College Preparatory Curriculum for All:
Consequences of Ninth-Grade Course Taking in Algebra and English on Academic Outcomes in Chicago

Elaine Allensworth
Takako Nomi
Nicholas Montgomery
Consortium on Chicago School Research at the University of Chicago

Valerie E. Lee
University of Michigan

December 3, 2008

Keywords: Curriculum, High school, Algebra, Ninth grade

We gratefully acknowledge support for this work from the Institute for Educational Sciences, U.S. Department of Education, grant # R305R060059. However, the responsibility for the findings and conclusions reported in this manuscript is the authors', not the U.S. Department of Education. We gratefully acknowledge the support of Nicole Tysvaer and Kristen Neishi in the production of this article. For information on this study, contact the first author <elainea@ccsr.uchicago.edu>.

College-Preparatory Curriculum for All:
Consequences of Ninth-Grade Course Taking in Algebra and English on Academic Outcomes in Chicago

Keywords: Curriculum, High school, Algebra, Ninth grade, Tracking

Abstract

There is a national movement to universalize the high-school curriculum so that all students graduate prepared for college. Here we evaluate a policy in Chicago that ended remedial classes and mandated college-preparatory coursework for all students. Using an interrupted time-series cohort design with multiple comparisons, we found that the policy reduced inequities in ninth grade coursework by entering ability, race/ethnicity, and special education status. Although more students completed ninth grade with credits in Algebra and English I, failure rates increased, grades declined slightly, test scores did not improve, and students were no more likely to enter college. Although few benefits resulted from universalizing college preparatory coursework among ninth graders, neither did dropout rates increase. Possible explanations are discussed.

Background

Two Philosophical Stances about High-School Curriculum

What should students learn in high school? Should all students learn the same skills and content, or should coursework reflect students' abilities, their motivation to learn, and their plans for the future? Such questions have been hotly debated since the inception of secondary schooling in the United States. These debates rest on a philosophical continuum. On one end is a view that all students—regardless of their educational or occupational futures—should experience intellectually challenging coursework that prepares them equally well for college or work. Supporters of this view argue for a "constrained academic curriculum" that does not differentiate students by ability, performance, or future plans. The other end of the philosophical continuum draws on a social efficiency argument—that schools have a duty to sort and match students to their future places in the social and economic system. Social efficiency advocates argue that offering only academic courses in high school overlooks two realities about students: (1) they enter high school with different intellectual capacities and skills, and (2) they aspire to disparate occupations. To address these realities, they advocate a "differentiated curriculum" that includes a broad range of academic and vocational offerings at different levels of rigor.

The social efficiency argument has predominated throughout the 20th century, reflected in a typical public high-school curriculum that is broad and diffuse—with many different courses in any subject and many of the same courses at varying difficulty levels (Angus & Mirel, 1999; Bryk, Lee, & Holland, 1993; Cremin, 1961; Kliebard, 1995; Lee & Ready, 2007; Oakes, 1985, 2005; Powell, Farrar, & Cohen, 1985; Yonezawa, Wells, & Serna, 2002). In the last two decades, the process through which U.S. high-school students are mapped onto courses has evolved from rigid curricular tracking to seemingly more flexible curricular choice. Despite changes in how students are mapped to coursework, the differentiated curriculum remains ubiquitous—resulting in substantial variation in students' academic experiences, within schools and across them (Angus & Mirel, 1999; Lucas, 1999; Oakes, 1985, 2005; Powell, et al., 1985).

Though originally viewed as a more democratic model of schooling, the differentiated curriculum has resulted in considerable social stratification in educational opportunities and outcomes (Lee & Ready, 2007; Yonezawa et al., 2002). A large volume of research in the 1990s documented strong links between students' academic and social backgrounds and their course taking (e.g., Lieberman, 1995; Newmann and Associates, 1996; Lee, 2002). In comprehensive high schools, students with strong academic skills and advantaged social backgrounds typically choose college-oriented course sequences, whereas students with weak academic skills and less advantaged or non-white backgrounds, often take low-level courses (Oakes, 1985, 2005; Lee, 2002). Currently, among educational researchers, there are virtually no advocates for the continuation of rigid tracking, although opinions differ about what is preferable. Most writings are critical of grouping students by ability (e.g., Argys, Rees, & Brewer, 1996; Gamoran & Mare, 1989; Lucas, 1999; Oakes, 2005). A much smaller group of writings suggest there may be advantages to homogeneous classes for organizational reasons, but that low-track students should receive more challenging coursework and better instruction than they receive under traditional tracking (e.g., Hallinan, 1994; Loveless, 1999). Although there is disagreement about the practice of differentiating coursework by student ability, both perspectives suggest low-ability students should be exposed to more rigorous coursework than has been typical.

The Call for Increased Rigor

Although researchers have voiced concerns about the social stratification that is inherent the differentiated curriculum, policy makers have targeted their criticisms of the high-school curriculum on its lack of rigor. Criticisms of low academic standards came to a head in the early 1980s with the *Nation at Risk* report, which described U.S. public secondary schools as "a sea of mediocrity" (National Commission on Excellence in Education, 1983). For the next two decades, changes in the U.S. and world economies have invited a crescendo of claims that too few students—especially those in urban schools—are graduating from high school with the skills needed for college and the workforce. More recently, policy makers have concluded that the skills needed for success in the workforce are the *same* as those needed to succeed in college (National Diploma Project, 2004). Thus, low-level coursework is increasingly viewed as insufficient to prepare any students for life after high school.

The criticisms raised in the 1980s about low academic standards, together with concerns about the current workforce, have motivated a national movement calling for rigorous high school course requirements. The National Governor's Association (2005a) recommended toughening high school graduation requirements to insist on college-preparatory coursework for everyone. Policy reports from ACT (2004) and the American Diploma Project (2004) have advocated increasing science and mathematics coursework and raising standards to improve alignment between secondary and post-secondary curricula. Policy makers have been following their recommendations.

At the state level, New York tightened its graduation requirements for all high school students beginning in 2001, followed by Texas in 2003—both states now mandate all students to complete a college-prep course sequence (Debray, 2005; Sipple, Killeen, & Monk, 2004). Thirteen states now require a college-prep curriculum, and 16 more states plan to adopt such requirements in the near future (Achieve, Inc., 2007). One large school district—Chicago—has been in the vanguard of this movement. In 1997 the Chicago Public Schools (CPS) mandated that all students enroll in a college-preparatory curriculum, eliminating the large array of remedial courses that were available. In this study we evaluate this Chicago policy.

Research Linking College-Prep Coursework to Student Outcomes

Several lines of research provide support for the reforms described above. A strand of research that gained prominence in the 1980s and 1990s was conceptualized within a school-effects framework, viewing curriculum structure as a measure of school academic organization. This group of studies used nationally representative samples of high schools and students and multilevel statistical methods to link high-school curriculum structure to student outcomes. The work began within a comparative Catholic/public framework, where researchers focused on differences between the constrained academic curriculum typical of Catholic high schools and the diffuse curriculum ubiquitous in public secondary schools (Bryk et al., 1993; Lee & Bryk, 1988, 1989). This research strand subsequently expanded beyond sector differences to explicitly tie curriculum structure to student outcomes (e.g., Lee, 2002; Lee, Burkam, Chow-Hoy, Smerden & Geverdt, 1998; Lee, Croninger, & Smith, 1997; Lee & Smith, 1995; Lee, Smith, & Croninger, 1997). These studies concluded that students attending schools offering a constrained academic curriculum—with few remedial courses and where most students take college-preparatory courses—benefited in two ways: (1) achievement gains were greater, and (2) learning was distributed more equitably by race, ethnicity and SES.

Support for policies requiring rigorous courses also comes from studies that directly link students' course-taking and their achievement. These studies document strong relationships between the courses students take in high school and their performance on academic tests and in college. For example, studies of curricular tracking consistently find that students in college-preparatory tracks have higher academic outcomes than those in general and vocational tracks (e.g., Lee & Bryk, 1988; Oakes, 1985, 2005). Several studies have shown that students who take advanced courses perform better on standardized tests than those without advanced coursework (Attewell & Domina, 2008; Chaney, Burgdorf & Atash, 1997; Gamoran & Hannigan, 2000; St. John, Musoba, Gross & Chung, 2004). Other studies have shown that students who complete a rigorous high school curriculum have better college outcomes than counterparts who complete less-demanding coursework (ACT, 2004; Adelman, 1999; Horn & Kojaku, 2001).

Distinct from work on the constrained curriculum (i.e. the *content* of courses and the *structure* of the curriculum) is work linking the *quantity* of required courses to student outcomes. Several studies have shown that simply mandating a minimum number of courses for graduation does not necessarily lead to better student outcomes (ACT, 2007; Clune & White, 1982; Hoffer; 1997; Teitelbaum, 2003). These findings are consistent with studies on curricular organization, which report that *which courses* students take is more important than the *number of courses* they complete (e.g., Lee, 2002; Lee, Burkam, et al., 1998; Lee, Croninger, & Smith, 1997; Lee & Smith, 1995; Lee, Smith, & Croninger, 1997). Replacing remedial coursework with college preparatory coursework is consistent with this research.

Recent Research on Tracking and De-Tracking

As policy organizations have called on states and districts to increase graduation requirements for all students, there has been virtually no discussion of how this could affect classroom organization within schools. Yet, by requiring all students to take the same courses, policies that universalize college preparatory coursework will lead many schools to group students more heterogeneously. Changes in ability grouping may have effects on student outcomes beyond the effects of changing curricular content. As noted earlier, many scholars have documented poor instructional environments in low-track classes (e.g., Oakes, 2005; Gamoran & Mare, 1989; Lucas, 1999; Powell, Farrar, & Cohen, 1985; Rosenbaum, 1976); this work suggests that low-ability students may learn more in mixed-ability classrooms. However, other work suggests that achievement is generally lower in heterogeneous classes, particularly for high-ability students (e.g., Argys, Rees, & Brewer,1996). In a study of newly detracked social studies classes, Rosenbaum (1999) reported that the most able students became bored and disaffected more than they had been in tracked classes. Loveless (1999) also points out that detracking may result in potential disadvantages for students in average- and high-tracks, loss of academically-talented students, and negative effects on low-ability students' self-esteem.

Prior studies suggest that successful detracking efforts may require fundamental changes in the organization of schools. Schools often face many difficulties when they attempt to eliminate tracking, including resistance from parents, technical difficulties of teaching heterogeneous classrooms, and a lack of instructional improvement due to teachers' low expectations for students (Wells & Oakes, 1996; Rubin, 2008). Although some schools have successfully detracked classrooms and improved instruction for low-ability students (Boaler & Staples, 2008; Oakes, 2005; Rubin, 2008), characteristics of such schools seem to be exceptional--with a shared belief in diversity among staff, successful professional development that led teachers to use inclusive pedagogical practices, and additional supports for struggling students (e.g., extra support courses). Thus, a policy of universalizing college-preparatory courses may have little chance of success if it does not address such issues as professional development around instruction, widespread support for the policy among the school community (teachers and parents), and extra support for low-ability students.

Shortcomings in Prior Research for Supporting the Current Policy

A large volume of research suggests that constraining the curriculum students follow to be college-focused will improve their academic outcomes. Yet, research on detracking has suggested that there may be cultural and structural limitations to universalizing a curriculum so that all students receive rigorous instruction. Furthermore, for a number of reasons, the existing research is limited in its applicability to the case of a *universal mandate* with which *all* schools are required to change their curricular offerings, and *all* students are required to take college preparatory classes. First, virtually all prior studies have suffered from some degree of selection bias. Second, prior research has paid little attention to differential effects by ability. Finally, the findings developed from data on national samples may not generalize to schools with chronic low performance and weak instructional capacity.

Selection bias. Most of the research supporting a college preparatory curriculum has compared student achievement between schools that already enrolled all students in college-preparatory courses to schools that did not (c.f., Lee, Burkam, et al. 1998; Lee, Croninger, & Smith, 1997; Lee, Smith, & Croninger, 1997; Lee & Smith, 1995). However, schools that had developed the capacity to enroll all students in college preparatory coursework may have been different than other schools in unmeasured ways—for example, they may have had a culture that was committed to diversity in education, or a mission to prepare all students for college. These unmeasured school differences could have affected students' outcomes, rather than the differences in coursework. Although prior studies have generally controlled for students' backgrounds and school composition, these adjustments cannot capture those structural and cultural impediments that limit some schools from successfully engaging all students in college preparatory coursework and which may also affect students' outcomes.

Likewise, studies based on comparisons of students in college preparatory tracks, or who took advanced coursework, compared to others students who did not, have been subject to selection bias at the student level. The observed benefits have been based on select students who chose college preparatory classes, or were counseled into them, who likely had unmeasured characteristics that affected their outcomes, such as high motivation or parental support. One recent study attempted to account for selection bias with propensity score matching--achievement and college enrollment outcomes were compared between matched samples of statistically similar students who did and did not follow a college-preparatory curriculum (Attewell & Domina, 2008). The authors found effect sizes that were substantially smaller or non-existent compared to those in previous studies, even using the same data and achievement outcomes. However, the variables available for that study were still not sufficient to eliminate the possibility of selection bias.[1] Furthermore, as the authors themselves state, their findings may not generalize to the case when a mandate or policy requires schools to offer college-prep coursework for all students.

Differential policy effects by ability. A universal policy assumes that all students can rise to the challenge of more demanding classes. Yet, schools typically offer remedial coursework for a reason--they believe that some students would struggle in college-prep classes. Very low-ability students could be particularly likely to become disengaged or fail when required to take challenging classes. They may even drop out before graduation—thus, negating any benefits from rigorous ninth grade coursework. Most studies of the high school curriculum have used a linear control for ability, rather than exploring differential effects. One prior study which did estimate differential effects of college-prep coursework found that students at or below the 20^{th} ability percentile benefited less than more able students from taking college-prep classes on a test of math achievement in grade (Gamoran & Hannigan, 2000). There may also be adverse effects on higher-ability students under the policy if teachers of college-preparatory courses modify their content and pacing to accommodate low-ability students who would otherwise have been in separate remedial classrooms. Evidence on dilution effects on curricula is mixed; the effects likely depend on the particular context and capacity of affected schools (see review in Teitelbaum, 2003).

Generalizability to urban schools. Furthermore, because much of the prior work on a universal curriculum is based on large national samples of schools, it may not be generalizable to schools in particularly challenging contexts. There may be substantial structural demands from curricular policies in schools with large numbers of low-achieving students, such as those in large urban districts like Chicago. For example, schools with large numbers of students in remedial tracks may lack sufficient qualified staff to teach a large expansion of college-preparatory courses. They may lack the resources to invest in professional development that would help teachers develop more inclusive pedagogy for incorporating many low-skill students in college-preparatory courses. In Chicago, prior to this policy, 19 percent of ninth graders failed their ninth grade English course, a quarter failed their math course, and students averaged over three weeks of course absence per semester. In schools with such high levels of failure and absenteeism, it may be particularly difficult to effectively increase instructional rigor in a way that promotes better academic outcomes for all students.

The policy mandate in Chicago provides an ideal opportunity to avoid the limitations of prior research. The fact that the Chicago reform applies to *all students* in all schools allows us to study the effects of requiring college-prep curriculum without selection bias. Large numbers of observations and detailed data on prior achievement allow us to estimate differential effects of a constrained curriculum on students entering high school with different levels of ability. Moreover, by estimating the effects of the policy on students and schools that would not ordinarily take/offer college preparatory courses, many of which struggle with very low achievement and weak capacity, we show the effects of mandating college preparatory coursework in a very challenging context.

A Conceptual Model for Studying Curriculum Effects

We used two important guides to structure our investigation of the effects of the new high school curriculum policy in Chicago: (1) the literature in which the topic is embedded (and the shortcomings of applying this literature to the Chicago context) and (2) a conceptual model that describes the mechanisms that may link the policy to the set of student outcomes we investigate. As conceptualized in our model, any policy effects on students' academic outcomes must flow through changes in the instructional program of the school: the courses, content, and pedagogy that students receive (see Figure 1). Whether a student begins a college preparatory sequence depends on both the structure of his or her school's instructional program (e.g., whether it requires college-prep or remedial courses) and the student's own background characteristics, which would influence his or her placement within the instructional program. The courses in which students enroll, combined with their response to instruction in those courses, shape academic learning in the ninth grade. In theory, mandating a constrained academic curriculum removes instructional variability so that all students have the same course experiences which should prepare them for advanced coursework in later grades. This should result in higher achievement both early and late in high school, particularly among students who would otherwise take less demanding courses. Ultimately, following such a curriculum should improve students' post-secondary outcomes.

*** Insert Figure 1 about here ***

Yet, as shown in Figure 1, the external policy mandate depends on a number of mediating factors if it is to affect outcomes proximally (at the end of Grade 9) and distally (at the end of high school and after graduation). Beneficial effects will occur only *if* schools' instructional programs adjust to the policy as expected, *if* students respond to changing instruction with better performance in the freshman year, and *if* improved freshman year performance leads to better academic outcomes in later years. However, some schools may have difficulty enacting substantial changes in their instructional programs in ways that benefit students, and students who would otherwise take remedial classes may not respond as expected.

Research Questions

In this study, we examine the consequences of universalizing a college preparatory curriculum on students' outcomes by comparing cohorts of students who attended the same Chicago high schools before and after policy implemention. We focus on two mandatory ninth grade courses: Algebra I and English I, as ninth grade coursework serves as a gatekeeper for more advanced study, and remedial ninth grade coursework was common pre-policy in both subjects. By basing this study in Chicago, we are specifically studying the effects of requiring college-prep coursework in an urban setting with a long history of chronic low performance. We limit our study to the consequences of the policy on students' academic outcomes, but recognize that the manner of implementation (e.g., changes in instructional demand and content in ninth grade classes) mediate the policy effects on student outcomes.

Research Question 1: Effects on Course-taking. To what extent did enrollment in ninth-grade college-preparatory courses increase as a result of the policy mandate, and how did the social distribution of course-taking by students' race, ability level, and disability status change between pre- and post-policy periods? With this question we discern the extent to which schools responded to the external mandate, and whether this resulted in a more equitable distribution of course-taking in these subjects based on students' background characteristics.

Research Question 2: Course Enrollment Effects on Student Outcomes. Did students' academic outcomes improve by taking college-prep instead of remedial classes, and did the effects differ by students' academic abilities as they began high school? This is a narrow question, showing the effect of taking one type of class versus another (college-prep vs. remedial), and applies only to students whose coursework was affected by the policy. As we hypothesize that students with weak academic skills may have the most difficulty adjusting to more demanding courses, we examine enrollment effects separately by students' incoming skills.

However, knowing the effect of taking a college prep class instead of a remedial class is not sufficient for evaluating the effects of the policy. The effects of the policy on any given student not only depend on how that student's achievement is different if she or he takes a college-prep class instead of a remedial class, but also on her or his likelihood of taking a college-preparatory class in the absence of the policy. For example, among average-ability students, taking Algebra I instead of remedial math might greatly affect their math grades (this is the enrollment effect), but because few students with average ability would have taken remedial math in the absence of the policy, the total policy effect on average-ability students would be small. Furthermore, the policy could have affected students' outcomes in ways other than changing their enrollment, such as by affecting climate and instruction in college preparatory classes. These effects would also accrue to students whose course enrollment was not affected by the policy. Therefore, our third research question discerns the broader policy effects.

Question 3: Overall Policy Effects on Student Outcomes. To what extent did the policy affect students' academic outcomes overall, and how did the effects differ for students entering high school with different abilities? The total policy effects incorporate the effects of college preparatory enrollment (discerned with Question 2) with students' likelihood of having their coursework affected by the policy (discerned with Question 1). The total policy effects also allow for unexpected consequences of the policy—such as changes in the content, rigor or composition of the college prep classes—that could influence the outcomes of all students. The policy effects in the analyses for Research Question 3 are, thus, more comprehensive than the enrollment effects from Question 2.

Method

The Chicago policy mandated college preparatory coursework for all students in all high schools beginning with students entering high school in 1997. In the ninth grade, students were required to take Algebra I and English I in the ninth grade (or a higher course in the math or English sequence, such as geometry, Algebra II, or English II). Remedial courses were eliminated in both subjects. We examine the effects of these changes in ninth-grade English and mathematics requirements, although the curriculum mandate was much more extensive.[2]

Sample and Data

Chicago has the third-largest school system in the United States. The student population is about 50 percent African-American, 38 percent Latino, 9 percent White, and 3 percent Asian. Approximately 85 percent of students are eligible for free/reduced priced lunches. In our statistical analyses, we include all CPS high schools in existence before and after the policy was implemented (n=59 schools). We use data on the entire population of students entering those high schools as first-time ninth graders over one decade: from the cohort entering in the fall of 1994 to the cohort entering in fall 2004. The cohorts range in size from 21,587 students in 1997 to 26,197 students in 2004.

We draw on a detailed longitudinal data archive containing complete administrative records for each student in each semester, semester-by-semester course transcripts, elementary and high school achievement test scores, data from the National Student Clearinghouse (NSC) on college enrollment, and from the 2000 U.S. Census. That these data are linked by student IDs allows us to analyze change over time in individual students' performance, and to control for changes in the types of students entering the high schools each year.

Measures

We constructed measures at the student, cohort, and school level to capture: (1) the effects of the policy on students' course-taking; (2) students' academic outcomes; (3) students' characteristics as they entered high school; and (4) control variables for cohort- and school-level characteristics that could otherwise influence our estimates of policy effects.

Measuring the policy. Our first step in measuring the policy was to determine which ninth graders enrolled in the college-prep courses (Algebra I, English I) in each cohort. We captured enrollment using information from grade and transcript files on course titles, levels (remedial, regular, and honors), and 6-digit course code designations. Students were coded as taking English/math college-preparatory courses if they took Algebra I/English I, or a course that was higher in the college preparatory sequence, such as Geometry or English II.[3]

To capture pre- and post-policy changes in course-taking and outcomes, we developed cohort-level dummy variables distinguishing four policy periods: (1) pre-policy cohorts [before 1997]; (2) the first year of the policy [1997]; (3) a mid-policy period [1998-2000]; and (4) a late policy period [after 2000]. We used these cohort indicators to compare enrollment and academic outcomes in post-policy years to pre-policy years. For simplicity, we only present the findings for mid-policy years in the tables.[4]

We measured change in college-preparatory enrollment at the school level in two ways. Our first indicator—used to address Research Question 2—captured the degree to which course enrollment changed for students with different incoming ability levels. For each high school we computed the proportion of students enrolled in college-prep courses pre-policy in each of four ability groups, and then computed the change in enrollment between pre- and post-policy periods (e.g., the percentage of very low ability students enrolled in college-prep classes in the school post-policy minus the percentage of very low ability students enrolled in college-prep classes in that school pre-policy, etc.). The second indicator—used to address Research Question 3—was a simple dummy-coded variable of whether or not the school was affected by the policy. We considered schools that enrolled at least 25% of their lowest-ability students in remedial coursework pre-policy as influenced by the policy (coded 1), whereas those that already enrolled 75% or more of their lowest-ability students pre-policy were coded 0, as they were largely unaffected by the policy—all (or almost all) of their students would have taken college preparatory courses in the absence of the policy. [5]

Student outcome measures. Reflecting the multiple outcomes shown in Figure 1, we considered 15 student-level dependent variables, measured both at the end of ninth grade and at the end of high school. The ninth grade outcomes included dummy-coded indicators for receiving credit in English/math *college-preparatory courses* (Algebra I or higher; English I or higher); failing a ninth-grade English/math course (regardless of level); continuous variables representing English/math course grades (on the traditional 4-point scale), number of English/math course absences, and English/math scores on the Tests of Academic Proficiency (TAP), given at the end of the ninth grade.[6] Long-term outcomes include final grade point average (GPA), dummy variables for high school graduation, earning credits in higher-level math classes (post-geometry and post-Algebra II), and enrollment in a four-year college within a year after high school.

While our focus is on the effects of courses in the ninth grade, we include long-term outcomes because ninth grade coursework often determines what students will take in subsequent years. By beginning college-preparatory sequences in ninth grade, students should have greater opportunities to take advanced coursework than if they waited until tenth grade to begin those sequences. For example, students who do not begin high school in algebra will not have enough years in high school to allow them to take pre-calculus. In addition, increased failure rates in ninth grade sets students up for a much higher risk of dropping out of school; this higher risk might not be seen until students reach an age at which they are likely to drop out.[7]

Student-level control variables. Research has shown that both course enrollments and student outcomes are associated with students' ability as they enter high school. To precisely measure this construct, we created composite measures of latent ability (one for math and one for reading) using a vector of students' annual testing history in the Iowa Tests of Basic Skills from third through eighth grade.[8] After standardizing the latent ability scores across all cohorts simultaneously, we created four dummy-coded ability categories: Group 1 (latent ability -0.5 SD below the mean or lower); Group 2 (-0.5 SD to the mean); Group 3 (from the mean to 0.5 SD above; and Group 4 (more than 0.5 SD above the mean).[9] We used these dummy-coded ability indicators to capture policy effects for students with different incoming abilities. We also created a set of continuous variables for English/math ability separately within each ability level, where students in other ability groups were coded zero.[10] Including these variables within each ability category allowed us to more precisely control for student ability and to adjust for potential shifts in the distribution of students in each ability group over time.

Our analytic models also included controls for age at high-school entry, gender, race/ethnicity, residential mobility prior to high school, special education eligibility, and English as Second Language (ESL) status, measured with dummy variables. In addition, we controlled for socio-economic status with two variables constructed from the U.S. census data on students' residential block groups (linked by students' home addresses: 1) concentration of poverty--a composite of male unemployment rate and the percentage families under the poverty line; 2) social status--a composite of the median family income and the average educational attainment.

Cohort-level control variables. We were concerned that changing compositions of students over time in a school could influence outcomes in ways that could be mistaken for policy effects.[11] For example, teachers may adjust instruction if the average ability levels of students in the school change over time. Therefore, we controlled for the average ability of students entering each school in each cohort.

School-level control variables. We began by considering a full set of variables for schools' structure and social composition, including measures of school size, the racial-ethnic and socioeconomic compositions, and schools' academic compositions—the proportions of students in special education, average incoming ability, ability heterogeneity, and whether or not the school was primarily a vocational or magnet school. Almost all of the school-level control variables were subsequently omitted from final models due to non-significance.

Analysis

Our analyses are presented in three parts. To address Research Question 1 (course-taking), we show changes in English I and Algebra I enrollments over time, including changes in enrollment by race and special education status. Using hierarchical models, with students nested within cohorts nested within schools, we estimated the policy effects on course enrollment adjusted for changes over time in students' background characteristics.[12] However, as the statistically adjusted results were similar to the descriptive results, we present the descriptive results for simplicity.

The analyses to address Research Questions 2 and 3 use an interrupted time series design with cohort comparisons to isolate policy effects on student outcomes. As discussed below, there was a clear shift in college-preparatory course enrollment post-policy, reaching nearly 100 percent for all schools by the mid-policy years. We take advantage of this shift to compare the outcomes of students in post-policy cohorts, almost all of whom enrolled in college preparatory courses, to previous cohorts in which many students took remedial courses.

One disadvantage of a cohort/interrupted time series design is that it could lead to false conclusions about the effects of the policy if there were other policy or programmatic changes in post-policy years which affected student outcomes. In fact, there were a number of policy changes in CPS over this period, including policies implemented in 1996 to require students to pass a standardized test to move on to ninth grade, and to hold schools accountable for students' test scores. A cohort approach by itself could confound the effects of these 1996 policies with the 1997 policy being evaluated here. Fortunately, the way that schools structured their course offerings pre-policy provided a natural comparison group of CPS schools that were not affected by the policy. Our analysis of ninth-grade course enrollment patterns showed considerable variability across all types of schools in pre-policy remedial course enrollment among students with the same ability levels.[13] Only schools that offered remedial courses pre-policy were affected by the mandate to end remedial coursework, while all schools would be affected by other CPS policies. Therefore, we were able to compare changes in students' outcomes in schools that were affected by this particular policy to changes in students' outcomes in schools not affected by this policy—the comparison schools serve as a control for other reforms occurring simultaneously. Combining cross-sectional and longitudinal comparisons allowed for much more confidence in the results of the analyses than either method would allow on its own.

Thus, there are two levels of comparison in our analyses, providing a difference-in-difference approach. First, we estimated how students in each school performed in post-policy cohorts compared to students with the same incoming ability in the same school pre-policy. We then compared these cohort differences in schools that were affected by the policy (because they initially offered remedial classes) to schools that were not affected by the policy (because they already enrolled all students in college-prep coursework).[14] Our analyses used 3-level hierarchical models, with students nested in cohorts nested in schools (see Appendix A). The analysis for Research Question 2 shows how comparable students' academic outcomes were different if they started high school in college-prep classes instead of remedial classes. The analysis for Research Question 3 shows the total effect of the policy on students' academic outcomes. The key difference in the analyses is in how we measure changes in course enrollment. To discern enrollment effects (RQ2), the key variable is the percentage change in enrollment in college preparatory classes (English, math, or both) for each ability group within each school, compared to pre-policy years. To discern policy effects (RQ3), the key variable is a simple dummy-coded indicator of whether the school was affected by the policy (i.e., whether it was a school that enrolled low-ability students in remedial courses in the absence of the policy).

Results

Research Question 1: Course Enrollment

Once the curriculum policy mandated college-preparatory courses and removed remedial course offerings, a large shift in ninth-grade course enrollment occurred. Figure 2 displays the proportion of ninth graders in each cohort enrolled in English I and Algebra I (vertical axes), based on students' ability levels upon entering high school (horizontal axes). From Figure 2, three trends are clear: (1) By 2000 virtually all CPS ninth graders were enrolled in both English I and Algebra I; (2) the policy most strongly influenced course enrollment among low-ability students but had almost no effect on course enrollment among students of high ability since such students had previously enrolled in college preparatory coursework before policy was enacted; and (3) implementation moved more rapidly in mathematics than in English.

*** Insert Figure 2 and Table 1 about here ***

These trends are reflected in the numerical results in Table 1. During the post-policy years, very close to 100% of students not eligible for special education services were enrolled in Algebra I, regardless of ability; 96 percent enrolled in English I. Gaps in course enrollments by race/ethnicity that existed prior to the policy largely closed post-policy. Although very low-ability students eligible for special education services were not in full compliance by the third policy year (2000), special education students' enrollment in college preparatory courses was much more strongly affected by the policy than regular education students, as they had the lowest college prep enrollment rates prior to the policy's implementation (40-42 percent).

Though the policy brought large shifts in course enrollment, it is possible that the observed changes were superficial: schools could simply rename remedial courses while students' experiences remained the same. Policies that mandate a specific curriculum assume that schools will respond by offering and enrolling students in the prescribed classes. However, we cannot assume that all schools will comply with such mandates—schools may be constrained by issues of capacity or culture. Although it is beyond the scope of this article to provide a full analysis of the instructional effects of the policy, we provide some evidence that the policy did have some substantive effects on the students' classroom experiences. As described in Appendix B, there was evidence of reduced tracking with the policy. On average, students with low incoming abilities were in classrooms with higher mean abilities post-policy than students with similar incoming skills pre-policy. In addition, fewer ninth grade math teachers reported spending little instructional time on Algebra. English teachers were less likely to report using textbooks, and more likely to assign students to read novels, poetry, non-fiction, and plays/scripts. Although we doubt that all algebra and English I classes had equally rigorous curriculum, the reform did seem to lead to some changes in the instructional experiences of low-ability students.

As pre-policy remedial course enrollment was strongly defined by students' academic abilities, we expected that schools serving mostly low-ability students would have been most likely to enroll their low-ability students in remedial coursework in the pre-policy years, whereas schools serving more high-achieving students would have been less likely to offer remedial coursework pre-policy. However, this was not the case. Once students' individual academic and social background characteristics were taken into account, only a few school characteristics were even slightly associated with the rates at which they enrolled students in remedial courses pre-policy.[15] (We do not include these tables here because of the preponderance of no-difference findings; they are available from the authors.) After taking many school-level characteristics into account, considerable variation in pre-policy college preparatory course enrollment remained between schools that otherwise served students of comparable ability. We found full enrollment in college-prep coursework in many schools in the pre-policy period that served predominantly low-ability students, whereas many schools serving generally high-ability students had substantial enrollment in remedial coursework among their low-ability ninth graders. We capitalized on this unexpected finding—that pre-policy remedial course enrollments were essential random across schools—by incorporating a second school-level contrast based on pre-policy college-prep course enrollments into our statistical models. This strengthened our analyses by providing a natural control group to incorporate into our time-series analyses.

Research Question 2: Course Enrollment Effects on Outcomes

The analyses addressing the second research question indicate whether students' outcomes changed as a result of taking college preparatory classes instead of remedial classes, and whether the effects differed by students' initial skills. Coefficients from the statistical models (given in Appendix A) are difficult to interpret; therefore, we show the results of the models in the form of a simulation.[16] Table 2 shows the changes in academic outcomes accompanying a 20 percentage point increase in college-prep enrollment (e.g., the effects of moving from 80% Algebra enrollment pre-policy to 100% Algebra enrollment post-policy). We group the 15 academic outcomes into three categories: (1) ninth-grade performance in mathematics, (2) ninth grade performance in English, and (3) outcomes measured at the end of high school. Because the original coefficients are not directly comparable, we converted the original units into two types of metrics: (a) school-level effect sizes for comparability in the left panel; and (b) meaningful units (e.g., percentage points, test score points) in the right panel. For simplicity, we show only the mid-policy period contrast (1998-99).[17] Results of the statistical tests are only indicated in the left-hand panel, although they apply equally to the right-hand panel.

The results shown in Table 2 hold constant the degree of enrollment change across the ability groups at 20 percentage points, allowing for a direct comparison of enrollment effects. Thus, we can determine whether very low-ability students were affected differently than average-ability students by enrolling in Algebra I or English I instead of remedial courses. In reality, a 20 percentage-point change is unrealistically high for average-ability students and is atypically low for the lowest-ability students in affected schools. Because so few students in the highest-ability group took remedial courses pre-policy, we do not report results for this group in Table 2, although they were included in the statistical models.

*** Insert Table 2 about here ***

Ninth-grade mathematics outcomes. Students in all ability groups were more likely to earn credit in Algebra I by the end of ninth grade with the policy. We would expect this with increased enrollment in Algebra, unless failure rates increased. However, beyond gaining course credit, there were no observable benefits to enrolling in Algebra I instead of remedial math. Moreover, there were some adverse consequences for both low- and average-ability students.

Across all ability groups, a 20 percentage point increase in Algebra I enrollment resulted in a 10% increase in students earning Algebra credit in ninth grade (from 8.0% to 11.6%, depending on ability level). This is consistent with an observed Algebra pass rate of about 50% among low-ability students. Math failure rates increased among low-ability students (3.0 percentage points, $p<.01$) and also among average-ability students (8.9 percentage points, $p<.05$). Students' math grades also decreased across ability groups by moderate amounts, declining the most among average-ability students (0.18 grade points; $p < .05$). Absenteeism increased in ninth-grade math among average-ability students (1.6 more days; $p<.05$). Math test scores were unaffected by taking Algebra instead of remedial math, although it is possible that the test was not sensitive to the change in curriculum. Only seven of 48 questions on the TAP exam test Algebra knowledge.

Ninth-grade English outcomes. Unlike outcomes in math, there were no adverse consequences from enrolling in English I instead of remedial English, although there were also few benefits. Students at all ability levels were much more likely to earn English I credit ($p<.01$), and their failure rates were unaffected. Course absence was very slightly less for the two lower ability groups by about a third of a day a year ($p<.05$). Neither English GPA nor reading test scores were affected. It is possible that the English test was insensitive to the curriculum; although it contained literary material that is more likely to be used in English I than in remedial English classes, such as nonfiction passages and poetry, about half of the questions were based on short and simple passages.

Longer-term outcomes. Although students were more likely to finish ninth grade with credits in Algebra I and English I, there were few effects on later outcomes. Students in the two lowest-ability groups were slightly more likely to earn upper-level math credits beyond geometry (a 20 percentage point increase in taking Algebra led to a three percentage point increase in lowest-ability students earning credits beyond geometry, $p<.01$), but not beyond Algebra II. Even though, in the post-policy period, students could potentially take the college-prep math sequence up to pre-Calculus (because they started the math sequence in ninth rather than tenth grade), students entering high school with low math abilities were not more likely to do so. Taking college-preparatory courses had no influence on graduating from high school, a finding we return to later. Echoing the lower grades in their ninth grade mathematics class, final high-school GPAs went down slightly ($p<.05$) for all but the lowest-ability students. Perhaps because their grades were lower, the probability of attending a 4-year college after high-school also decreased slightly for all but the lowest-ability students by less than one percentage point ($p<.05$).

Research Question 3: Overall Effects from the Policy

The results shown in Table 2 tell an incomplete story; they estimate course enrollment effects but do not take into account the likelihood that any student's course enrollment was affected by the policy. Nor do they capture unintended policy effects, including those that may influence students who would have taken college-prep classes even in the absence of the policy (e.g., changing the composition and content of college-prep classes). Thus, our final analyses compare those schools that offered remedial classes pre-policy to those that did not. We present the results in Table 3 in their natural metrics, parallel to the right-hand panel in Table 2.

** Insert Table 3 about here **

Ninth-grade math outcomes. The pattern of policy effects is similar to that seen in Table 2, however, it is clear from Table 3 that the lowest-ability students' academic outcomes were most strongly affected by the policy. This is reasonable—these students' course enrollments were most strongly affected by the policy. Although course enrollment effects on outcomes were stronger for average-ability students than low-ability students (Table 2), average-ability students were less likely to change their enrollment as a result of the policy because few were taking remedial courses pre-policy. Thus, overall policy effects on average-ability students' academic outcomes shown in Table 3 are small. Compared to the pre-policy years, post-policy students in the two lower-ability groups were more likely to earn credits for Algebra I or higher in ninth grade—increases of 8.8 and 7.4 percentage points for the lowest- and low-ability groups (p<.01). However, failures for lowest-ability students increased by 7.4 percentage points (p<.05) in the post- compared to the pre-policy period. Average-ability students were absent more often (3.14 days, p<.05), and the lowest-ability students' math GPAs declined by .15 points (p<.01). Achievement scores in mathematics at the end of ninth grade were unaffected.

Ninth-grade English outcomes. As a result of the policy, students in all but the high-ability group were considerably more likely to earn English I credit by the end of ninth grade than in the pre-policy period, particularly the lowest-ability students (increases of 12 to 36%, p<.01). This reflects the extensive use of remedial English in pre-policy years. The two low-ability groups were marginally more likely to fail their ninth grade English course (a 4.0 percentage point increase, but p>.05) and their English grades were slightly lower (by .11 to .15 points, p<.10). Absenteeism in English I classes increased somewhat post-policy, particularly among average- and high-ability students (by about 2 days, p<.05). Reading test scores also declined slightly for the highest achieving students, but were otherwise unaffected by the policy.

Longer-term outcomes. The policy effects on long-term outcomes were small to non-existent. A few of the pre/post-policy comparisons emerge as significant, but not in a consistent way across ability groups. The risk of Type 1 errors arising from multiple comparisons makes us less confident that these differences in long-term outcomes are real (they may be random noise). Students in the lowest ability group showed slightly higher GPAs over their high-school years ($p<.05$) and were slightly more likely to graduate; other ability groups showed no change in these outcomes. Students in the second-lowest ability group were 2.8 percentage points less likely to attend a 4-year college than they were in the pre-policy period ($p<.01$); other ability groups showed no change in college entrance. Thus, a policy that replaced remedial with college-prep coursework in the ninth grade for all students did not prepare more students for college—overall, students were no more likely to obtain credits in upper-level math courses, or attend a four-year college. Very low ability students were only slightly more likely to graduate.[18]

Discussion

Summary of Findings: Some Good but Mostly Bad News

The CPS policy of College-Prep for All was fully implemented within a few years, with almost all students enrolled in college-preparatory courses in English and mathematics in their first year of high school. Thus, the policy successfully eliminated the low-level coursework that had been common among low-ability students. The social distribution of course-taking also became more equitable; enrollment differences in college-prep coursetaking by race largely disappeared and enrollment gaps based on incoming ability and special education eligibility declined dramatically. Any program evaluation must consider whether schools responded to the policy—and our evidence suggests that they did. That is positive news.

The extant literature, which identified generally positive effects on students from taking college preparatory coursework, suggested that we would find larger policy benefits for lower-ability students, because their course enrollment was most affected. On one set of outcomes—

obtaining credit for ninth-grade Algebra I and English I—we found substantial benefits of the policy for such students. Over a third more of the students who began high school in the lowest-ability group earned credit in English I as a result of the policy, and almost one tenth more earned Algebra I credit by the end of their first high-school year. Moreover, their graduation rates did not decline when they began high school taking college preparatory courses instead of remedial courses—more good news.

Although it may not be reasonable to proclaim that no change in graduation rates is a positive finding, we suggest that this particular no-difference finding is important. Many educators have been reluctant to increase demands on high-school students, worrying that this could alienate them from school and encourage them to leave school before graduating. The policy evaluated in this study profoundly increased the demands made on low-ability students without the hypothesized concomitant effect of driving more of them out of high school. In the context of Chicago, where dropout rates approach 50 percent, there is great need to improve graduation rates. Requiring all students to take college-prep courses rather than accumulating credits through remedial coursework did not aggravate an already high dropout rate.

The remainder of the news is not good. On most of this wide array of outcomes—measured both at the end of ninth grade and at the end of high school—the policy had few positive effects and several negative effects. Ninth-grade mathematics grades declined and math failure rates increased among the students who took Algebra I instead of remedial math. Achievement scores in mathematics and reading were unaffected. Students were no more likely to obtain advanced math credits beyond Algebra II. Nor were they more likely to graduate from high school (except those in the lowest ability group) or to attend college. Absenteeism actually increased among average- and high-ability students. Thus, most of the benefits of the "College

Prep for All" policy suggested by the extant research were unrealized. Given the promise of such a policy, we asked ourselves why these results might be so disappointing. The remainder of our discussion posits some intertwined explanations.

Why Were Improvements in Academic Outcomes Unrealized as Expected?

Extant research was limited in its application to a universal mandate. Whereas the body of prior research on curriculum effects is extensive, the existing evidence could not conclude that college-preparatory courses for all students would benefit everyone. Earlier we mentioned the issue of selectivity bias. Most prior studies made use of nationally representative samples of students attending U.S. comprehensive high schools. In these samples, students who typically completed rigorous course sequences were those whose families had selected particular schools for them to attend (sometimes by careful choice of residence), those who were particularly motivated within their schools, and those who performed well in earlier grades. Schools that enrolled all students in college preparatory courses without a mandate to do so likely had greater capacity in terms of their structure, culture or instructional practices to support a universal curriculum. Because the policies in the current curriculum reform movement require that *all students* take demanding courses in *all schools*, it should not be surprising that the benefits are not as positive as previous studies would suggest.

Instructional content may be less important than quality. There is a large and growing body of research that links the quality of instruction to student learning. In elementary schools, there is agreement that all children should learn to read and compute but much less agreement about how these skills should be taught. For example, although everyone agrees that children should learn to read well, fierce debates rage between advocates of phonics and whole language.

In high schools, the debates more often concern which courses schools should offer and to whom they should offer those courses, rather than the way the courses should be taught. This policy mandate was narrowly focused on the curriculum, on the types of courses offered. Thus, the focus was more on *what was taught* in the courses rather than on *how it was taught*. Despite the change in course content, students may not have been any more engaged in learning math and English than they were prior to the policy. At the very least, it seems reasonable that teaching courses with high-level content to students without a record of high-level performance requires some substantial changes in the process of instruction. Yet, the policy was silent on this issue. Teachers who had taught remedial courses were suddenly required to teach college-prep courses. Even the teachers who had taught such courses prior to the policy were subsequently confronted with the problem of how to teach the same content to different types of students.

Classroom composition became more heterogeneous in college-preparatory courses. Prior to the implementation of this policy, ninth grade courses were highly tracked in many Chicago high schools. Although this study does not explicitly examine the effects of tracking, this policy did lead to more heterogeneous grouping of students. Research on successful detracking has emphasized the importance of instruction and of teacher buy-in for success in mixed-ability high-school mathematics classrooms (e.g., Boaler & Staples, 2008; Gamoran & Weinstein, 1998). This curriculum policy was not explicitly intended to detrack classes, and it was not accompanied by practices which have been suggested as helpful for schools that are detracking. In later years, CPS did initiate some of these practices to try to support struggling students, including double-period classes in ninth grade English and math, and we show the results of these efforts in other work.

One particular problem noticed in case studies of detracking is the challenge of teaching heterogeneous classes; teachers often target instruction to the hypothetical middle student (Bar and Dreeben, 1983; Rosenbaum, 1999). We suspect that problems typical of detracking were evident in Chicago's English I and Algebra I classes. Such problems may impede learning and sometimes engender negative behaviors, such as misbehaving in class. This could explain the higher absenteeism rates in Algebra I and English I, and the decline in English test scores, for more able students (see Table 3).

The policy also led students who would have taken remedial classes to take math and English with higher-ability peers than they would have in the absence of the policy. This change in students' relative abilities, compared to their classroom peers, likely increased their probability of failure. The possibility of negative effects on students' self-esteem was noted by Loveless (1999), and this study provides some suggestive evidence that is consistent with this hypothesis. Few studies on tracking have examined the effects on grades, instead of test scores. Further work that our research team is pursuing shows that students with the same ability levels have lower grades when in classrooms with higher-ability peers. This might explain the larger course enrollment effects on average-ability students—they would have been the highest-ability students in remedial classes but would be lower-ability students when mixed into college-prep classes (see Table 2).

Some students had difficulty handling high-level content, but not just because of weak academic skills. It is possible to draw a conclusion from our results that not all students are well served by college-prep courses. Reflecting earlier views on how high schools should be organized—and the ubiquitous nature of the comprehensive high school—it could be argued that these courses are beyond the capability of some students, or that intellectually demanding courses are not helpful to students if they don't want them, don't want to work hard to master the content, or don't see them as useful for their futures. Such perspectives seemed to lie behind resistance to detracking in some schools (Oakes & Wells, 1996; Rubin, 2008). However, at a time when the vast majority of U.S. students want a college degree, and when the workforce demands higher skills than in previous decades, all students need the opportunity to learn high-level skills before leaving high school. Furthermore, despite its apparent logic, our findings do not support this argument.

The short-term findings suggest that lower-ability students are more likely to struggle in college-prep classes, but the long-term findings do not show adverse effects (even if they also show few benefits). Thus, the policy generally had a null effect—no gains, but no costs. This does not suggest that low-skill students cannot handle college-prep coursework, but rather that they do not benefit from it any more than they do from remedial coursework. Further, average-ability students showed more adverse consequences from taking college-prep courses instead of remedial courses than did low-skill students, suggesting that the higher failure rates were not entirely explained by students' insufficient academic skills. Instead, it suggests that more attention be paid to how students are taught and the quality and depth of the tasks in which students are engaged, rather than the content of what they are taught (see a case study that further suggests this in Boaler & Staples, 2008). It also shows that poor academic performance does not just arise solely from weak academic skills.

There is a need to attend to non-cognitive skills and behavior in high school and earlier grades. It seems unreasonable to mandate a policy that makes greater cognitive demands on students and not suggest changes in students' preparation before they leave their K-8 schools. In this respect, the timing of this policy corresponded well with improvements in student learning that were occurring in Chicago's elementary schools. Over this time period, achievement by the eighth grade was improving and more students were entering high school with the academic skills that would be expected for success in college-prep coursework. However, weak academic skills are not the primary source of course failure in Chicago schools--students' academic behaviors (attendance and completing homework) are eight times more predictive of failure than their test scores (Allensworth & Easton, 2007). For this policy to succeed, these behaviors must be instilled in earlier grades and further developed when students get to high school.

High course absence rates and low levels of engagement are reflected in both the pre- and post-policy period in Chicago high-school students' grades. Pre-policy, the average ninth-grade math or English grade in CPS was below a C (about a 1.8 average in each subject on the typical 4-point scale). For students with very low abilities, the average grade was a D+ (about 1.5). These disturbingly low grades did not improve post-policy. If students are earning Ds in their courses, can we really expect the content that they are barely learning to matter? As long as students continue to be minimally engaged in their courses and to attend school irregularly, we should not expect substantial improvements in learning. Getting the content and structure of courses right is just the first step. Real improvements in learning require strategies that get students excited about learning, attending school and classes regularly, and working hard in whatever course they are taking.

Final Comments

What we offer here is, in effect, a large array of "no difference" findings for an important policy initiative about the high-school curriculum. Given the data we have used to estimate our findings, it would seem that the most common explanation for no-difference findings—inadequate statistical power—is not applicable here. Because we have employed whole cohorts of high school students, sample sizes are quite large (at least in terms of numbers of students). Our analytic methods are rigorous and robust to alternative analytic strategies. We have considerable confidence that our findings are, in fact, real.

Our evaluation suggests that the policy of universalizing high-level coursework for all students does not live up to the high hopes held for it, at least in the case of an urban district with chronic low-performance where improvements in student performance are most desperately needed. Despite disappointing results, we are unwilling to say at this point that this is not a good policy, or to conclude that high-level intellectual content is not appropriate for all students. In fact, we laud CPS for having a strong belief that all their students can succeed in a curriculum that demands a lot from everyone. Although this study does not provide support for a constrained curriculum, it also does not support the argument for social efficiency. The differentiated curriculum was not serving students well—even when they took remedial coursework large numbers of students failed those courses and eventually dropped out. What we are willing to say is that curricular policies need to be accompanied by other profound changes in the educational system—with greater attention to instruction, and with concomitant efforts to improve the academic behaviors that have been shown to be associated with better school performance. High schools need to implement deeper changes if all students are to be successful in engaging in high-level intellectual content and the increased learning.

In future work we will further investigate the mechanisms that may explain why this policy did so little to improve students' outcomes. We will examine the effects of detracking (mixed ability grouping) itself on high- and low-ability students, separate from the effects of universalizing the curriculum. We will also examine the ways in which schools staffed the expanded course offerings in the college preparatory sequence. Finally, we will examine whether the policy had differential effects based on schools' organizational structure and capacity. Schools with more supports for students, for example, may have more successfully helped low-ability students master more rigorous courses. Schools with a strong professional community or instructional leadership by the principal might have better responded to the mandate. The problem may reside in how schools implemented the new curriculum, and future work will test this.

References

Achieve, Inc. (2007). *Closing the expectations gap 2007: An annual 50-state progress report on the alignment of high school policies with the demands of college and work*. Retrieved February 15, 2008 from http://www.achieve.org/node/844.

ACT, Inc. (2004). *Crisis at the core: Preparing all students for college and work.* Iowa City, IA: ACT Research Report Series.

ACT, Inc. (2007) *Benefits of a High School Core Curriculum for Students in Urban High Schools.* Retrieved February 15, 2008 from www.act.org/research/policymakers/pdf/core_curriculum.pdf.

Adelman, C. (1995). *The New College Course Map and Transcript Files: Changes in Course-Taking and Achievement, 1972-1993* (No. PE-95-8001). Washington, DC: National Institute on Postsecondary Education, Libraries and Lifelong Learning.

Allensworth, Elaine and John Q. Easton. 2007. *What Matters for Staying On-Track and Graduating in Chicago Public High Schools: A Close Look at Course Grades, Failures and Attendance in the Freshman Year*. Consortium on Chicago School Research, Chicago, Illinois.

American Diploma Project. (2004). *Ready or not: Creating a high school diploma that counts.* Washington, DC: Achieve, Inc.

Angus, D.& Mirel. J. (1999). *The failed promise of the American high school, 1890-1995.* New York: Teachers college Press.

Argys, L.M., Rees, D.I., & Brewer, D.J. (1996). Detracking America's schools: Equity at zero cost? *Journal of Policy Analysis and Management*, *15*(4), 623-645.

Attewell, P. & Domina, T. (2008). Raising the bar: Curricular intensity and academic performance. *Educational Evaluation and Policy Analysis, (30)*1, 51-71.

Boaler, J. & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record 110*(3), 608-645.

Boaler, J. (1997). Setting, social class, and survival of the quickest. *British Educational Research Journal, 23*(5), 575-595.

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

Chaney, B., Burgdorf, K., & Atash, N. (1997). Influencing Achievement Through High School Graduation Requirements. *Educational Evaluation and Policy Analysis, 19*(3), 229-244.

Clune, W. H., & White, P. A. (1992). Education reform in the trenches: increased academic course taking in high schools with lower achieving students in states with higher graduation requirements. *Educational Evaluation and Policy Analysis, 14*(1), 2-20.

Cremin, L.A. (1961). *The transformation of the school: Progressivism in American education.* New York: Knopf.

Debray, E. (2005). A comprehensive high school and a shift in New York State policy: A Study of early implementation. *The High School Journal, 89*(1), 18-45.

Gamoran, A., & Hannigan, E. C. (2000). Algebra for Everyone? Benefits of College-Preparatory Mathematics for Students with Diverse Abilities in Early Secondary School. *Educational Evaluation and Policy Analysis, 22*(3), 241-254.

Gamoran, A., and Mare, R. D. (1989). Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality? *American Journal of Sociology*, *94*, 1146-1183.

Gamoran, A., and M. Weinstein. (1998). Differentiation and Opportunity in Restructured Schools. American Journal of Education, 106 (May): 385-415.

Hallinan, M. (1994). Tracking: From theory to practice. *Sociology of Education, 67*(2), 79-84.

Hoffer, T. B. (1997). High school graduation requirements: effects of dropping out and student achievement. *Teachers College Record, 98*, 584-607.

Horn, L., & Kojaku, L. K. (2001). High School Academic Curriculum and the Persistence Path through College: Persistence and Transfer Behavior of Undergraduates 3 Years after Entering 4-Year Institutions. *Education Statistics Quarterly, 3*(3), 65-72.

Kliebard, H.M. (1995). *The struggle for the American curriculum 1893-1958.* New York: Routledge.

Lee ,V.E.& Ready, D.D. (2007). *Schools within schools: Possibilities and pitfalls of high school reform.* New York: Teachers College Press.

Lee, V. E., & Bryk, A. S. (1988). Curriculum tracking as mediating the social distribution of high school achievement. *Sociology of Education, 61*, 78-94.

Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education, 62*, 172-192.

Lee, V. E., & Smith, J. B. (1995). Effects of high school restructuring and size on early gains in achievement and engagement. *Sociology of Education, 68*, 241-270.

Lee, V. E., Burkam, D. T., Chow-Hoy, T., Smerdon, B. A., & Geverdt, D. (1998). *High school curriculum structure: Effects of course-taking and achievement in mathematics for high school graduates. An examination of data from the National Education Longitudinal Study of 1988.* (Working Paper No. 98-09) Washington, DC: U.S. Department of Education Office of Educational Research and Development, National Center for Educational Statistics.

Lee, V. E., Croninger, R. G., & 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., & Croninger, R. G. (1997). How high school organization influences the equitable distribution of learning in mathematics and science. *Sociology of Education, 70*, 128-150.

Lee, V.E. (2002). *Restructuring high schools for equity and excellence: What works.* New York: Teachers College Press.

Lieberman, A. E. (1995). *The work of restructuring schools: Building from the ground up*. New York: Teachers College Press.

Loveless, T. (1999). *The tracking wars: State reform meets school policy.* Washington, DC: Brooking Institution.

Lucas, S.R. (1999). *Tracking inequality: Stratification and mobility in American high schools.* New York: Teachers College Press.

Miller, Shazia Rafiullah, and Elaine M. Allensworth. 2002. "Progress and Problems: Student Performance in CPS High Schools, 1993 to 2000." Chapter 2 (pp. 51-87) in *Reforming Chicago's High Schools: Research Perspectives on School and System Level Change*, Valerie Lee (ed.), Consortium on Chicago School Research, Chicago.

National Center for Education Statistics. (2005). *The nation's report card.* Retrieved January 15, 2008 from http://nces.ed.gov/nationsreportcard.

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

National Diploma Project (2004). *Ready or not: Creating a high school diploma that works.* Achieve, Inc.: Washington, DC.

National Governors Association Center for Best Practices. (2005a) *Getting it done: Ten steps to a state action agenda*. Retrieved November 08, 2005 from http://www.nga.org/Files/pdf/05warnerguide.pdf.

Oakes, J. (1985). *Keeping track: How schools structure inequality.* New Haven: Yale University Press.

Oakes, J. (1994). More than misapplied technology: A normative and political response to Hallinan on tracking. *Sociology of Education, 67*(2), 84-91.

Oakes, J. (2005). *Keeping track: How schools structure inequality* (2nd ed.). New Haven: Yale University Press.

Powell, A.G., Farrar, E., & Cohen, D.K. (1985). *The shopping mall high school: Winners and losers in the educational marketplace.* Boston: Houghton Mifflin.

Rosenbaum, J. E. (1976). *Making inequality: The hidden curriculum of high school tracking*. New York: Wiley.

Rosenbaum, J.E. (1999). If tracking is bad, is detracking better? *American Educator, 23*(4), 24-29. 47.

Rubin, B.C. (2008). Detracking in context: How local constructions of ability complicate equity. *Teachers College Record, 110*(3), 646-699.

Sipple, J. W., Killeen, K., & Monk, D. H. (2004). Adoption and adaptation: school district responses to state imposed learning and graduation requirements. *Education Evaluation and Policy Analysis, 26*(2), 143-168.

St. John, E. P., Musoba, G. D., Gross, J. P. K., & Chung, D. G. (2004). *Academic preparation and college success: Analyses of Indiana's 2000 high school class* (No. 04-03). Indianapolis: Lumina Foundation & Indiana Commission for Higher Education.

Tietelbaum, Peter. (2003). The Influence of High School Graduation Requirement Policies in Mathematics and Science on Student Course-Taking Patterns and Achievement. *Education Evaulation and Policy Analysis*, 25(1): 31-57.

Wells, A. S., & Oakes, J. (1996). Potential pitfalls of systemic reform: Early lessons from research on detracking. *Sociology of Education, 69*(Extra Issue: Special Issue on Sociology and Educational Policy: Bringing Scholarship and Practice Together), 135-143.

Yonezawa, S., Wells, A.S., & Serna, I. (2002). Choosing tracks: "Freedom of choice" in detracking schools. *American Educational Research Journal, 39*(1), 37-67.

Endnotes

[1] Propensity score analysis can only claim to eliminate selection bias when there are such rich data available for consideration that it is inconceivable that factors could remain to account for both outcomes and selection. Although they had students' grades available for development of propensity scores, Attewell and Domina could not control for many factors which could conceivably affect outcome and coursework (e.g., parental involvement, motivation, teacher support).

[2] The policy specified four years of specific English courses (survey literature, American literature, European literature, world literature), three years of specific math courses (algebra, geometry, advanced algebra), three years of science (biology, earth /space or environmental science, chemistry or physics), and three years of social science (world studies, U.S. history, elective). However, classifying students by whether they completed a college-preparatory sequence introduces problems of selection and attrition. We focus on ninth grade coursetaking to avoid these issues.

[3] To classify courses as remedial, college-prep, or elective, we examined several fields on student transcript records--the course level, the course code, the course name, as well as the curriculum designation associated with each course code in central office curriculum records. The curriculum designation provides information on whether a particular course code satisfies a district graduation requirement, and which requirement is fulfilled by that course. If the course level in the transcript file indicated a course was "Basic" or "Essential", or if the curriculum designation for the course identified it as an elementary level course, we coded the course as remedial. If the course title or curriculum designation indicated that the course was a support course (e.g., 'reading workshop', 'math workshop', 'math support'), and the student was not simultaneously taking a college-prep course in that subject (English or math), we coded it as a remedial course. To mark courses as college-prep, we used similar methods of looking at the course level, the curriculum designation, and the course title to determine if a course was English I-IV, or Algebra, Algebra II/Trig, Geometry, or an advanced math course. We marked all courses as electives that we had not marked as remedial or college-prep (e.g., journalism).

[4] In preliminary analyses, we compared all post-policy years to all pre-policy years. However, when we presented these findings, we heard concerns that the results may have been affected by implementation issues in the first year of the policy, or by another phase of curricular reform that occurred several years after the 1997 policy (after 2000). By separating out post-policy periods, we could see that the effects were similar in all three periods. We present the mid-policy period results because this is the period that should not have been affected by these other issues. Tables with results from all three post-policy periods are available from the authors.

[5] We only use lowest-ability (group 1) students for this definition because these students would be enrolled in remedial classes if they were available at the school. Including students of higher ability levels would confound our definition, as it would not only depend on whether schools offered remedial classes, but also on what proportion of students in the school were low-ability.

[6] Results were similar if the positive skew of the absence variable was reduced through a log transformation, so the untransformed variable is used for ease of interpretation.

[7] There is a strong and consistent relationship between freshman year course failures and whether a student eventually graduates, however, few students drop out before age 16 (Allensworth and Easton, 2007).

[8] A two-level HLM, nesting years within students, modeled each student's learning trajectory; level 1 included variables for grade and grade-squared which were allowed to vary across students. There was also a dummy variable representing a repeated year in the same grade, to adjust for learning that occurred the second time in a grade, and a different dummy variable for repeating the eighth grade year so that additional learning that occurred when eighth grade was repeated could be added into a student's latent score. Before running the HLM, students' test scores were equated through Rasch analysis to remove form and level effects.

[9] This categorization was guided by preliminary analyses. We found that students with ability scores above .5 had nearly 100 percent enrollment rates in college-reparatory courses pre-policy, and these students were categorized together. All the other students were grouped by .5 standard deviation intervals.

[10] We initially used just one continuous variable which was centered on the mean for each ability group to avoid collinearity between continuous and dummy ability variables. However, we found that using separate continuous ability variables allowed for more precision in controlling for achievement because the slope between latent ability and enrollment/outcomes could vary by ability group.

[11] This was a particular concern because a number of schools showed substantial changes in the types of students they served over this period. Therefore, we not only control for students' individual incoming ability and the average abilities of students in their cohort in their school, but we also make comparisons across schools based on pre-policy remedial course enrollment so as to control for systemwide changes that were unrelated to the policy to end remedial coursework.

[12] These analyses are available from the authors. They show the actual policy effect net of changes in the background characteristics of students entering the schools.

[13] No measured school characteristics were related to the degree to which schools enrolled students in remedial coursework, including type of school (magnet, vocational, neighborhood), size, average incoming ability level, or demographic composition.

[14] The degree of college-preparatory enrollment is calculated separately for enrollment in English and math to predict the corresponding English and math outcomes. For non-subject specific outcomes we combine English and math course enrollment, with full college preparatory enrollment counted as being enrolled in both subjects.

[15] For example, average incoming achievement alone, racial composition and school demographic characteristics, were not related to pre-policy course enrollment. Only schools with both higher average ability and more heterogeneous ability distributions were slightly less likely to enroll low achieving students in college prep math courses than other schools. These small differences were driven by just a few schools.

[16] Coefficients from the full models are available from the authors.

[17] We looked for different policy effects in each period out of concern that the policy may have not been fully implemented in the first year, and that the later cohorts would be affected by policies implemented in later years. However, our interpretations of policy effects are similar across all three post-policy periods.

[18] It may seem counter-intuitive that graduation rates would increase when freshman failure rates rise. However, prior to the policy many low-ability students did not take enough credits in their freshman year to potentially graduate in four years. By requiring students to take more classes in their freshman year, including Algebra, students were more likely to obtain enough credits to graduate in four years (see Miller & Allensworth, 2002).

Figures

Figure 1. Conceptual model of How Curricular Policy Influences Student Outcomes

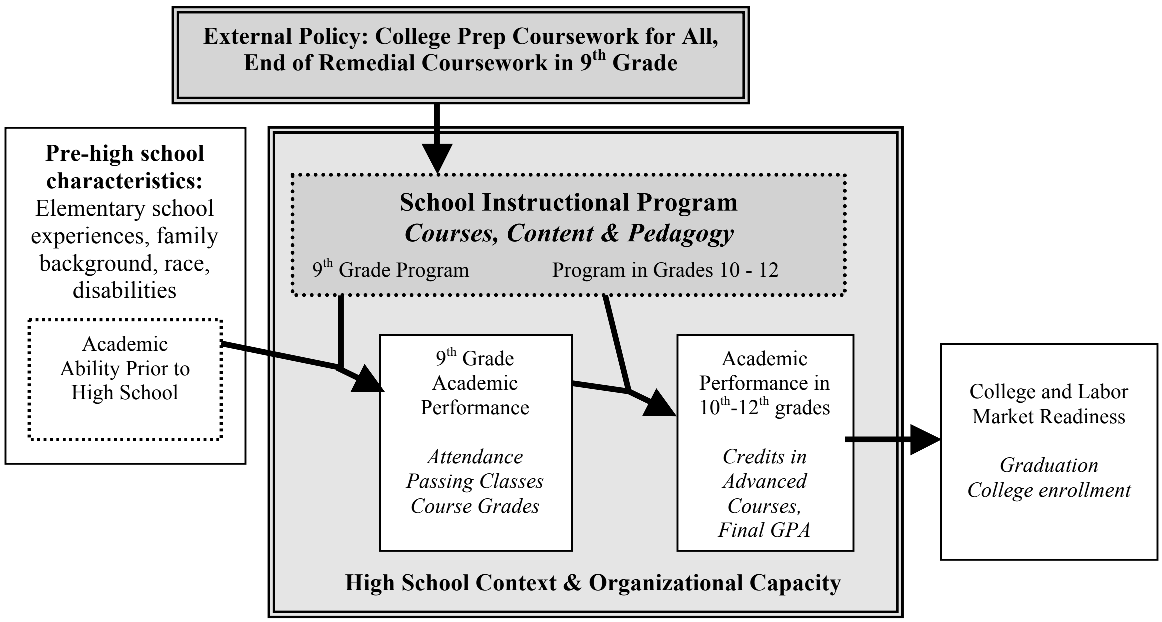
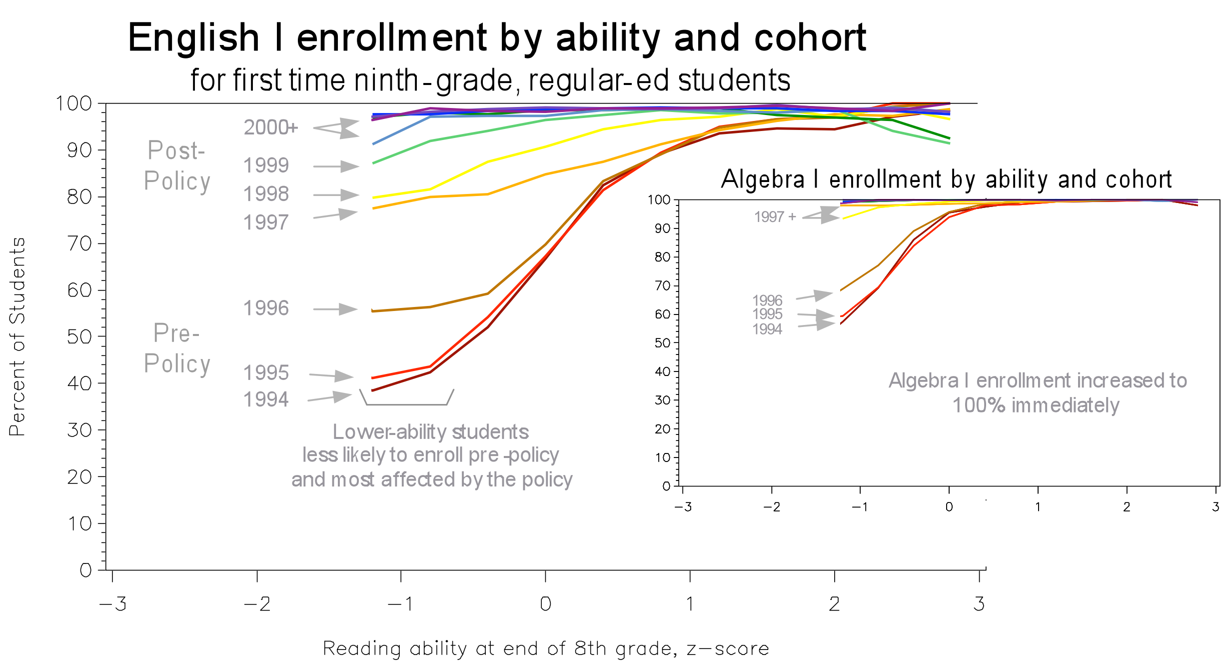


Figure 2. Enrollment in College-Prep English and Math Courses by Freshman Cohort



Note: English I enrollment includes the small percentage of ninth graders who took other collegeprep English (i.e., English II-IV). Likewise, Algebra I enrollment includes ninth graders who took college-prep math beyond Algebra I (e.g., Geometry, Algebra II, etc)

Table 1. Ninth Grade College Prep Math and English Enrollment Pre- and Post-Policy

	All Students				Regular Students				Special Education Students			
	Math		English		Math		English		Math		English	
Policy Period	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Overall	81%	97%	64%	95%	85%	99%	66%	96%	48%	89%	42%	93%
Ability Group												
Lowest	64%	91%	45%	93%	71%	97%	46%	93%	42%	86%	40%	93%
Low	88	97	59	94	89	98	60	94	71	95	47	94
Average	96	99	77	96	96	99	78	96	89	97	59	96
High	98	99	92	98	99	99	93	98	96	98	81	96
Race/Ethnicity												
African American	79%	96%	64%	96%	84%	98%	67%	97%	48%	87%	45%	94%
Latino	82	98	62	94	86	99	64	95	49	92	41	92
White	88	99	64	95	90	99	70	96	53	96	31	93
Asian	83	98	64	96	91	99	66	96	47	93	43	93

The post-policy statistics in the table are averaged across all of the post-policy years. Compliance was close to 100% by 2000 for all but lowest-ability students eligible for special education services.

Table 2. Effects of Ninth Grade College-Prep Math and English Enrollment on Academic Outcomes

	Effect of 20 Percentage Point Increase in Enrollment (Effect Sizes)					Effect of 20 Percentage Point Increase in Enrollment (Percentage Points, Days, Grade Points)				
Ninth Grade Math	Algebra or Higher Credit	Math Course Failure	Math Course Absences	Math GPA	Math Test Scores	Algebra or Higher Credit	Math Course Failure	Math Course Absences	Math GPA	Math Test Scores
Lowest Ability	1.11**	0.31**	0.00	-0.25**	0.06	8.9%	3.0%	-	-0.06	-
Low Ability	1.22**	0.41*	0.08~	-0.34*	-0.02	11.6%	3.5%	0.70	-0.08	-
Average Ability	0.96*	1.24*	0.21*	-0.76*	-0.07	8.0%	8.9%	1.60	-0.18	-
Ninth Grade English	English I Credit	English Course Failure	English Course Absences	English GPA	Reading Test	English I Credit	English Course Failure	English Course Absences	English GPA	Reading Test
Lowest Ability	1.65**	0.07	-0.05*	-0.09	0.01	14.7%	-	-0.33	-	-
Low Ability	1.56**	0.07	-0.05*	-0.09~	0.00	12.8%	-	-0.32	-0.02	-
Average Ability	1.60**	0.14	-0.05	-0.09	-0.03	12.4%	-	-	-	-
Long-Term	Post-Geometry Credit	Advanced Math Credit (beyond Algebra II)	Graduate	Final GPA	Attend 4-Year College	Post-Geometry Credit	Advanced Math Credit (Non-Algebra II)	Graduate	Final GPA	Attend 4-Year College
Lowest Ability	0.40**	0.07	-0.14	0.07	0.03	3.3%	-	-	-	-
Low Ability	0.35~	0.10	-0.05	-0.07~	-0.09~	3.5%	-	-	-0.02	-0.4%
Average Ability	0.05	-0.14	0.05	-0.14*	-0.13*	-	-	-	-0.04	-0.7%

~p<.10 *p<.05 **p<.01: This table shows the effects for period 2 (1998-1999) only; similar effects were observed in other post-policy periods. We calculated values by taking the difference between the pre/post policy outcomes by the degree to which enrollment changed in the school, compared to schools with little enrollment change. The values were converted into their natural metric. Effect sizes were calculated by multiplying the mid post-policy percent change coefficient by 2 (for a 20% change) and dividing that value by the school level standard deviation in the respective outcome from the fully unconditional models. If the coefficient was not statistically significant at p<.10 no value is displayed in the right side of the table. All coefficients from the full model are available from the authors.

Table 3. Total Policy Effects on Academic Outcomes

Policy Effects on Student Outcomes by Ability Level					
Ninth Grade Math	Algebra or Higher Credit	Math Course Failure	Math Course Absences[i]	Math GPA[ii]	Math Test Scores[iii]
Lowest Ability	8.8% **	7.7% *	2.24	-0.15 **	0.63
Low Ability	7.4% **	3.6%	2.40	-0.07	0.13
Average Ability	1.0%	1.3%	3.14 *	-0.01	-0.17
High Ability	-0.4%	1.2%	1.67	0.10	0.50
Ninth Grade English	English 1 Credit	English Course Failure	English Course Absences[i]	English GPA[ii]	English Test Scores[iii]
Lowest Ability	35.8% **	4.1%	1.15	-0.15 *	-0.61
Low Ability	28.2% **	4.0% ~	1.46 ~	-0.11 ~	-0.64
Average Ability	11.7% **	2.1%	2.11 *	-0.06	-0.54
High Ability	-3.2%	2.6%	1.70 *	-0.07	-1.11 ~
Long-Term	Post-Geometry Credit	Advanced Math (Non-Algebra II) Credit	Graduate	Final GPA[ii]	Attend 4-Year College
Lowest Ability	1.0%	-0.3%	4.6% ~	0.10 *	0.6%
Low Ability	0.6%	-2.0%	-1.7%	-0.05	-2.8% **
Average Ability	-0.3%	-2.5%	0.2%	-0.04	-1.1%
High Ability	-0.9%	-1.9%	-0.7%	0.00	-2.0%

~p<.10 *p<.05 **p<.01: This table shows the effects for period 2 (1998-1999) only. Values were calculated by taking the difference between the pre/post policy change for schools that changed enrollment and those that did not change enrollment. The values were converted into their natural metric.

[i]Days [ii]Grade points [iii]Normal curve equivalents

Appendix A. Statistical Models

Both sets of analyses of enrollment/policy effects on academic outcomes use three-level hierarchical models, with students nested within cohorts within schools. The student-level model to estimate the outcome Y for student *i* in cohort *j* in school *k* is written as:

$$Y_{ijk} = \pi_{1jk}(\text{ability level1})_{ijk} + \pi_{2jk}(\text{ability level2})_{ijk} + \pi_{3jk}(\text{ability level3})_{ijk} + \pi_{4jk}(\text{ability level4})_{ijk} + \sum_{p=1}^{P} \pi_{4+pjk}(X)_{ijk} + e_{ijk}$$

where X is a vector of student-level control variables (incoming ability, race, mobility, age, etc.).

This model does not include an intercept; cohort effects are estimated at each ability level independently. The first four coefficients (π_{1jk}, π_{2jk}, π_{3jk}, and π_{4jk}) provide the mean outcome (e.g., test score, course failure, college enrollment) for students in each cohort in each school at each ability level, controlling for individual background characteristics ($X_{1\ldots p}$). At the cohort-level, we specify these means as a function of cohort year, controlling for the academic composition of students in that school in that cohort. For each ability level *m*:

$$\pi_{jk} = \beta_{m0k} + \beta_{m1k}(\text{early post-policy})_{jk} + \beta_{m2k}(\text{mid post-policy})_{jk} + \beta_{m3k}(\text{late post policy})_{jk} + \beta_{m3k}(\text{cohort average latent ability})_{jk} + r_{mj,}$$

The intercept β_{m0k} represents the average *pre-policy* outcome at ability level *m* in school *k*, and the coefficients $\beta_{m1k,}$ $\beta_{m2k,}$ and β_{m3k} represent the <u>change</u> in the average outcome for students in ability group *m* at each school from pre-policy to the respective post-policy period. If there was no policy effect, these coefficients should be equal to zero. If the policy had an effect on the outcomes of students in ability group *m*, these coefficients should be different than zero.

At the school-level, we estimated the average pre-policy outcomes (β_{m0k}) and the average post-policy change in outcomes ($\beta_{m1k,}$ $\beta_{m2k,}$ and β_{m3k}) as a function of school characteristics. Schools in which few students in a given ability group enrolled in college prep courses pre-policy should have shown more change in outcomes for that ability group than schools where almost all students in the group enrolled in college prep courses before the policy. (If all of the students of that ability already enrolled in college prep courses pre-policy the policy should have had no effect.) In the analysis of enrollment effects, we included a variable at the school level representing the degree to which college prep course enrollment changed for students in that ability group in that school compared to pre-policy levels. Initial models included variables representing school characteristics, but these variables were removed for parsimony as they did not change the estimates of policy effects:

$\beta_{m0k} = \gamma_{m00} + \gamma_{m01}$(% college-prep enrollment pre-policy for group m)$_k + u_{00k}$
$\beta_{m1k} = \gamma_{m10} + \gamma_{m11}$(change in % college-prep enrollment post-policy for group m)$_k$
$\beta_{m2k} = \gamma_{m20} + \gamma_{m21}$(change in % college-prep enrollment post-policy for group m)$_k$
$\beta_{m3k} = \gamma_{m30} + \gamma_{m31}$(change in % college-prep enrollment post-policy for group m)$_k$

In the above models, the intercept γ_{m00} represents average pre-policy outcome for students in ability group *m* in schools with 100% pre-policy college-preparatory course enrollment (i.e., schools that did not have to change the ways they enrolled students in ability group m with the policy because they already enrolled them all in college preparatory courses without the policy). Post-policy intercepts (γ_{m10}, γ_{m20}, and γ_{m30}) represent, respectively, the average early, mid, and late post-policy changes in outcomes for schools with no changes in remedial/college-prep enrollment. The coefficients of interest are γ_{m11}, γ_{m21}, and γ_{m31}; these represent the extent to which changes in college preparatory course enrollment were associated with changes in academic outcomes in early, mid, and late post-policy periods, respectively, for students at each ability level. If enrolling in college preparatory courses instead of remedial courses affected students' outcomes, we should see that schools that changed the most in terms of the degree to which they enrolled students in college preparatory versus coursework also showed the largest changes in students' outcomes. The numbers in Table 2 are based on the coefficients $\gamma_{m11,}$ $\gamma_{m21,}$ and $\gamma_{m31.}$

The analyses of total policy effects are similar to those of enrollment effects, but instead of using change in enrollment by ability group as the key independent variable at level three, we use a dummy variable indicating whether the school was affected by the policy. Schools are considered to have been affected by the policy if they enrolled at least 25 percent of their lowest-ability (group 1) students in remedial courses pre-policy. (Schools did not enroll average-ability students in remedial classes unless they also enrolled low-ability students in remedial classes. If a school already enrolled almost all of very low ability students in college preparatory courses pre-policy, it would not be substantially affected by the policy—its students took college preparatory courses in the absence of the policy.) The coefficients of interest are the same as in the previous analyses, but they represent the total effect of the policy on schools that did not already enroll all students in college preparatory courses prior to the policy.

Schools with many low-ability students would likely experience more course programming changes and demands on school capacity than schools with few low-ability students, and the degree of change in the school might affect all students' outcomes. Therefore, we also included a variable for the percent of low-ability students in the school, and an interaction of the percent low-ability with whether the school was affected by the policy. The numbers in Table 3 are based on the coefficients $\gamma_{m11,}$ $\gamma_{m21,}$ and γ_{m31} from models that use the dummy variables for whether a school was affected by the policy rather than the percentage change in college preparatory enrollment.

Appendix B. Evidence of Substantive Change in Ninth Grade English and Math Classrooms

We would expect no policy effects if high schools had dropped remedial course titles without making more substantive changes in the curriculum. However, we have some evidence that the curricular changes that occurred in response to the policy were not merely cosmetic.

More heterogeneous ninth grade math and English classes. There is evidence of some detracking with implementation of the policy. Pre-policy, the lowest-ability students were, on average, in math classrooms where average test scores were .16 standard deviations lower than the school average; post-policy, their classrooms had mean achievement levels close to the school average achievement, only .04 standard deviations lower. (See Table B-1.) Similar patterns can be seen in English classrooms.

More algebra content in ninth grade math classes. In the year prior to the policy, and in the year after the policy, CPS teachers answered surveys about their schools and their instructional practices.[1] Responses from participants suggest that there was a shift in content coverage in ninth grade math classes, see Table B-2.[2] In both years, the typical (median) ninth grade math classroom spent about 60 percent of instructional time on algebra topics. However, in the year before the policy (1996-97), classrooms at the 25th percentile covered algebra for only about one-third of class time, while in the second year of the policy (1998-99), classrooms at the 25th percentile spent almost half (48 percent) of instructional time on algebra topics. This is what

---

[1] Surveys were sent to all CPS high schools, and response rates were moderate—51 percent responded to the 1996-97 survey, while 44 percent responded to the 1998-99 survey. While only about half of teachers responded to the surveys, participating schools and teachers were representative of the district as a whole.

[2] Indicators of the percent of instructional time spent on algebra for each teacher were constructed by dividing the total time teachers said they spent on algebra or advanced topics by the total time they said they spent on all topics. Examples of topics on which teachers responded include: *Non-algebra* - Associative, communicative, distributive properties; Absolute value; Multi-digit addition; Multi-digit subtraction; Multi-digit multiplication; Long division; Operations with decimals; Operations with fractions; Operations with negative numbers; Exponents and roots; *Algebra* - Equations with 1 unknown; Solving inequalities; Simplifying algebraic expressions; Factoring algebraic expressions; Equations of lines; Solving quadratics; Graphing equations; Solving 2 equations with 2 unknowns

we would expect from the policy—regular algebra classrooms should have changed little in terms of algebra content, while fewer classes should have had minimal algebra content.

More demanding reading materials in ninth grade English classes. English teachers who participated in the surveys reported using different types of reading materials in 1998-99, compared to 1996-97, and these changes were in the direction that would be expected if more students were taking college preparatory English instead of remedial English. Fewer post-policy teachers than pre-policy teachers reported assigning their students to read textbook chapters (39 percent postpolicy; 55 percent prepolicy), while more post-policy teachers reported assigning their students to read novels (96 percent postpolicy; 85 percent prepolicy), short stories (100 percent versus 91 percent), poetry (94 percent versus 77 percent), non-fiction (87 percent versus 69 percent), and plays/film scripts (76 percent versus 70 percent).

Appendix Table B-1. Changes in classroom average ability levels by students' incoming ability levels.

		Lowest Ability	Low Ability	Average Ability	High Ability
Math classroom average achievement	Pre-policy	-0.16	0.03	0.20	0.53
	Post-policy	-0.04	0.00	0.08	0.37
English classroom average achievement	Pre-policy	-0.31	-0.01	0.21	0.65
	Post-policy	-0.24	-0.10	0.03	0.42

Ability is measured with students' standardized 8th grade test scores in math or English subtracted from the school mean. A value of "0" for class average achievement means that students are in classrooms with ability levels at the school average; a value of 0.5 means students are in classrooms with achievement levels a half of a standard deviation above the school mean. Pre-policy rows show classroom ability levels among students entering high school from 1994 to 1996. Post-policy rows represent students entering in post-policy period 2 (1998 and 1999).

Appendix Table B-2. Percentage of Instructional Time Spent on Algebra in Ninth Grade Math Classes

Classrooms at the:	1996-97	1998-99
90th Percentile	87%	87%
75th Percentile (Top quartile)	73%	77%
50th (Median)	57%	59%
25th (Bottom quartile)	34%	48%
10th Percentile	9%	21%