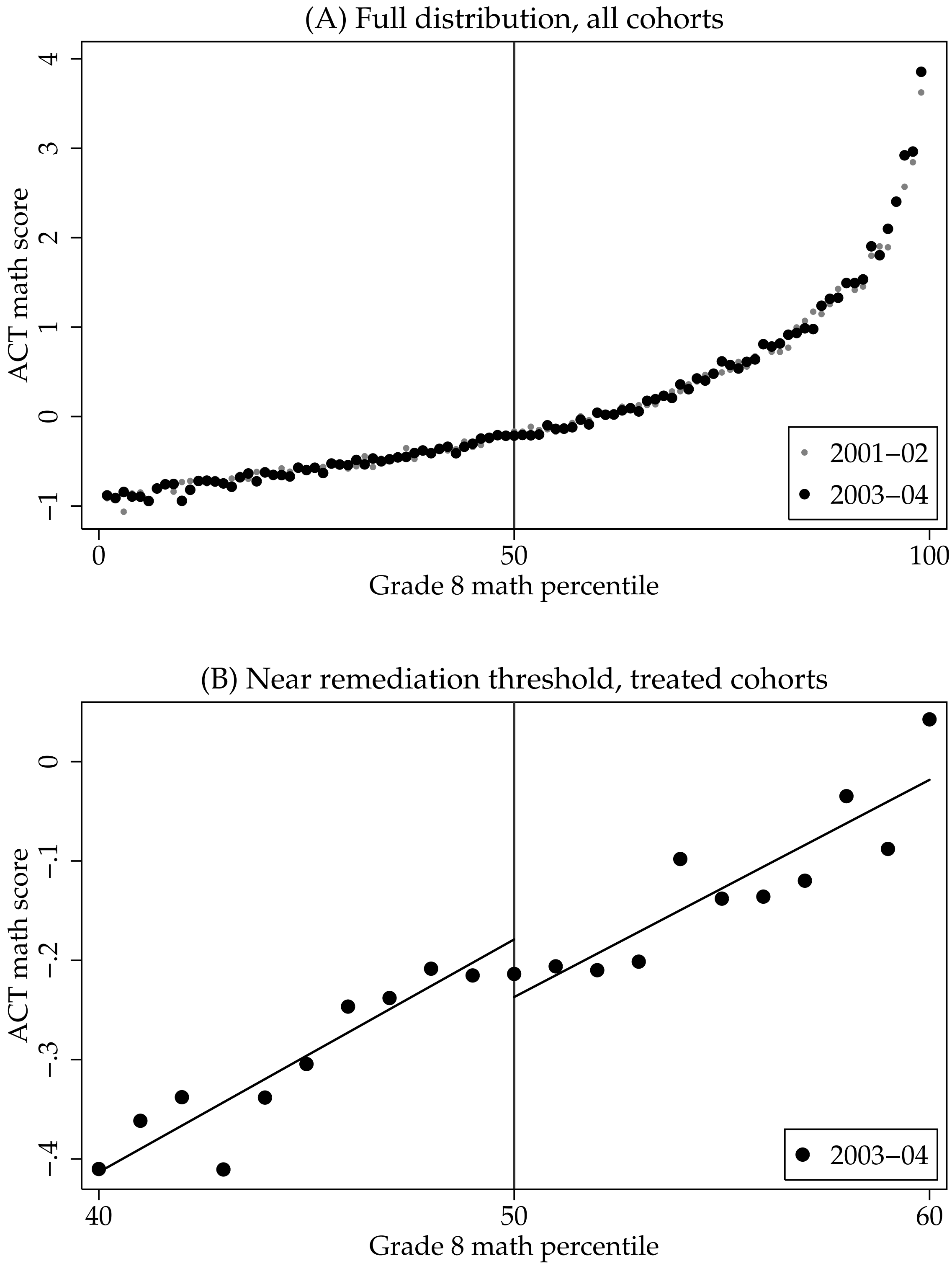


Figure 8: Graduated in four years

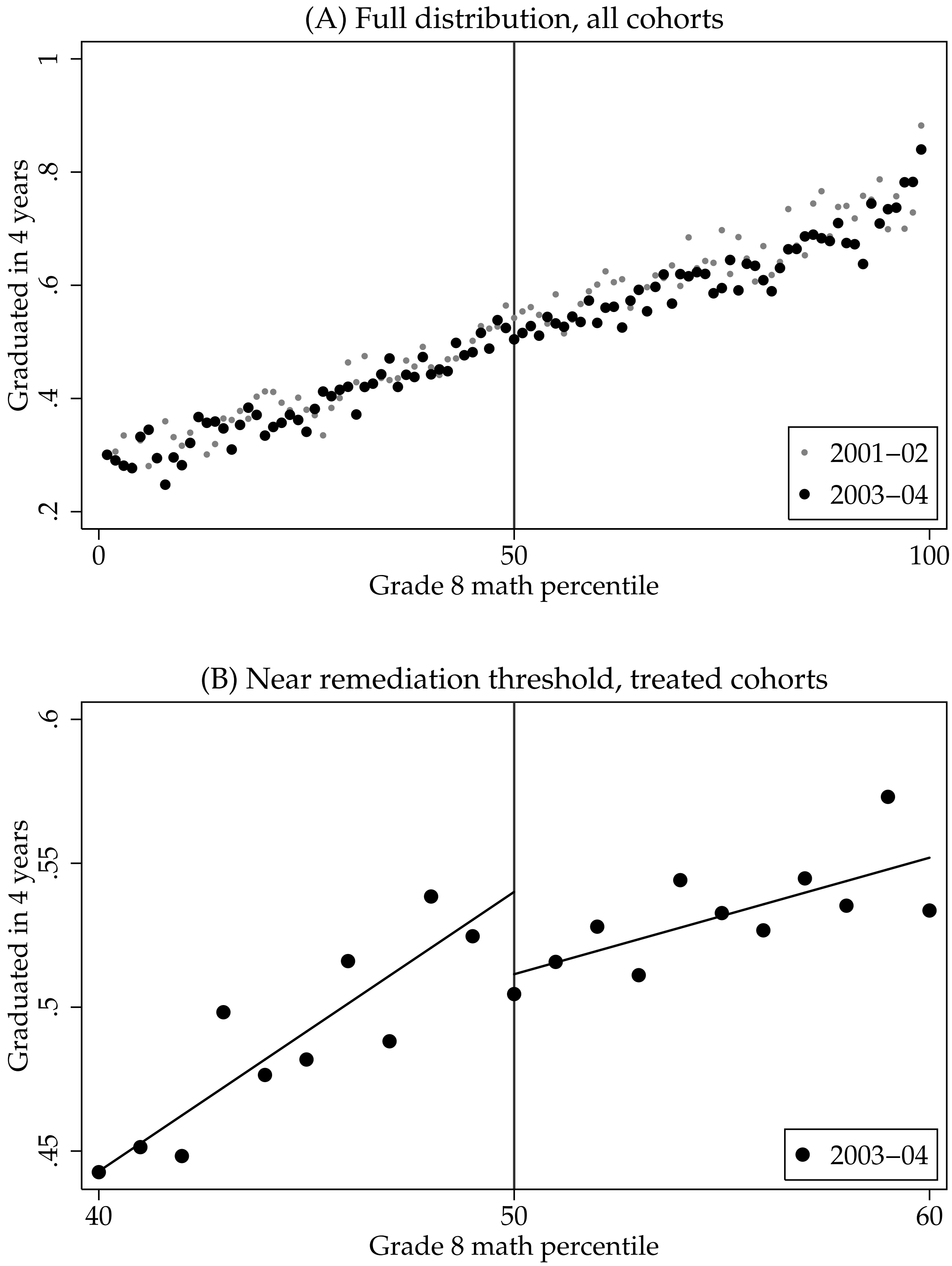


Figure 9: Enrolled in college

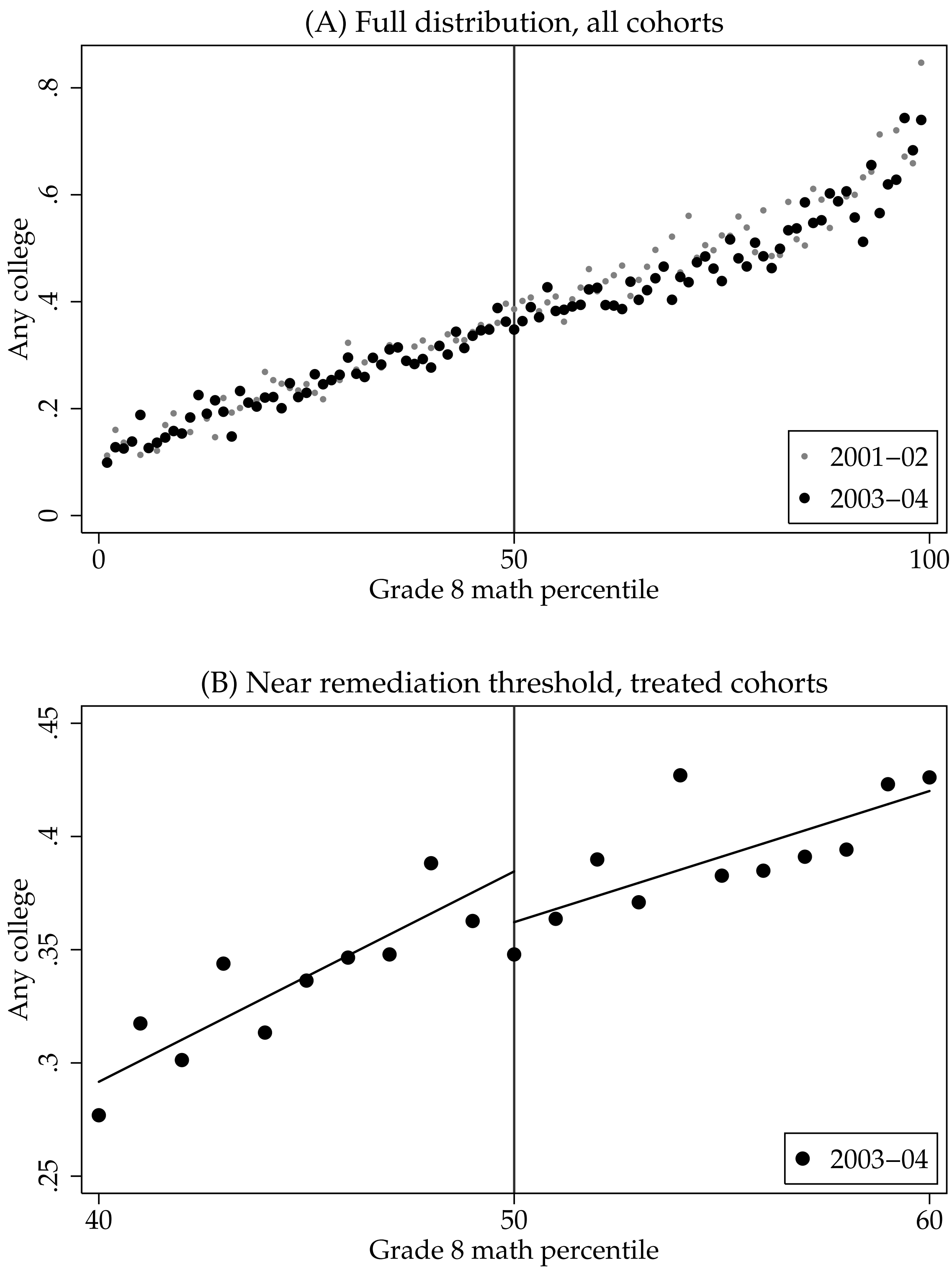


Table 1: Summary Statistics

Cohort	(1) 2001	(2) 2002	(3) 2003	(4) 2004
(A) Demographics				
Female	0.50	0.50	0.50	0.50
Black	0.57	0.54	0.59	0.56
Hispanic	0.33	0.35	0.32	0.35
Special education	0.20	0.20	0.19	0.20
(B) Remediation				
Grade 8 math percentile	42.41	47.63	45.80	45.37
Eligible for remediation	0.61	0.52	0.55	0.56
Remediated	0.00	0.00	0.42	0.43
Freshman math courses	0.96	0.98	1.39	1.41
SD of peer skill	17.23	16.93	14.47	15.21
(C) Achievement				
A or B in algebra	0.32	0.35	0.32	0.34
Passed algebra	0.82	0.84	0.81	0.82
Fall 10 math score	0.00	-0.00	-0.00	-0.00
Fall 11 math score	0.00	0.00	-0.00	-0.00
ACT math score	-0.00	0.00	-0.00	-0.00
(D) Attainment				
Graduated in 4 years	0.50	0.50	0.47	0.50
Any college	0.36	0.35	0.34	0.34
Two-year college	0.19	0.18	0.17	0.17
Four-year college	0.17	0.17	0.17	0.17
N	19,530	19,008	20,357	21,007

Notes: Mean values of each variable are shown by cohort.

Table 2: Eligibility as an Instrument for Remediation

Y = remediated	(1) No controls	(2) Demographic controls	(3) High school fixed effects
(A) DD			
Eligible * after	0.594***	0.593***	0.584***
	(0.022)	(0.022)	(0.022)
(B) DD by cohort			
Eligible * 2003	0.606***	0.605***	0.600***
	(0.024)	(0.024)	(0.024)
Eligible * 2004	0.582***	0.581***	0.569***
	(0.028)	(0.028)	(0.028)
N	79,902	79,902	79,902
(C) RD			
Eligible for remediation	0.400***	0.399***	0.406***
	(0.038)	(0.039)	(0.037)
(D) RD by cohort			
Eligible * 2003	0.477***	0.477***	0.486***
	(0.039)	(0.039)	(0.037)
Eligible * 2004	0.322***	0.320***	0.326***
	(0.047)	(0.048)	(0.046)
N	11,546	11,546	11,546

Notes: Heteroskedasticity robust standard errors clustered by initial high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Each column in each panel represents the first-stage regression of a remediation dummy on eligibility as determined by eighth grade math score. Panels (A)-(B) are difference-in-difference specifications using all of the cohorts. Eligibility and cohort indicators are included but not shown. Panels (C)-(D) are regression discontinuity specifications using only the treated cohorts. It fits straight lines on both sides of the threshold using a bandwidth of 10 percentiles. Cohort indicators are included but not shown. Column (2) adds to column (1) controls for gender, race, census block poverty and socioeconomic status, and eighth grade math and reading scores.Column (3) adds to column (2) high school fixed effects.

Table 3: First stage heterogeneity by ability

Y = remediated	(1) Both cohorts	(2) 2003 cohort	(3) 2004 cohort
(A) DD			
Eligible * very low math skill	0.536***	0.513***	0.560***
	(0.034)	(0.037)	(0.038)
Eligible * low math skill	0.694***	0.704***	0.684***
	(0.022)	(0.024)	(0.027)
Eligible * medium math skill	0.628***	0.690***	0.568***
	(0.025)	(0.024)	(0.035)
N	79,902	58,894	59,545
(B) RD			
Eligible * below average reader	0.393***	0.498***	0.287***
	(0.039)	(0.043)	(0.049)
Eligible * above average reader	0.429***	0.564***	0.294***
	(0.041)	(0.045)	(0.052)
N	11,546	5,760	5,786

Notes: Heteroskedasticity robust standard errors clustered by initial high school are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$). Each column in each panel represents the first-stage regression of a remediation dummy on eligibility as determined by eighth grade math score. Panel (A) is a difference-in-difference specification using all of the cohorts. Eligibility and cohort indicators are included but not shown. Panel (B) is a regression discontinuity specification using only the treated cohorts. It fits straight lines on both sides of the threshold using a bandwidth of 10 percentiles. Cohort indicators are included but not shown. Columns (2)-(4) include interactions of eligibility with student characteristics. All regressions include controls for gender, race, census block poverty and socioeconomic status, eighth grade math and reading scores, and indicators for observations in which any of those variables have been imputed.

Table 4: Remediation, Freshman Coursework and Peer Ability

	(1) Math courses	(2) Academic courses	(3) Other courses	(4) Total courses	(5) Mean peer ability	(6) St. dev. of ability	(7) Distance to mean peer	(8) Class size
(A) DD								
Remediated	0.986***	-0.098	-0.705***	0.184***	-12.884***	0.229	2.871***	0.692
	(0.010)	(0.063)	(0.073)	(0.050)	(1.322)	(0.582)	(0.628)	(0.632)
(B) DD by cohort								
Remediated * 2003	0.978***	-0.114*	-0.704***	0.160***	-14.194***	-0.046	3.184***	0.968
	(0.013)	(0.068)	(0.072)	(0.059)	(1.414)	(0.690)	(0.670)	(0.798)
Remediated * 2004	0.994***	-0.081	-0.705***	0.208***	-11.564***	0.506	2.555***	0.415
	(0.009)	(0.073)	(0.092)	(0.061)	(1.676)	(0.568)	(0.669)	(0.636)
μ	0.957	3.377	2.363	6.696	35.594	17.029	12.737	19.177
N	79,902	79,902	79,902	79,902	79,902	79,793	79,902	79,902
(C) RD								
Remediated	0.991***	-0.180*	-0.676***	0.134	-19.442***	-3.074***	3.218***	2.435***
	(0.021)	(0.095)	(0.086)	(0.084)	(1.430)	(0.759)	(1.214)	(0.765)
(D) RD by cohort								
Remediated * 2003	0.977***	-0.178**	-0.680***	0.118	-19.502***	-3.329***	3.352***	2.365***
	(0.022)	(0.089)	(0.085)	(0.078)	(1.411)	(0.826)	(1.050)	(0.837)
Remediated * 2004	1.015***	-0.184	-0.670***	0.162	-19.336***	-2.622***	2.980*	2.560***
	(0.023)	(0.121)	(0.100)	(0.109)	(2.109)	(0.860)	(1.730)	(0.877)
μ	1.213	3.512	2.252	6.977	51.451	15.817	10.717	20.893
N	11,546	11,546	11,546	11,546	11,546	11,540	11,546	11,546

Notes: Heteroskedasticity robust standard errors clustered by initial high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Each column in each panel represents the instrumental variables regression of the listed outcome on a remediation indicator, where the first-stage is given in table 2. Panels (A)-(B) are difference-in-difference specifications with all of the cohorts. Panels (C)-(D) are regression discontinuity specifications with only the treated cohorts.

Table 5: Impacts of Remediation on Math Coursework

	(1) A or B in algebra	(2) Passed algebra	(3) Passed geom.	(4) Passed trig.	(5) Math GPA 10-12
(A) DD					
Remediated	0.085***	0.015	0.043**	0.029*	0.056
	(0.023)	(0.015)	(0.020)	(0.017)	(0.037)
(B) DD by cohort					
Remediated * 2003	0.088***	0.012	0.050**	0.025	0.058
	(0.025)	(0.017)	(0.022)	(0.018)	(0.042)
Remediated * 2004	0.083***	0.017	0.036	0.034*	0.054
	(0.027)	(0.017)	(0.029)	(0.020)	(0.044)
μ	0.235	0.775	0.438	0.364	1.210
N	79,902	79,902	79,902	79,902	71,122
(C) RD					
Remediated	0.145***	0.050*	-0.016	0.055	0.156*
	(0.047)	(0.031)	(0.049)	(0.039)	(0.082)
(D) RD by cohort					
Remediated * 2003	0.144***	0.040	0.005	0.049	0.161**
	(0.047)	(0.026)	(0.044)	(0.038)	(0.078)
Remediated * 2004	0.146**	0.070	-0.052	0.067	0.149
	(0.058)	(0.045)	(0.067)	(0.051)	(0.107)
μ	0.279	0.834	0.490	0.477	1.393
N	11,546	11,546	11,546	11,546	10,475

Notes: Heteroskedasticity robust standard errors clustered by initial high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Each column in each panel represents the instrumental variables regression of the listed outcome on a remediation indicator, where the first-stage is given in table 2. Panels (A)-(B) are difference-in-difference specifications with all of the cohorts. Panels (C)-(D) are regression discontinuity specifications with only the treated cohorts.

Table 6: Impacts of Remediation on Math Test Scores

	(1) Fall 10 algebra	(2) Fall 11 algebra	(3) Spring 11 ACT math	(4) Fall 10 geometry	(5) Fall 11 geometry
(A) DD					
Remediated	0.084***	0.025	0.003	0.033	0.028
	(0.025)	(0.024)	(0.029)	(0.030)	(0.031)
(B) DD by cohort					
Remediated * 2003	0.147***	0.013	0.046	0.080**	-0.008
	(0.027)	(0.028)	(0.030)	(0.036)	(0.030)
Remediated * 2004	0.020	0.037	-0.039	-0.016	0.066*
	(0.030)	(0.030)	(0.036)	(0.034)	(0.040)
μ	-0.540	-0.533	-0.535	-0.392	-0.414
N	53,649	49,915	44,799	53,647	49,915
(C) RD					
Remediated	0.088	0.153***	0.152**	0.012	0.106
	(0.070)	(0.059)	(0.071)	(0.080)	(0.084)
(D) RD by cohort					
Remediated * 2003	0.082	0.142**	0.139**	0.025	0.059
	(0.064)	(0.056)	(0.064)	(0.072)	(0.077)
Remediated * 2004	0.098	0.172**	0.172**	-0.010	0.186*
	(0.089)	(0.078)	(0.087)	(0.105)	(0.109)
μ	-0.106	-0.126	-0.214	-0.103	-0.131
N	8,227	7,446	6,716	8,227	7,446

Notes: Heteroskedasticity robust standard errors clustered by initial high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Each column in each panel represents the instrumental variables regression of the listed outcome on a remediation indicator, where the first-stage is given in table 2. Panels (A)-(B) are difference-in-difference specifications with all of the cohorts. Panels (C)-(D) are regression discontinuity specifications with only the treated cohorts.

Table 7: Impacts of Remediation on Attainment

	(1) Graduated in 4 years	(2) Enrolled in college	(3) Two year college	(4) Four year college
(A) DD				
Remediated	0.029**	0.029***	0.015*	0.014
	(0.014)	(0.011)	(0.009)	(0.011)
(B) DD by cohort				
Remediated * 2003	0.042**	0.033**	0.006	0.028**
	(0.017)	(0.014)	(0.012)	(0.012)
Remediated * 2004	0.017	0.025*	0.025**	-0.000
	(0.016)	(0.015)	(0.011)	(0.012)
μ	0.408	0.254	0.168	0.085
N	79,902	79,902	79,902	79,902
(C) RD				
Remediated	0.091**	0.076*	0.062	0.014
	(0.046)	(0.043)	(0.041)	(0.027)
(D) RD by cohort				
Remediated * 2003	0.109**	0.088**	0.053	0.035
	(0.043)	(0.039)	(0.039)	(0.025)
Remediated * 2004	0.060	0.054	0.078	-0.023
	(0.059)	(0.058)	(0.053)	(0.040)
μ	0.505	0.348	0.179	0.169
N	11,546	11,546	11,546	11,546

Notes: Heteroskedasticity robust standard errors clustered by initial high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Each column in each panel represents the instrumental variables regression of the listed outcome on a remediation indicator, where the first-stage is given in table 2. Panels (A)-(B) are difference-in-difference specifications with all of the cohorts. Panels (C)-(D) are regression discontinuity specifications with only the treated cohorts.

Table 8: Robustness checks

	(1) Passed algebra	(2) Passed geometry	(3) Passed trig.	(4) Fall 10 algebra	(5) Fall 11 algebra	(6) Spring 11 ACT math	(7) Graduated in 4 years	(8) Enrolled in college
(A) DD								
No controls	0.013	0.046**	0.031*	0.104***	0.046	0.025	0.038***	0.043***
	(0.015)	(0.021)	(0.017)	(0.035)	(0.034)	(0.039)	(0.015)	(0.012)
Demographic controls	0.012	0.041**	0.028*	0.085***	0.025	0.003	0.025*	0.028**
	(0.015)	(0.020)	(0.017)	(0.025)	(0.024)	(0.029)	(0.014)	(0.011)
High school fixed effects	0.008	0.046**	0.027	0.073***	0.019	0.005	0.022*	0.025**
	(0.014)	(0.020)	(0.016)	(0.027)	(0.024)	(0.028)	(0.013)	(0.011)
(B) RD								
No controls	0.041	-0.030	0.039	0.069	0.128**	0.140*	0.078	0.056
	(0.032)	(0.049)	(0.039)	(0.073)	(0.060)	(0.074)	(0.050)	(0.047)
Demographic controls	0.048	-0.017	0.055	0.085	0.152***	0.152**	0.093**	0.075*
	(0.030)	(0.049)	(0.039)	(0.070)	(0.059)	(0.071)	(0.046)	(0.043)
High school fixed effects	0.040	-0.037	0.058	0.044	0.120**	0.126*	0.085*	0.064
	(0.029)	(0.047)	(0.038)	(0.065)	(0.059)	(0.073)	(0.045)	(0.043)
Linear, bandwidth = 5	0.053	0.075	0.095	0.071	0.264**	0.168	0.127*	0.119*
	(0.053)	(0.077)	(0.069)	(0.107)	(0.104)	(0.110)	(0.075)	(0.061)
Quadratic, bandwidth = 15	0.060	0.026	0.086	0.071	0.168**	0.185**	0.134**	0.137**
	(0.044)	(0.069)	(0.059)	(0.084)	(0.084)	(0.090)	(0.065)	(0.065)

Notes: Heteroskedasticity robust standard errors clustered by initial high school are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$).

Table 9: Heterogeneity by skill

	(1) Passed algebra	(2) Passed geometry	(3) Passed trig.	(4) Fall 10 algebra	(5) Fall 11 algebra	(6) Spring 11 ACT math	(7) Graduated in 4 years	(8) Enrolled in college
(A) DD								
Remediated * very low math skill	-0.015	0.042	0.039	-0.050	0.004	0.002	0.037**	0.045***
	(0.019)	(0.029)	(0.024)	(0.034)	(0.032)	(0.035)	(0.018)	(0.014)
Remediated * low math skill	0.034**	0.051***	0.029*	0.102***	0.007	-0.009	0.020	0.021*
	(0.015)	(0.019)	(0.016)	(0.026)	(0.024)	(0.027)	(0.015)	(0.011)
Remediated * medium math skill	0.010	0.014	0.006	0.116***	0.035	0.001	0.021	0.012
	(0.013)	(0.017)	(0.016)	(0.028)	(0.026)	(0.026)	(0.017)	(0.015)
p	0.014	0.073	0.314	0.000	0.537	0.896	0.613	0.080
N	79,902	79,902	79,902	53,649	49,915	44,799	79,902	79,902
(B) RD								
Remediated * below average reader	0.060*	0.008	0.081*	0.094	0.182***	0.221***	0.130**	0.119**
	(0.036)	(0.054)	(0.044)	(0.084)	(0.065)	(0.074)	(0.056)	(0.051)
Remediated * above average reader	0.041	-0.039	0.030	0.080	0.128*	0.089	0.052	0.033
	(0.029)	(0.049)	(0.042)	(0.066)	(0.066)	(0.077)	(0.043)	(0.043)
p	0.439	0.146	0.142	0.812	0.351	0.011	0.042	0.018
N	11,546	11,546	11,546	8,227	7,446	6,716	11,546	11,546

Notes: Heteroskedasticity robust standard errors clustered by initial high school are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$). Panels (A) and (B) are respectively the difference-in-difference and regression discontinuity specifications with demographic controls used in previous tables, with further interactions by skill. In panel (A), students are divided into skill groups as defined by 8th grade reading score. In panel (B), students are divided into skill groups as defined by 8th grade math score.

Table 10: Spillovers into other subjects

	(1) Fall 10 science	(2) Fall 11 science	(3) ACT science	(4) Fall 10 verbal	(5) Fall 11 verbal	(6) ACT verbal	(7) Nonmath GPA 10-12
(A) DD							
Remediated	0.095***	0.143***	0.105***	-0.018	-0.001	0.018	0.029
	(0.031)	(0.037)	(0.025)	(0.025)	(0.027)	(0.025)	(0.031)
(B) DD by cohort							
Remediated * 2003	0.027	0.241***	0.126***	-0.008	0.014	0.044	0.027
	(0.031)	(0.042)	(0.027)	(0.025)	(0.030)	(0.028)	(0.034)
Remediated * 2004	0.164***	0.040	0.084**	-0.029	-0.017	-0.007	0.031
	(0.040)	(0.040)	(0.034)	(0.032)	(0.031)	(0.030)	(0.035)
μ	-0.434	-0.449	-0.510	-0.474	-0.476	-0.537	1.565
N	53,441	49,765	44,758	53,828	50,056	44,806	71,906
(C) RD							
Remediated	-0.005	0.033	0.035	0.101*	0.049	0.182***	0.159*
	(0.083)	(0.086)	(0.084)	(0.062)	(0.068)	(0.065)	(0.084)
(D) RD by cohort							
Remediated * 2003	0.051	0.093	0.069	0.096*	0.076	0.169***	0.160**
	(0.076)	(0.081)	(0.078)	(0.057)	(0.061)	(0.059)	(0.074)
Remediated * 2004	-0.104	-0.069	-0.019	0.110	0.002	0.203**	0.155
	(0.110)	(0.110)	(0.106)	(0.081)	(0.095)	(0.092)	(0.119)
μ	-0.048	-0.050	-0.172	-0.070	-0.008	-0.127	1.774
N	8,202	7,429	6,715	8,246	7,460	6,716	10,546

Notes: Heteroskedasticity robust standard errors clustered by initial high school are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$). Each column in each panel represents the instrumental variables regression of the listed outcome on an indicator for remediation, where the first-stage is given in table 2. Panels (A)-(C) are regression discontinuity specifications with the two treated cohorts. Panels (D)-(F) are difference-in-difference specifications with all four cohorts. Specifications and included controls are the same as those in table 2.