Educational Evaluation and Policy Analysis
December 2009, Vol. 31, No. 4, pp. 367–391
DOI: 10.3102/0162373709343471

College Preparatory Curriculum for All: Academic Consequences of Requiring Algebra and English I for Ninth Graders in Chicago

Elaine Allensworth
Takako Nomi
Nicholas Montgomery
University of Chicago

Valerie E. Lee
University of Michigan

There is a national movement to universalize the high school curriculum so that all students graduate prepared for college. The present work evaluates a policy in Chicago that ended remedial classes and mandated college preparatory course work for all students. Based on an interrupted time-series cohort design with multiple comparisons, this study found that the policy reduced inequities in ninth grade course work 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 slightly declined, test scores did not improve, and students were no more likely to enter college. In sum, few benefits resulted from universalizing college preparatory course work among freshmen, but dropout rates did not increase. Possible explanations are discussed.

Keywords: *curriculum, high school, college preparatory, ninth grade, algebra, tracking*

WHAT should students learn in high school? Should all students learn the same skills and content, or should course work reflect students' abilities, their motivation to learn, and their plans for the future? Such questions have been debated hotly 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 course work 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

We gratefully acknowledge support for this work from the Institute for Educational Sciences, U.S. Department of Education, Grant No. R305R060059. However, the responsibility for the findings and conclusions reported in this manuscript are those of 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. We also thank the anonymous reviewers for their thorough and thoughtful feedback. For information on this study, contact Elaine Allensworth (e-mail: elainea@ccsr.uchicago.edu).

by offering academic courses only, high schools overlook two realities about students: First, they enter high school with different intellectual capacities and skills; second, they aspire to disparate occupations. To address these realities, proponents of social efficiency 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, as reflected in a typical public high school curriculum that is broad and diffuse—with different courses in any subject and with similar courses at varying levels of difficulty (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 course work, 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).

Although 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 course taking (e.g., Lee, 2002; Lieberman, 1995; Newmann and Associates, 1996). 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 (Lee, 2002; Oakes, 1985, 2005). 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 suggests that there may be advantages to homogeneous classes, for organizational reasons, but that low-track students should receive more challenging course work and better instruction than what they receive under traditional tracking (e.g., Hallinan, 1994; Loveless, 1999). Despite disagreement about the practice of differentiating course work by student ability, both perspectives suggest that low-ability students be exposed to more rigorous course work than that which has been typical.

The Call for Increased Rigor

While researchers have voiced concerns about the social stratification that is inherent in the differentiated curriculum, policymakers 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 suggesting 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, policymakers have concluded that the skills needed for success in the workforce are the same as those needed to succeed in college (American Diploma Project, 2004). Thus, low-level course work increasingly is viewed as being insufficient to prepare any students for life after high school.

The criticisms raised in the 1980s about low academic standards, with concerns about the current workforce, have motivated a national movement calling for rigorous high school course requirements. The National Governors Association (2005) recommended toughening high school graduation requirements to insist on college preparatory course work for everyone. Policy reports from ACT (2004) and the American Diploma Project (2004) have advocated increasing science and mathematics course work and raising standards to improve alignment between secondary and postsecondary curricula. Policymakers 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). In fact, 13 states now require a college prep curriculum, and 16 more plan to adopt such requirements in the near future (Achieve, 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 Course Work 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, as well as multilevel statistical methods, to link high school curriculum structure to student outcomes. The work began within a comparative Catholic school–public school 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, Smerdon, & 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: First, achievement gains were greater; second, learning was distributed more equitably by race, ethnicity, and socioeconomic status.

Support for policies requiring rigorous courses also comes from studies that directly link students' course taking and achievement. These studies document strong relationships between the courses that 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 do 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 do those without advanced course work (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 do their counterparts who complete less-demanding course work (ACT, 2004; Adelman, 1995; 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 course type is more important than course quantity—that is, it is not the number of courses that students complete but which ones (e.g., Lee, 2002; Lee, Croninger, et al., 1997; Lee et al., 1998; Lee & Smith, 1995; Lee, Smith, et al., 1997). Replacing remedial course work with college preparatory course work is consistent with this research.

Recent Research on Tracking and Detracking

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 course work 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., Gamoran, Nystrand, Berends, & LePore, 1995; Lucas, 1999; Oakes, 2005; Powell et al., 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 et al., 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, in the loss of academically talented students, and in 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 owing to teachers' low expectations for students (Rubin, 2008; Wells & Oakes, 1996). Although some schools have detracked successfully classrooms and improved instruction for low-ability students (Boaler, 1997; Boaler & Staples, 2008; Oakes, 1994; 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 that students follow to be college focused will improve their academic outcomes. Yet, research on detracking has suggested cultural and structural limitations to universalizing a curriculum such 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 and schools that did not (cf., Lee, Croninger, et al., 1997; Lee et al., 1998; Lee & Smith, 1995; Lee, Smith, et al., 1997). However, schools that had developed the capacity to enroll all students in college preparatory course work may have been different from 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 course work. Although prior studies generally have 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 course work and so may affect students' outcomes.

Likewise, studies based on comparisons of students who were in college preparatory tracks or who took advanced course work, as compared to 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 and 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 nonexistent than 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 course work 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 course work for a reason; namely, 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 course work. Rather than explore differential effects, most studies of the high school curriculum used a linear control for ability. One prior study did estimate differential effects of college prep course work (Gamoran & Hannigan, 2000); it found that students at or below the 20th ability-based percentile benefited less than more-able students from taking college prep classes on a test of math achievement in grade. 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 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 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 such as 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. Such schools may also 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, before this policy, 19% of ninth graders failed their ninth-grade English course; a quarter failed their math course; and students averaged over 3 weeks of course absence per semester. In schools with such high levels of failure and absenteeism, it may be particularly difficult to increase effectively 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 and offer college preparatory courses, many of which struggle with low achievement and weak capacity, we show the effects of mandating college preparatory course work in a challenging context.

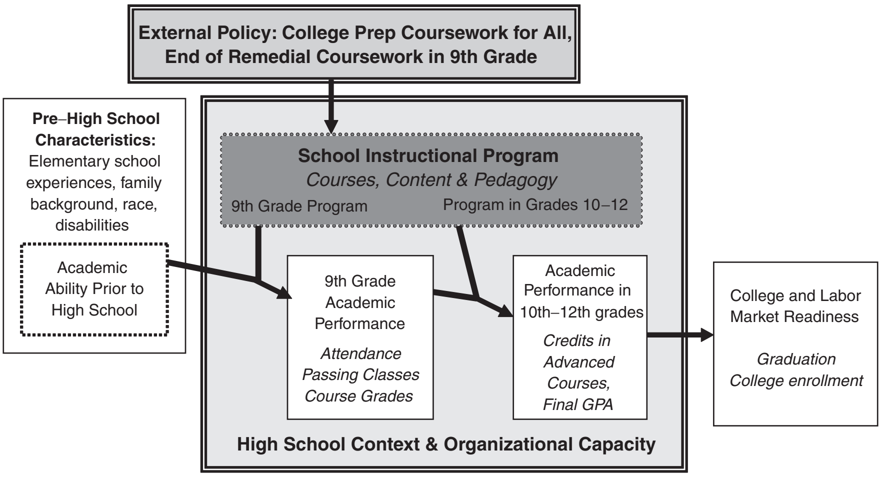

FIGURE 1. *Conceptual model of how curricular policy influences student outcomes.*

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: first, the literature in which the topic is embedded (and the shortcomings of applying this literature to the Chicago context); second, a conceptual model that describes the mechanisms that may link the policy to the set of student outcomes we investigated. 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 the structure of his or her school's instructional program (e.g., whether it requires college prep or remedial courses) and the characteristics of the student's background, 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 such that all students have the same course experiences, which should prepare them for advanced course work in later grades. This should result in higher achievement early and late in high school, particularly among students who would otherwise take less-demanding courses. Following such a curriculum should ultimately improve students' postsecondary outcomes.

Yet, as shown in Figure 1, the external policy mandate depends on a number of mediating factors if it is to affect outcomes both 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—because ninth-grade course work serves as a gatekeeper for more advanced study and remedial ninth-grade course work was common prepolicy in both subjects. By basing this study in Chicago, we are specifically studying the effects of requiring college prep course work 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 postpolicy 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 the academic outcomes of students improve by taking college prep classes instead of remedial classes, and did the effects differ by their 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 versus remedial), and it applies only to students whose course work 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 depend not only 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 course work 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, which 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 course work 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 (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% African American, 38% Latino, 9% White, and 3% Asian. Approximately 85% 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). 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, and elementary and high school achievement test scores, as well as data from the National Student Clearinghouse on college enrollment and 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 the following: the effects of the policy on students' course taking, students' academic outcomes, the characteristics of students as they entered high school, and 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 six-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 postpolicy changes in course taking and outcomes, we developed cohort-level dummy variables distinguishing four policy periods: prepolicy cohorts (before 1997), the first year of the policy (1997), a midpolicy period (1998–2000), and a late policy period (after 2000). We used these cohort indicators to compare enrollment and academic outcomes in postpolicy years to prepolicy years. For simplicity, we present only the findings for midpolicy 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 levels of incoming ability. For each high school, we computed the proportion of students enrolled in college prep courses prepolicy in each of four ability groups; then, we computed the change in enrollment between pre- and postpolicy periods (e.g., the percentage of very low-ability students enrolled in college prep classes in the school postpolicy minus the percentage of very low-ability students enrolled in college prep classes in that school prepolicy). The second indicator (used to address Research Question 3) was a simple dummy-coded variable of whether the school was affected by the policy or not. We considered schools that enrolled at least 25% of their lowest-ability students in remedial course work prepolicy as influenced by the policy (coded 1), whereas those that already enrolled 75% or more of their lowest-ability students prepolicy were coded 0, because they were largely unaffected by the policy—that is, 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) and for failing their ninth-grade English/math course (regardless of level), as well as 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, 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 (postgeometry and post–Algebra II), and enrollment in a 4-year college within a year after high school.

Although our focus is on the effects of courses in the ninth grade, we include long-term outcomes because ninth-grade course work often determines what students take in subsequent years. By beginning college preparatory sequences in ninth grade, students should have greater opportunities to take advanced course work than if they wait 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 precalculus. 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 course enrollments and student outcomes are associated with ability as students enter high school. To measure precisely 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 standard deviations below the mean or lower; Group 2, –0.5 standard deviations to the mean; Group 3, from the mean to 0.5 standard deviations above; and Group 4, more than 0.5 standard deviations 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 within each ability level, where students in other ability groups were coded zero.[10] Including these variables within each ability category allowed us to control more precisely 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 before high school, special education eligibility, and ESL (English as second language) status, as measured with dummy variables. In addition, we controlled for socioeconomic status with two variables constructed from the U.S. census data on students' residential block groups, linked by students' home addresses: first, concentration of poverty, a composite of male unemployment rate and the percentage of families under the poverty line; second, social status, a composite of the median family income and the average educational attainment.

Cohort-level control variables. We were concerned that the 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 students' average ability levels change over time. Therefore, we controlled for the average incoming latent ability of students in each school in each cohort. This variable has the same value for students of all ability levels within a school in the same ninth-grade 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 the academic compositions—that is, proportions of students in special education, average incoming ability, ability heterogeneity, and whether or not the school was a vocational or magnet school. Almost all the school-level control variables subsequently were omitted from final models owing to nonsignificance on policy effects.

Analysis

Our analyses are presented in three parts. To address Research Question 1 (course taking), we show enrollment changes in English I and Algebra I 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, because 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 postpolicy, reaching nearly 100% for all schools by the midpolicy years. We take advantage of this shift to compare the outcomes of students in postpolicy cohorts (almost all of whom enrolled in college preparatory courses) to those of 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 postpolicy years that 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 prepolicy 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 prepolicy remedial course enrollment among students with the same ability levels.[13] Only schools that offered remedial courses prepolicy were affected by the mandate to end remedial course work, whereas all schools would be affected by other CPS policies. Therefore, we were able to compare (a) changes in students' outcomes in schools that were affected by this particular policy to (b) changes in students' outcomes in schools not affected by this policy; that is, the comparison schools serve as a control for other reforms occurring simultaneously. Combining cross-sectional and longitudinal comparisons allowed for more confidence in the results of the analyses than what either method would allow on its own.

Thus, our analyses comprise two levels of comparison, providing a difference-in-difference approach. First, we estimated how students in each school performed in postpolicy cohorts, as compared to students with the same incoming ability in the same school prepolicy. 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 course work).[14] Our analyses used three-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 students 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 lies in how we measure changes in course enrollment. To discern enrollment effects (Research Question 2), the key variable is the percentage change in enrollment in college preparatory classes (English, math, or both) for each ability group within each school, as compared to that of prepolicy years. To discern policy effects (Research Question 3), 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: First, by 2000, virtually all CPS ninth graders were enrolled in both English I and Algebra I; second, the policy most strongly influenced course enrollment among low-ability students but had almost no effect on course enrollment among students of high ability, given that such students had

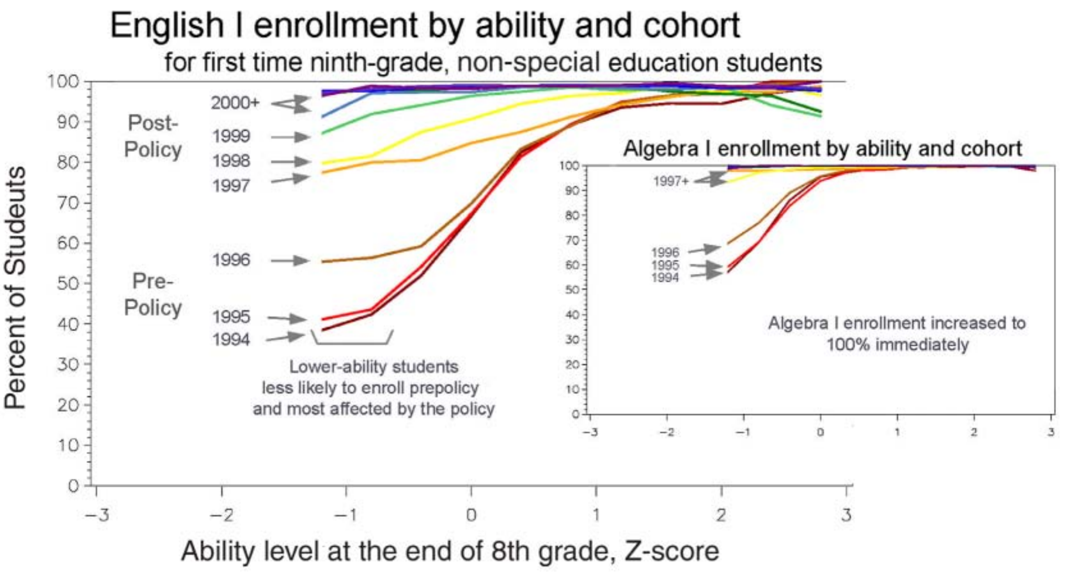

FIGURE 2. *Enrollment in college prep English and math courses by freshman cohort.*

TABLE 1
Ninth-Grade College Prep Math and English Enrollment Prepolicy and Postpolicy (in percentages)

	All students				Non-special education students				Special education students			
	Math		English		Math		English		Math		English	
	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

Note. The postpolicy statistics are averaged across all the postpolicy years. Compliance was close to 100% by 2000 for all but the lowest-ability students eligible for special education services.

previously enrolled in college preparatory course work before policy was enacted; and, third, implementation moved more rapidly in mathematics than in English.

These trends are reflected in the numerical results in Table 1. During the postpolicy years, close to 100% of students not eligible for special education services were enrolled in Algebra I, regardless of ability; 96% were enrolled in English I. Gaps in course enrollments by race/ethnicity that existed prepolicy largely closed postpolicy. Although very low-ability students eligible for special education services were not in full compliance by the third policy year (2000), the policy much more strongly affected the enrollment of special education students into college preparatory courses than it did that of regular education students, given that the former had the lowest college prep enrollment rates before the policy's implementation (42% to 48%).

Although the policy brought large shifts in course enrollment, the observed changes were potentially superficial: Schools could simply rename remedial courses while maintaining students' experiences. 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 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 postpolicy than were students with similar incoming skills prepolicy. 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 report assigning students to read novels, poetry, nonfiction, 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.

Because prepolicy remedial course enrollment was defined strongly 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 course work in the prepolicy years, whereas schools serving more high-achieving students would have been less likely to offer remedial course work prepolicy. 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 schools enrolled students in remedial courses prepolicy.[15] We do not include these tables here because of the preponderance of no-difference findings (tables are, however, available upon request). After many school-level characteristics were taken into account, considerable variation in prepolicy college preparatory course enrollment remained between schools that otherwise served students of comparable ability. We found full enrollment in college prep course work in many schools in the prepolicy period that served predominantly low-ability students, whereas many schools serving generally high-ability students had substantial enrollment in remedial course work among their low-ability ninth graders. We capitalized on this unexpected finding—considerable variability in prepolicy college prep course enrollment among schools with similar observed characteristics—by incorporating a second school-level contrast based on prepolicy 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 (described 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 an increase in college prep enrollment by 20 percentage points (e.g., the effects of moving from 80% algebra enrollment prepolicy to 100% algebra enrollment postpolicy). We group the 15 academic outcomes into three categories: ninth-grade performance in mathematics, ninth-grade performance in English, and 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: school-level effect sizes for comparability (in the left panel) and meaningful units such as percentage points and test score points (in the right panel). For simplicity, we show only the midpolicy period contrast (1998–1999).[17] Results of the statistical tests are indicated in only the lefthand panel, although they apply equally to the righthand panel.

TABLE 2

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

Ability	Effect of increased enrollment (+20 percentage points): Effect sizes					Percentages, days, grade points				
Math	Algebra credit	Failure	Absences	GPA	Test scores	Algebra credit	Failure	Absences	GPA	Test scores
Lowest	1.11**	0.31**	0.00	–0.25**	0.06	8.9%	3.0%	—	–0.06	—
Low	1.22**	0.41**	0.08†	–0.34*	–0.02	11.6%	3.5%	0.70	–0.08	—
Average	0.96*	1.24*	0.21*	–0.76*	–0.07	8.0%	8.9%	1.60	–0.18	—
English	English I credit	Failure	Absences	GPA	Reading test	English I credit	Failure	Absences	GPA	Reading test
Lowest	1.65**	0.07	–0.05*	–0.09	0.01	14.7%	—	–0.33	—	—
Low	1.56**	0.07	–0.05*	–0.09†	0.00	12.8%	—	–0.32	–0.02	—
Average	1.60**	0.14	–0.05	–0.09	–0.03	12.4%	—	—	—	—
Long term	Post-geometry credit	Adv math credit[a]	Graduate	Final GPA	4-year college	Post-geo credit	Adv math credit[a]	Graduate	Final GPA	4-year college
Lowest	0.40**	0.07	–0.14	0.07	0.03	3.3%	—	—	—	—
Low	0.35†	0.10	–0.05	–0.07†	–0.09†	3.5%	—	—	–0.02	–0.4%
Average	0.05	–0.14	0.05	–0.14*	–0.13*	—	—	—	–0.04	–0.7%

Note. This table shows the effects for Period 2 only (1998–1999); similar effects were observed in other postpolicy periods. We calculated values by taking the difference between the pre- and postpolicy 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-postpolicy percentage 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 at http://epa.sagepub.com/supplemental.

a. Beyond Algebra II.

$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$.

The results 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 change of 20 percentage points is unrealistically high for average-ability students and atypically low for the lowest-ability students in affected schools. Because so few students in the highest-ability group took remedial courses prepolicy, we do not report results for this group in Table 2, although they were included in the statistical models.

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 finding 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 low- and average-ability students.

Across all ability groups, an increase in Algebra I enrollment by 20 percentage points 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 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 7 of 48 questions on the Tests of Academic Proficiency 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 (e.g., nonfiction passages and poetry), about half 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; specifically, an increase of 20 percentage points in taking algebra led to an increase of 3 percentage points in lowest-ability students' earning credits beyond geometry ($p < .01$); however, this did not hold beyond Algebra II. Even though students in the postpolicy period could take the college prep math sequence up to precalculus (because they started the math sequence in ninth grade 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 found in the ninth-grade mathematics class, final high school GPAs went down slightly ($p < .05$) for all but the lowest-ability students. Perhaps because the grades were lower, the probability of attending a 4-year college after high school decreased slightly for all but the lowest-ability students by less than 1 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. They also do not 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 prepolicy to those that did not. We present the results in Table 3 in their natural metrics, parallel to the righthand panel in Table 2.

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 for 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 prepolicy. Thus, overall policy effects on average-ability students' academic outcomes, as shown in Table 3, are small. Compared to those of the prepolicy years, students in the two lower-ability groups postpolicy 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

TABLE 3
Total Policy Effects on Ninth-Grade Academic Outcomes by Ability Level

Ability	Student outcomes				
Math	Algebra credit	Failure	Absences[a]	GPA[b]	Test scores[c]
Lowest	8.8%**	7.7%*	2.24	–0.15**	0.63
Low	7.4%**	3.6%	2.40	–0.07	0.13
Average	1.0%	1.3%	3.14*	–0.01	–0.17
High	–0.4%	1.2%	1.67	0.10	0.50
English	English I credit	Failure	Absences[a]	GPA[b]	Test scores[c]
Lowest	35.8%**	4.1%	1.15	–0.15*	–0.61
Low	28.2%**	4.0%†	1.46†	–0.11†	–0.64
Average	11.7%**	2.1%	2.11*	–0.06	–0.54
High	–3.2%	2.6%	1.70*	–0.07	–1.11†
Long term	Post-geometry credit	Adv math[d]	Graduate	Final GPA[b]	4-year college
Lowest	1.0%	–0.3%	4.6%†	0.10*	0.6%
Low	0.6%	–2.0%	–1.7%	–0.05	–2.8%**
Average	–0.3%	–2.5%	0.2%	–0.04	–1.1%
High	–0.9%	–1.9%	–0.7%	0.00	–2.0%

Note. This table shows the effects for Period 2 only (1998–1999). Values were calculated by taking the difference between the pre- and postpolicy change for schools that changed enrollment and for those that did not change enrollment. The values were converted into their natural metric. Coefficients are available in their original units at http://epa.sagepub.com/supplemental.
a. Days.
b. Grade points.
c. Normal curve equivalents.
d. Beyond Algebra II.
†$p < .10$. *$p < .05$. **$p < .01$.

points ($p < .05$) in the postpolicy period compared to the prepolicy period. Average-ability students were absent more often (3.14 days; $p < .05$), and the lowest-ability students' math GPAs declined by 0.15 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—particularly, the lowest-ability students (increases of 12% to 36%; $p < .01$). This finding reflects the extensive use of remedial English in prepolicy years. The two low-ability groups were marginally more likely to fail their ninth-grade English course (an increase of 4.0 percentage points, but $p > .05$), and their English grades were slightly lower (by 0.11 to 0.15 points; $p < .10$). Absenteeism in English I classes increased somewhat postpolicy, 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 nonexistent. A few of the prepolicy–postpolicy comparisons emerge as being significant but not in a consistent way across ability groups. The risk of Type I errors arising from multiple comparisons makes us less confident that these differences in long-term outcomes are real (i.e., 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 prepolicy period ($p < .01$); other ability groups showed no change in college entrance. Thus, a policy that replaced remedial with college prep course work 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 4-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 policy ending remedial classes was implemented fully 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 course work that had been common among low-ability students. The social distribution of course taking also became more equitable: Enrollment differences in college prep course taking 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. This is positive news.

The extant literature, which identified generally positive effects on students from their taking college preparatory course work, suggested that we would find larger policy benefits for lower-ability students because their course enrollment would be most affected. On one set of outcomes—namely, 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 10% 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 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%, there is great need to improve graduation rates. Requiring all students to take college prep courses rather than accumulate credits through remedial course work did not aggravate an already high dropout rate.

The remainder of the news is not good. On most of this array of outcomes, measured at the end of ninth grade and at the end of high school, the policy had few positive effects and several negative effects. Among the students who took Algebra I instead of remedial math, ninth-grade mathematics grades declined and math failure rates increased. 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 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 mention 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 schools for them to attend (sometimes by careful choice of residence), those who were 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 there is 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 course offerings and their potential students rather than the way that such 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 rather than 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 before 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 before the policy subsequently were 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. Before the implementation of this policy, ninth-grade courses were tracked highly 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 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 that have been suggested as being 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 (Nomi & Allensworth, 2009).

One problem noticed in case studies of detracking is the challenge of teaching heterogeneous classes; that is, teachers often target instruction to the hypothetical middle student (Bar & 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, which 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, as compared to those of their classroom peers, likely increased their probability of failure. Loveless (1999) noted the possibility of negative effects on students' self-esteem, and this study provides some suggestive evidence 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, which might explain the larger course enrollment effects on average-ability students—that is, they would have been the highest-ability students in remedial classes but the 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—one could argue that these courses are beyond the capability of some students or that intellectually demanding courses are not helpful to students if they do not want them, do not want to work hard to master the content, or do not see them as being useful for their futures. Such perspectives seem to lie behind resistance to detracking in some schools (Oakes & Wells, 1996; Rubin, 2008). However, at a time when the majority of U.S. students want a college degree and when the workforce demands higher skills than it has 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 show few benefits). Thus, the policy generally had a null effect—that is, no gains but no costs. This does not suggest that low-skill students cannot handle college prep course work but rather that they do not benefit from it any more than they do from remedial course work. Furthermore, 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 explained entirely by students' insufficient academic skills. Instead, the findings suggest that more attention be paid to how students are taught and to the quality and depth of the tasks in which students are engaged rather than the content of what they are taught (for a case study, see 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 noncognitive skills and behavior in high school and earlier grades. It seems unreasonable to mandate a policy that makes greater cognitive demands on students and to 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 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 course work. However, weak academic skills are not the primary source of course failure in Chicago schools; students' academic behaviors (attendance and completing homework) are 8 times more predictive of failure than their test scores are (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 prepolicy period and the postpolicy period in Chicago high school students' grades. Prepolicy, the average ninth-grade math or English grade in CPS was below a C (an average of about 1.8 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 postpolicy. 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 engage minimally in their courses and irregularly attend school, 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 feeling excited about learning, attending school and classes regularly, and working hard in whatever course they are taking.

Conclusion

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 that we have used to estimate our findings, it would seem that the most common explanation for no-difference findings—namely, 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 course work 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 course work, 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 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) 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 helped low-ability students more successfully 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.

Appendix A
Statistical Models

Both sets of analyses of enrollment and 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

$$Y_{ijk} = \pi_{1jk}(\mathit{Ability\ Level\ 1})_{ijk} + \pi_{2jk}(\mathit{Ability\ Level\ 2})_{ijk} + \pi_{3jk}(\mathit{Ability\ Level\ 3})_{ijk} + \pi_{4jk}(\mathit{Ability\ Level\ 4})_{ijk} + \sum_{P=1}^{P} \pi_{4+pjk}\,(\mathbf{X})_{ijk} + e_{ijk},$$

where $\mathbf{X}$ 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 (i.e., 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 ($\mathbf{X}_{1\ldots p}$). 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,

$$\pi_{mjk} = \beta_{m0k} + \beta_{m1k}(\mathit{early\ postpolicy})_{jk} + \beta_{m2k}(\mathit{mid\text{-}postpolicy})_{jk} + \beta_{m3k}(\mathit{late\ postpolicy})_{jk} + \beta_{m4k}(\mathit{cohort\ average\ latent\ ability})_{jk} + r_{mjk}.$$

The intercept β_{m0k} represents the average prepolicy outcome at ability level m in school k, and the coefficients β_{m1k}, β_{m2k}, and β_{m3k} represent the change in the average outcome for students in ability group m at each school from prepolicy to postpolicy. These postpolicy coefficients do not necessarily represent policy effects, because factors other than the policy might

(continued)

have affected postpolicy changes in the outcomes. For example, test scores may have improved in all schools systemwide because of a new emphasis on test-based accountability during the postpolicy period; but this would have nothing to do with the policy that we are studying. To estimate the policy effects from ending remedial course work, we compared postpolicy changes in the outcomes for each school (β_{m1k}, β_{m2k}, and β_{m3k}) by the degree to which schools increased college preparatory course enrollment postpolicy. If the policy had an effect, then outcomes should have changed at a different rate in schools that had to respond to the policy (i.e., schools that previously offered remedial classes) than in schools that did not (i.e., schools that already enrolled all students in college preparatory course work). Furthermore, schools with the largest changes in course enrollment should show the largest changes in academic outcomes. It is with these coefficients that we distinguish policy effects from other changes occurring in Chicago's schools during the same period.

In the analysis of college preparatory course enrollment effects, we estimated the following models:

$\beta_{m0k} = \gamma_{m00} + \gamma_{m01}$(percentage college prep enrollment prepolicy for group $m)_k + u_{00k}$;

$\beta_{m1k} = \gamma_{m10} + \gamma_{m11}$(change in percentage college prep enrollment postpolicy for group $m)_k$;

$\beta_{m2k} = \gamma_{m20} + \gamma_{m21}$(change in percentage college prep enrollment postpolicy for group $m)_k$;

$\beta_{m3k} = \gamma_{m30} + \gamma_{m31}$(change in percentage college prep enrollment postpolicy for group $m)_k$;

The average prepolicy outcome (β_{m0k}) in ability group m is estimated as a function of the percentage of prepolicy college prep enrollment for that ability group, which was centered on 100%. Thus, the intercept γ_{m00} represents the average prepolicy outcome for students in ability group m in schools with 100% prepolicy college preparatory course enrollment (i.e., schools that did not have to change course enrollment with the policy). The average postpolicy changes (β_{m1k}, β_{m2k}, and β_{m3k} for early, mid-, and late postpolicy periods, respectively) are estimated as a function of changes in percentage college prep enrollment postpolicy for ability group m. Thus, postpolicy intercepts γ_{m10}, γ_{m20}, and γ_{m30} represent, respectively, the average early, mid-, and late postpolicy changes in outcomes for schools with no changes in college preparatory enrollment. These intercepts pick up changes in outcomes that should not be attributed to the policy. The effects of enrolling students in college preparatory courses, instead of remedial courses, are captured by the coefficients γ_{m11}, γ_{m21}, and γ_{m31}, representing the extent to which changes in college preparatory course enrollment were associated with changes in academic outcomes beyond the changes observed in schools unaffected by the policy, 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 increased college preparatory enrollment the most also showed the largest changes in students' outcomes. The numbers in Table 2 are based on the coefficients γ_{m11}, γ_{m21}, and γ_{m31}. Initial models also included variables representing school characteristics, but these variables were removed for parsimony, given that they did not change the estimates of policy effects.

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 the school level, 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% of their lowest-ability students (Group 1) in remedial courses prepolicy. (Schools did not enroll average-ability students in remedial classes unless they enrolled low-ability students in remedial classes. If a school already enrolled almost all low-ability students in college preparatory courses prepolicy, 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 before the policy.

Schools with many low-ability students would likely experience more course programming changes and demands on school capacity than would schools with few low-ability students, and the degree of change in the school might affect all students' outcomes. Therefore, we included (a) a variable for the percentage of low-ability students in the school and (b) an interaction of the percentage of low-ability students with whether the school was affected by the policy. The numbers in Table 3 are based on the coefficients γ_{m11}, γ_{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. Prepolicy, the lowest-ability students were, on average, in math classrooms where average test scores were 0.16 standard deviations lower than the school average; postpolicy, their classrooms had mean achievement levels close to the school average achievement, only 0.04 standard deviations lower (see Table B1). Similar patterns can be seen in English classrooms.

TABLE B1
Changes in Classrooms' Average Ability Levels by Students' Incoming Ability Levels

		Ability			
Classroom	Period	Lowest	Low	Average	High
Math	Prepolicy	–0.16	0.03	0.20	0.53
	Postpolicy	–0.04	0.00	0.08	0.37
English	Prepolicy	–0.31	–0.01	0.21	0.65
	Postpolicy	–0.24	–0.10	0.03	0.42

Note. Prepolicy rows show classroom ability levels among students entering high school from 1994 to 1996. Postpolicy rows represent students entering in postpolicy Period 2 (1998 and 1999).

More Algebra Content in Ninth-Grade Math Classes

In the year before the policy and in the year after the policy, Chicago Public Schools teachers answered surveys about their schools and their instructional practices.[19] Responses from participants suggest a shift in content coverage in ninth-grade math classes (see Table B2).[20] In both years, the typical ninth-grade math classroom (50th percentile, or median) spent about 60% of instructional time on algebra topics. However, in the year before the policy (1996–1997), classrooms at the 25th percentile covered algebra for only about one third of class time, whereas in the second year of the policy (1998–1999), classrooms at the 25th percentile spent almost half (48%) of instructional time on algebra topics. This is what we would expect from the policy: Regular algebra classrooms should have changed little in terms of algebra content, whereas fewer classes should have had minimal algebra content.

TABLE B2
Percentage of Instructional Time Spent on Algebra in Ninth-Grade Math Classes by Classroom Percentile

Percentile	1996–1997	1998–1999
90th	87	87
75th	73	77
50th	57	59
25th	34	48
10th	9	21

More Demanding Reading Materials in Ninth-Grade English Classes

English teachers who participated in the surveys reported using different types of reading materials in 1998–1999, compared to 1996–1997, and these changes were in the direction that would be expected if more students were taking college preparatory English instead of remedial English. Fewer postpolicy teachers than prepolicy teachers reported assigning their students to read textbook chapters (39% postpolicy versus 55% prepolicy), whereas more postpolicy teachers reported assigning their students to read novels (96% versus 85%), short stories (100% versus 91%), poetry (94% versus 77%), nonfiction (87% versus 69%), and plays or film scripts (76% versus 70%).

Ability is measured with students' standardized eighth-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 that students are in classrooms with achievement levels a half standard deviation above the school mean.

Notes

[1]Propensity score analysis can claim to eliminate selection bias only 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 (2008) could not control for many factors that could conceivably affect outcome and course work (e.g., parental involvement, motivation, teacher support).

[2]The policy specified 4 years of specific English courses (survey literature, American literature, European literature, world literature), 3 years of specific math courses (algebra, geometry, advanced algebra), 3 years of science (biology, earth/space or environmental science, chemistry or physics), and 3 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 course taking to avoid these issues.

[3]To classify courses as *remedial*, *college prep*, or *elective*, we examined several fields on students' 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 course code satisfies a district graduation requirement and which requirement that course fulfils. If the course level in the transcript file indicated that 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 if the student was not taking a college prep course simultaneously 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 I, 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 postpolicy years to all prepolicy 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 postpolicy periods, we could see that the effects were similar in all three periods. We present the midpolicy period results because this is the period that should not have been affected by these other issues. Tables with results from all three postpolicy periods are available upon request.

[5]We use only lowest-ability students (Group 1) 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 because it would depend not only on whether schools offered remedial classes but also on what proportion of students in the school were of 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 & Easton, 2007).

[8]A two-level hierarchical linear model, 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 the model was run, students' test scores were equated through Rasch analysis to remove form and level effects.

[9]Preliminary analyses guided this categorization. We found that students with ability scores higher than 0.5 had nearly 100% enrollment rates in college preparatory courses prepolicy; as such, these students were categorized together. All the other students were grouped by intervals of 0.5 standard deviations.

[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 whom they served over this period, particularly in their incoming achievement levels. Therefore, we control for not only students' individual incoming abilities but also the average incoming abilities of students in their cohort within their school.

[12]These analyses are available upon request. 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 course work, including type of school (magnet, vocational, neighborhood), size, average incoming ability level, and 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 prepolicy course enrollment. Only schools with higher average ability distributions and more heterogeneous ability distribution were slightly less likely than other schools to enroll low-achieving students in college prep math courses. These small differences were driven by just a few schools.

[16]Coefficients from the full models are available at http://epa.sagepub.com/supplemental.

[17]We looked for different policy effects in each period, out of concern that the policy may have not been implemented fully 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 postpolicy periods.

[18]It may seem counterintuitive that graduation rates would increase when freshman failure rates rise. However, before the policy, many low-ability students did not take enough credits in their freshman year to potentially graduate in 4 years. By requiring students to take more classes in their freshman year, including algebra, they were more likely to obtain enough credits to graduate in 4 years (see Miller & Allensworth, 2002).

[19]Surveys were sent to all high schools within Chicago Public Schools, and response rates were moderate: 51% responded to the 1996–1997 survey whereas 44% responded to the 1998–1999 survey. Although only about half the teachers responded to the surveys, participating schools and teachers were representative of the district as a whole.

[20]Indicators of the percentage of instructional time spent on algebra for each teacher were constructed by dividing (a) the total time that teachers spent on algebra or advanced topics (as reported by teachers) by (b) the total time they spent on all topics. Sample topics on which teachers responded include the following:

Nonalgebra: associative, communicative, distributive properties; absolute value; multidigit addition; multidigit subtraction; multidigit multiplication; long division; operations with decimals; operations with fractions; operations with negative numbers; exponents and roots

Algebra: equations with one unknown, solving inequalities, simplifying algebraic expressions, factoring algebraic expressions, equations of lines, solving quadratics, graphing equations, solving two equations with two unknowns

References

Achieve. (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. (2004). *Crisis at the core: Preparing all students for college and work*. Iowa City, IA: Author.

ACT. (2007). *Benefits of a high school core curriculum for students in urban high schools.* Retrieved February 15, 2008, from http://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, E., & Easton, J. Q. (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*. Chicago: Consortium on Chicago School Research.

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

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*, 51–71.

Barr, R, & Dreeben, R. (1983). *How schools work.* Chicago: University of Chicago Press.

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

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.

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*, 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*, 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*, 241–254.

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

Gamoran, A., Nystrand, M., Berends, M., & LePore, P. C. (1995). An organizatinal analysis of the effects of ability grouping. *American Educational Research Journal*, *32*(4), 687–715.

Gamoran, A., & Weinstein, M. (1998). Differentiation and opportunity in restructured schools. *American Journal of Education, 106*, 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. (2002). *Restructuring high schools for equity and excellence: What works*. 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., 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*, 99–121.

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., & 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., 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.

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: Brookings Institution.

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

Miller, S. R., & Allensworth, E. M. (2002). Progress and problems: Student performance in CPS high schools, 1993 to 2000. In V. Lee (Ed.), *Reforming Chicago's high schools: Research perspectives on school and system level change* (pp. 51–87). Chicago: Consortium on Chicago School Research.

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

National Governors Association. (2005). *Getting it done: Ten steps to a state action agenda*. Retrieved November 8, 2005, from http://www.nga.org/Files/pdf/05warnerguide.pdf

Newmann, F. & Associates. (1996). *Authentic achievement: Restructuring schools for intellectual quality*. San Francisco: Jossey-Bass.

Nomi, Takako and Elaine Allensworth. (2009). “Double-Dose” Algebra as an Alternative Strategy to Remediation: Effects on Students’ Academic Outcomes. *Journal of Research on Educational Effectiveness, 2*, 111-148.

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

Oakes, J. & Wells, A. (1996). *Beyond the technicalities of school reform: Policy lessons from detracking schools.* Los Angeles, CA: Research for Democratic School Communities, University of California, Los Angeles Graduate School of Education and Information Studies.

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*, 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, IN: Lumina Foundation and Indiana Commission for Higher Education.

Teitelbaum, P. (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*, 31–57.

Wells, A. S., & Oakes, J. (1996). Potential pitfalls of systemic reform: Early lessons from research on detracking. *Sociology of Education, 69* (Special issue), 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.

Authors

ELAINE ALLENSWORTH is interim coexecutive director of the Consortium on Chicago School Research, at the University of Chicago, 1313 East 60th Street, Chicago, IL 60637; elainea@ccsr.uchicago.edu. She conducts research on high school graduation, high school reform, and college readiness.

TAKAKO NOMI is a senior research analyst at the Consortium on Chicago School Research, at the University of Chicago, 1313 East 60th Street, Chicago, IL 60637; nomit@ccsr.uchicago.edu. Her areas of specialization include urban education, education policy, inequality in education, school and classroom organization, and research methodology.

NICHOLAS MONTGOMERY is a research analyst at the Consortium on Chicago School Research, at the University of Chicago, 1313 East 60th Street, Chicago, IL 60637; nmontgomery@ccsr.uchicago.edu. He studies the effects of high school curriculum policy changes in the Chicago Public Schools and works with practitioners to develop scalable research-based data reports.

VALERIE LEE is a professor of education and a faculty associate at the University of Michigan’s Institute for Social Research, Ann Arbor (correspondence can be directed to 116 West Way, Mashpee, MA 02649). She teaches courses in research methods. Her research interests focus on organizational characteristics of schools (especially, high schools)—namely, characteristics related to excellence and equity.

Manuscript received June 17, 2008
Final revision received December 3, 2008
Accepted June 24, 2009