Ending Social Promotion:

Dropout Rates in Chicago after Implementation of the Eighth-Grade Promotion Gate

March 2004

By

Elaine Allensworth

CHARTING REFORM IN CHICAGO SERIES

CONSORTIUM ON CHICAGO SCHOOL RESEARCH

ACKNOWLEDGMENTS

The author is grateful to the many individuals who provided support and assistance in the production of this report. First and foremost, I thank Melissa Roderick who provided encouragement and feedback throughout all stages of this work. I also particularly thank Jenny Nagaoka whose work on the companion report, *Ending Social Promotion: The effects of retention,* made her the best source for information about the promotion-gate policy and its consequences for students. Consortium directors provided helpful advice on early drafts of this report, and I would particularly like to thank John Easton, Penny Sebring, and Stuart Luppescu for their thorough reviews. I would also like to thank the two outside reviewers, Aaron Pallas and Steve Raudenbush, for their many helpful comments. In addition, members of the Consortium Steering Committee provided useful feedback, and I particularly thank G. Alfred Hess for his comments. I am grateful to Melissa Dean for her careful editing and substantive review, and to Sandra Jennings for the beautiful graphics and the design of this report. Finally, this report would not have been possible without the work of the Consortium staff in developing and maintaining the necessary data and providing feedback through all stages of the analytic work. This study was made possible by the grants supporting the core research of the Consortium from the Joyce Foundation, the John D. and Catherine T. MacArthur Foundation, and the Spencer Foundation.

Ending Social Promotion:

Dropout Rates in Chicago after Implementation of the Eighth-Grade Promotion Gate

March 2004

Elaine Allensworth

TABLE OF CONTENTS

CONSORTIUM ON CHICAGO SCHOOL RESEARCH

John Booz

INTRODUCTION

Beginning with the 1995-96 eighth-grade class, the Chicago Public Schools (CPS) initiated a high-stakes testing policy that required students to meet a minimum score on the Iowa Tests of Basic Skills before being promoted to high school.[1] Those who failed either repeated the eighth grade or moved into new schools called "transition centers," which were created for students too old to remain in elementary school. About 1,800 eighth graders were held back from entering ninth grade that first year. Even more students were held back from entering ninth grade each of the following three years, with about 3,000 students repeating eighth grade or entering a transition center in the fall of 1997; 3,900 students in the fall of 1998; and 3,300 students in the fall of 1999.[2]

One of the biggest concerns about the implementation of a policy that holds students back is the increased risk that they will drop out of school. While the effects of system-mandated retention are not known, there is a great deal of research on grade retention resulting from the individual decisions of teachers and parents. Consistently, this research on teacher-initiated retention has shown that retained students are much more likely to drop out than students at the expected grade for their age.[3] Moreover, this research shows that retained students appear to be especially vulnerable to dropping out at young ages.[4]

Before the promotion-gate policy was implemented, Chicago already faced high dropout rates—about 43 percent of students dropped out of school by age 19.[5] Based on studies of teacher-initiated retention, there was substantial concern that because high-stakes testing would cause retention rates to rise, the policy would exacerbate the system's already severe dropout rate.

However, proponents of the policy believed the threat of retention would both encourage students to work harder and motivate teachers and schools to address better the educational needs of the lowest-achieving students. Because higher-achieving students are less likely to drop out than lower-achieving students, proponents argued that if the promotion gate brought about better achievement, it would, in turn, lower overall dropout rates.[6] The eighth-grade promotion gate, therefore, could have had two contradictory effects on dropout rates: raising the likelihood of dropping out by putting more students through the experience of retention, but lowering the likelihood of dropping out by improving achievement.

Existing studies on grade retention and dropping out are not sufficient to answer questions about the effects of a high-stakes promotion gate on dropout rates. [7] While previous research has shown a strong relationship between teacher-initiated retention and dropping out of school, this research may not be applicable to the high-stakes testing environment, where the context of retention is very different. A number of differences could be hypothesized, in terms of both the social-psychological effects of retention and the instructional responses of teachers and schools. For example, from the students' perspective, the explicit criteria for promotion under high-stakes testing might lead students to internalize failure more when retained, rather than blaming their school or teacher. Alternatively, students may be less likely to feel singled out or unfairly punished. If retention is relatively common in a school, as it is in some of the lowest-performing schools in Chicago, the stigma associated with being held back may not be as strong. Teachers may react differently to students retained under high-stakes testing than to those that have been singled out for retention. For example, teachers may be more likely to tailor instruction to a group of retained students than they would to a student retained alone. They also may be more likely to view retained students as victims of an external policy, rather than seeing deficits in particular students, and may be more sympathetic in their

How the Promotion-Gate Policy Works

Beginning in spring 1996, eighth-grade students were required to meet a minimum score on the Iowa Tests of Basic Skills (ITBS) in both reading and mathematics in order to be promoted to ninth grade. The test-score cutoff for promotion was set using the grade-equivalent metric, which is based on national norms. Using this metric, a student is considered on grade level if they obtain a score of their grade plus the number of months in the school year that have passed (e.g., a score of 8.8 would be on grade level for an eighth-grade student taking the exam in the eighth month of the school year). The cutoff score in the first year was two years below grade level, or 6.8 grade equivalents. In spring 1997, the eighth-grade cutoff score was raised to 7.0, and promotion gates were put in place at grades three and six. The district continued to increase the eighth-grade cutoff score each subsequent year, to 7.2 in 1998, 7.4 in 1999, and 7.7 in 2000. Students who did not meet the cutoff score in both reading and math were required to participate in a summer-school program called Summer Bridge, where they received instruction in reading and math aimed at helping them to pass the test. At the end of the summer, students took the ITBS again. Those who did not pass the second time either were retained in eighth grade or moved into a transition center. Transition centers were new schools designed specifically for students who failed the eighth-grade standards, but were too old to remain in elementary school.[1] Between 40 to 50 percent of students held back by the eighth-grade gate enrolled in transition centers each year, beginning in the 1997-98 school year.

Two groups of students—those who were in special education, and those who were in bilingual education for three years or less—were not held to the promotion criteria in the same ways as other students. The promotion decisions for special-education students were based on criteria outlined in their Individual Education Plans (IEP). At the start of the policy, bilingual students were excluded if they had been enrolled in a bilingual education program for less than three complete years as of the beginning of the school year. In 1999, the criterion was changed from three years to four years.

interactions. There are likely many mechanisms through which teacher-initiated retention leads to an elevated risk of dropping out that would differ from high-stakes testing-based retention. Furthermore, studies on teacher-initiated retention have been limited by the degree to which they could control for factors that spuriously inflate the relationship between dropping out and retention, such as overall disengagement from school, which leads to a greater likelihood of both retention and dropping out. These studies likely have produced inflated estimates of the effect of retention on the risk of dropping out. Therefore, the relationship of high-stakes testing-based retention to dropping out is very likely not the same as has been found in studies on teacher-initiated retention.

The potentially contradictory effects of high-stakes testing on dropout rates, both through rising retention rates and improving achievement, further complicate the question of the effects of high-stakes testing-based retention on dropping out. There are no studies that have tracked students over a number of years to determine the effects of this kind of promotion gate on the likelihood that students will eventually drop out.[8] This study fills the gap in knowledge about the consequences of the promotion gate on dropout rates by comparing dropout rates in Chicago before and after implementation of the eighth-grade gate. The report is organized around five central questions:

1. What happened to dropout rates after implementation of the eighth-grade promotion gate?
2. Did retention at the gate affect students' likelihood of dropping out?
3. Did retention at the gate lead students to drop out at an earlier age than they would have without the gate?
4. Did simultaneous improvements in student achievement lead more students to stay in school?
5. Were dropout trends different for subgroups of students: by race, gender, exclusion from testing, or age at which students encountered the promotion gate?

Until spring 2001, the eighth-grade promotion standards only incorporated test-score cutoffs, with no explicit provisions made for exceptions. Still, in each year, about one-third of the students who did not meet the cutoff at the end of the summer were promoted. Latino eighth graders were slightly more likely than African-American students to receive waivers.[2] Waiver rates were also highest among students close to the cutoff score, suggesting some hesitancy to use the strict test-score cutoff. Students' likelihood of being waived was also largely determined by the region of the city in which they lived. Region officers handled principals' waiver requests, and they varied considerably in their willingness to grant those waivers.[3] Because waivers were largely based on factors that could be statistically controlled (e.g., ethnicity or achievement), or were unrelated to dropping out (e.g., region), waivers had almost no influence on the results of the statistical analyses.[4]

[1] Students who will turn 15-years-old by December 1 of the school year are not allowed to enroll in elementary school. Transition centers were later renamed "Academic Preparation Centers."

[2] Roderick, et al. (1999).

[3] See Nagaoka and Roderick (2004), and Roderick, et al. (1999) for a more detailed description of the waiver patterns.

[4] In the companion report, *Ending Social Promotion in Chicago: The effects of retention*, the policy is described primarily for grades three and six. In practice, eighth-grade waivers were implemented slightly differently.

Methods of Analysis

The subjects of the study were seven cohorts of CPS students who were 13-years-old on September 1 of each year, from 1992 to 1998;[1] a total of 171,471 students across all seven cohorts.[2] Three cohorts were classified as prepolicy and four as postpolicy.[3] Each cohort was followed for at least four years, until its members were 17-years-old. All but the last two cohorts were followed for six years, until age 19. Dropout rates at age 17 were used as the primary outcome in this report so that four postpolicy cohorts could be studied. By following four cohorts, policy effects could be examined during the initial stages of implementation and after adjustments were made in later years.[4]

Students were classified as dropouts if they were no longer active CPS students as of September 30 of the school year in which they were 17-years-old (or 19-years-old for analyses that traced students to age 19) and their administrative records indicated that they left school for any of the following reasons: lost—could not be located; transferred to an evening school; lost—undeclared; exited IEP (rather than graduated); dropout self-declared; dropout for absences; did not arrive at school; entered/finished alternative school program (GED or alternative diploma); or no leave code recorded.[5] Students who transferred to another school system were included in the analyses only as long as they were in CPS (e.g., included in the analysis of 15-year-olds if in the system at age 15, and then not included in the analysis for age 16 and 17 if they transferred before reaching age 16). Because students are classified based on their status at a certain point in time, it is possible that students who are classified as dropouts in these analyses did eventually graduate.[6] Likewise, students who are classified as nondropouts may have dropped

---

[1] Students were grouped by age instead of grade because of the very strong relationship between age and dropping out, and to avoid selection bias due to earlier promotion gates.

[2] This is the entire population of CPS students meeting the criteria, not a sample. Therefore, the dropout rates that are presented are actual (not estimated) rates and all differences across cohorts and subgroups are significant. Some specific statistical analyses use only a subset of students (e.g., those with eighth-grade test scores), as described below under "Analysis."

[3] Most postpolicy cohort students were subject to the promotion gate. However, a small number of these students had skipped an elementary grade, so they were not subject to the gate. Likewise, some students in the prepolicy cohorts had been held back in elementary school, so they were subject to the eighth-grade gate. Classifying students by whether they were actually subject to the gate, however, would bias estimates of policy effects, since the lowest-achieving students from prepolicy cohorts would be moved to the postpolicy group, and the highest-achieving students from postpolicy cohorts would be moved to the prepolicy group. By defining cohorts strictly by age, such selection bias is not introduced.

[4] In the first year of the policy, students were held back without any prior changes in programs or instruction, and there were doubts about whether failing students would be retained. Support structures for failing students were introduced after the second year of the policy, with the third cohort through the gate.

[5] The most common dropout codes were "lost—not coming to school" and "dropout for absences." This remained the case across the cohorts, although there was one shift with fewer students coded as "lost—undeclared" and more students coded as "dropout—other" in the later cohorts. The percentage of students coded as dropouts who left for alternative school/GED programs varied between 16 and 19 percent across the cohorts, but there was no systematic increase or decline. Between 5 and 7 percent of the students classified as dropouts had no leave reason recorded in their administrative records. Over 70 percent of these students were classified as dropouts in administrative records one year later. The late classification likely represents a delay in categorizing students for whom status is uncertain. Students under age 16 are over-represented in this group, suggesting a hesitancy to classify students under age 16 as dropouts without verification. Over half of the remaining students with no leave reason recorded are classified as active students or graduates the following year. Others remain unclassified, or are coded as transfer or institutionalized students the following year.

[6] In each cohort, less than 1 percent of students coded as dropouts at age 17 actually re-enrolled and graduated by age 19, while 2 percent had re-enrolled and were still active students.

out and then re-enrolled in school. Dropouts may also have left school more than once, but each student is counted only once in the analyses, based on their status at the age being examined.

Data

Data were collected from three sources: CPS student administrative data files from each fall semester from 1992 to 2002, CPS students' scores on the Iowa Tests of Basic Skills (ITBS) from May 1990 to May 2001, and 1990 census STF3 files on census block groups.

Administrative data. The student administration files provided information on whether a student was enrolled each semester and the reason for leaving for students who were no longer enrolled. Student administration files also provided information on each student's age, gender, race (Asian, African-American, white, Latino, Native American), special-education classification, and home address. Students' home addresses were used to link each student to a particular block group within the city, which could then be linked to census data on the economic conditions of the student's neighborhood. Two indicators of economic conditions were used: concentration of poverty, which is a composite measure based on the percentage of males unemployed in the block group and the percentage of families below the poverty line; and average social status, which is a composite measure based on the percentage of employed persons who are managers, professionals, or executives and the average level of education among adults age 21 or older. Special-education status was used only to improve the estimation of dropout rates when modeling achievement effects. Special-education status was not included in demographic adjustments of dropout rates across cohorts because of the rise in classifications with high-stakes testing.

Student administration files also were used to determine whether students were retained in eighth grade. Students were classified as eighth-grade retainees for one of three reasons: they enrolled in eighth grade in the fall after being enrolled in eighth grade the previous year; they dropped out of school after failing the eighth-grade promotion gate; or they enrolled from an elementary school into a transition center. The second two conditions applied only to students who faced the eighth-grade promotion gate (the postpolicy cohorts).[7]

Test scores. Students' elementary school test scores on the ITBS were used to construct measures of students' achievement. A simple measure of students' achievement by the end of elementary school would be their eighth-grade spring test score. However, a single test score is not a precise measure of true ability (e.g., a student may experience an especially bad or good test day). In addition, postpolicy, eighth-grade test scores may have been affected by the high-stakes testing environment. To more accurately represent students' true achievement levels at the end of eighth grade, a more reliable score was constructed by adjusting students' scores for especially good or bad performance in eighth grade, compared to their performance on previous tests. This was done through two-level hierarchical linear models of students' achievement growth over all of the years that they were enrolled in grades three through eight in CPS.[8] From this model, measures of students' underlying eighth-grade ability were obtained for both reading and math.[9] For the current study, math and reading achievement were combined into one measure of overall achievement by averaging

[7] Students who dropped out after failing the eighth-grade test are included as retainees because they would have faced either retention, summer school, or both as a result of the implementation of high-stakes testing, and these factors may have contributed to their decision to leave school. Students who skipped eighth grade and attended a transition center are included as retainees because, prior to the high-stakes testing policy, they would have moved into high school instead of a transition center.

[8]The models were set up as observations within individuals. Level one fit students' achievement growth trajectories with variables for grade, grade-squared, and repeating a grade. Level two modeled the intercept and each slope without predictor variables, with the slope for repeating a grade fixed across individuals. Separate models were run for each cohort of first-time eighth-grade students in CPS. Before running the HLM, students' test scores were equated through Rasch analysis to remove form and level effects.

the standardized underlying reading and math scores. Test files also were used to determine whether a student was included in test reporting (i.e., included in public reports about CPS student achievement). Students could be excluded from public reporting because of bilingual or special-education exemptions. Postpolicy, students excluded from test reporting were held to different promotion standards than other students.

Analysis

The analysis is based on a matched-samples approach, comparing dropout rates among cohorts of students in CPS before and after the promotion gate. Particular attention is paid to changes in the relationship of dropping out with both retention and achievement after implementation of the promotion gate. These relationships are used to estimate the effect of high-stakes testing on dropout rates among both retained and nonretained students. Hierarchical linear models (HLM) are used to control for changes in demographics and school enrollment patterns in CPS over the period studied.[10] Details on the models and results of analyses are available in the Appendix.

All 171,471 CPS students who were 13-years-old between September 1992 and September 1998 were included in the analyses of dropout-rate trends. However, the analyses that discerned the relationships of retention and achievement with dropping out were limited to those students who were included in testing and enrolled in eighth grade at age 13—about two-thirds of the students, both prepolicy and postpolicy (113,937 students). Confining these particular analyses to this group allowed for the clearest comparison between cohorts for discerning the relationships of achievement and retention to dropping out.[11] All students, however, were used for comparisons of dropout rates across cohorts, with the last two cohorts included only in analyses up to age 17.

---

[9]It was suggested by one reviewer that observed scores should be used, rather than scores obtained from learning-growth models, since the gate should have motivated students to work hard and these learning gains would be "smoothed down" by the model. However, observed scores are not comparable across cohorts because the testing conditions were not the same once the promotion gate was put in place. If there were postpolicy learning gains in eighth grade that were smoothed by the model that were substantive enough to affect dropout rates, then the effects of those gains should be picked up as a main postpolicy effect. In fact, no main postpolicy effect exists. At the other extreme, another reviewer suggested that eighth-grade scores should not be used to model underlying achievement to eliminate any test gains due to the gate from the control for underlying achievement. I believe that the model used here to produce the underlying scores, which includes students' eighth-grade test scores, provides the best estimate of students' actual achievement at the end of eighth grade. Eighth-grade scores that are higher than would be expected given students' previous performance contribute to the latent score, allowing for some incorporation of the motivation effects on student achievement in eighth grade, but they are constrained substantially, minimizing score inflation due mostly to test effects.

[10] For the hierarchical analyses, students are grouped by the first high school or transition center that they attended in CPS. Before ever enrolling in a high school or a transition center, 2.3 percent of the students dropped out. These students were assigned to the high school most frequently attended by other students from their elementary school so that they could be included in the analyses.

[11] Students who were not yet in eighth grade at age 13 are not included because their grade progression and test-taking patterns could not be discretely classified. Students not subject to test reporting are not included because they were not subject to the policy in the same way as other students, and because their participation in testing changed considerably across the years studied. Analyzing only a subgroup of students introduces concern about selection bias. However, the size of this subgroup is similar prepolicy (66 percent) and postpolicy (65 percent), suggesting only minor selection differences. (While increasing numbers of students were excluded from testing with implementation of the promotion gate, most of these late classifications occurred among students already old-for-grade at age 13.) Importantly, it will be shown that the relationships of the control variables with dropout rates are the same for prepolicy and postpolicy analyses groups, suggesting that any slight selection differences did not introduce bias into the statistical analyses. For those comparisons that might be affected by selection bias, adjustments for the potential selection effects are provided.

Ending Social Promotion: A five-year study

Over the past five years, Consortium researchers have been evaluating the effects of the Chicago Public Schools' ending social promotion initiative, which combined sanctions, such as high-stakes testing-based retention and school accountability, with supports, such as optional after-school and mandatory summer-school programs. This report is one in a series that began in 1999 with the publication of *Ending Social Promotion: Results from the first two years*. An update to this report was published in 2000. The first two reports tracked important indicators of student progress, including the proportion of students who met the test-score cutoff at the end of the school year; at the end of the summer; and, for those who were retained, after a second time through the policy. These reports found that after institution of the policy, the proportion of students who had very low test scores (test scores below the promotion cutoff) declined significantly. Declines in the proportion of students with very low test scores were much greater in the sixth and eighth grades than in the third grade. Early data suggested that the Summer Bridge program and the second chance it afforded students to pass the test and avoid retention had been one of the most successful aspects of the policy, accounting for a large proportion of the improvements in passing rates. At the same time, early data suggested that retained students were struggling their second time through the promotion-gate grade and that their postretention achievement growth was no different than previously socially promoted youths.

Subsequent reports and research articles have looked more closely at these initial findings. Two articles, "The Grasshopper and the Ant: Motivational responses of low achieving students to high stakes testing" and "High Stakes Testing in Chicago: Effects on achievement in promotional gate grades," looked both qualitatively and quantitatively at the effects of the policy in shaping students' work effort, experiences in school, and achievement. Another report in the series, *Ending Social Promotion: Results from Summer Bridge* presents the results of a multiyear evaluation of Summer Bridge. This study examined the short- and long-term effects of Summer Bridge on student achievement and looked carefully at the nature and quality of Summer Bridge and students' reports of their experiences. An additional report, *Ending Social Promotion in Chicago: The response of teachers and students*, used Consortium surveys to examine teachers and students' assessments of the impact of the policy. This report investigated whether, after implementation of the policy, teachers spent more time on test preparation and aligned the content of their courses to the ITBS. It also examined trends in students' reports of the academic support they received from teachers and parents, their perceptions of the challenge of their coursework, and their participation in after-school programs.

In this final report, *Ending Social Promotion: Dropout rates in Chicago after implementation of the eighth-grade promotion gate,* and its companion report, *Ending Social Promotion: The effects of retention,* we focus on evaluating the effects of retention on student achievement and progress in school and ultimately on dropout rates. More technical presentations of these results can be found in two accompanying research papers: Allensworth (forthcoming) and Roderick and Nagaoka (forthcoming). In addition, Stone and Engel (forthcoming) examined qualitatively the experience of a group of students who were retained in Chicago, how teachers approached the retention year, and the degree to which retained students experienced different educational supports.

John Booz

CHAPTER 1

What Happened to Dropout Rates after Implementation of the Eighth-Grade Promotion Gate?

Dropout Rates Remained at Prepolicy Levels with the First Two Cohorts through the Gate, Then Declined with the Next Two Cohorts.

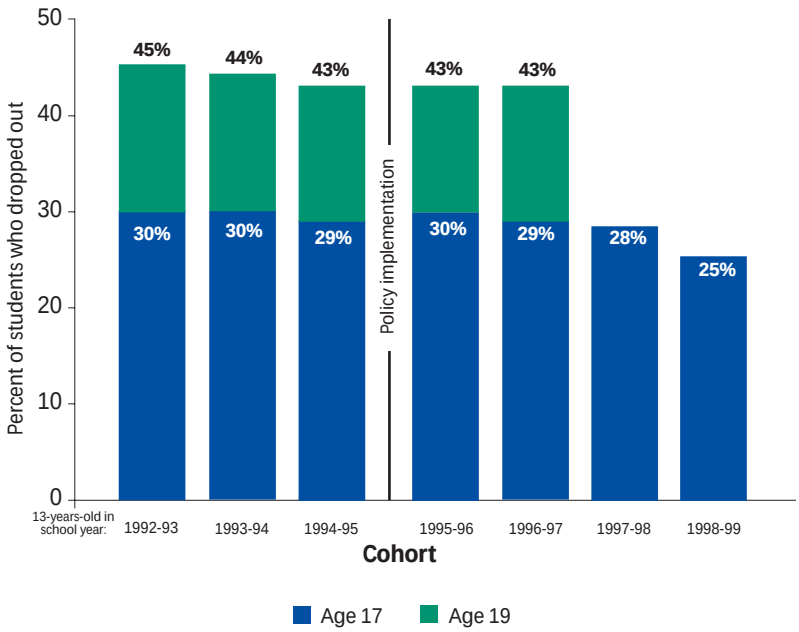
Prior to implementation of the eighth-grade promotion gate, 43 to 45 percent of students dropped out of school by age 19 (see Figure 1). Dropout rates at age 19 remained at about 43 percent with the first two cohorts through the promotion gate. Because the third and fourth cohorts through the gate have not yet reached age 19, their dropout rates at age 19 cannot yet be determined. If students are followed only until age 17, four postpolicy cohorts can be examined, including two cohorts that went through the mature policy. As with dropout rates at age 19, dropout rates at age 17 remained at prepolicy levels for the first two cohorts through the promotion gate—29 to 30 percent. Dropout rates then declined slightly with the third cohort through the gate and fell even further with the fourth cohort through the gate, to 25 percent.

Factors That May Have Affected Dropout Rates

Initial examination of dropout rates by cohort may suggest that implementation of the promotion gate had no negative consequences for students' likelihood of dropping out of school. However, there were a number of changes that occurred simultaneously in Chicago public schools that could have affected dropout rates. High-stakes testing could have had an

Figure 1

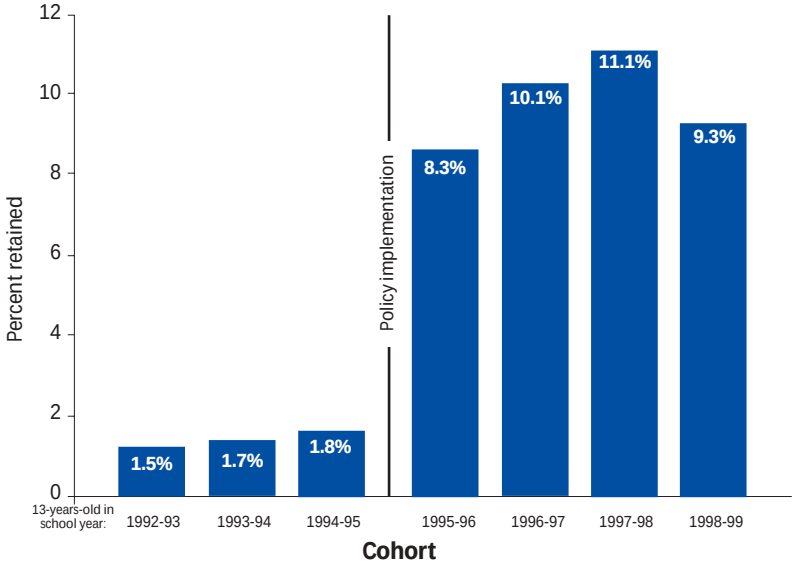
Dropout Rates at Age 17 and 19, by Cohort
All Students



Note: Students were 13-years-old on September 1 of their cohort year. For students at the expected grade at age 13, this was the fall of their eighth-grade year. Students were followed for four years to determine dropout rates by age 17, and for six years to determine rates by age 19.

Figure 2

Eighth-Grade Retention by Cohort
All Students



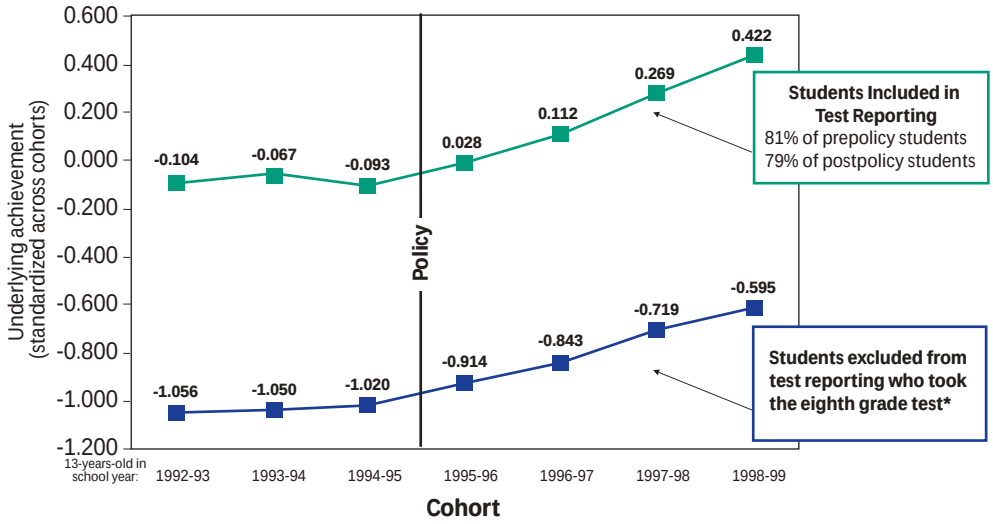
Note: Numbers are calculated from students who had not transferred to another school district by age 17. Students are counted as retained if they enrolled in eighth grade for a second year, dropped out after failing the eighth-grade promotional gate, or skipped from sixth to seventh grade to a transition center.

effect on dropout rates through increased eighth-grade retentions. High-stakes testing could also have affected achievement which could in turn, have affected dropout rates. In addition, there were a number of changes in Chicago's public schools that were unrelated to the promotion gate, such as demographic shifts and efforts at high-school redesign, which could have affected both achievement and dropout rates during this period. All of these factors need to be disentangled in order to understand the effects of the promotion gate on dropout rates.

Retention. Implementation of the eighth-grade promotion gate caused many more eighth-grade students to be held back from entering high school. Before implementation of the gate, just under 2 percent of students repeated eighth grade each year (see Figure 2). After the policy was initiated, retention rates surged in the first year to 8.3 percent. Because the cutoff scores for promotion were set higher each year, retention rates continued to rise with the next two cohorts, peaking at 11.1 percent. Retention rates declined with the fourth cohort to 9.3 percent because achievement increased substantially, and because more students were promoted despite not meeting the cutoff or were moved into special education and excluded from test reporting.[1]

Figure 3

Average Achievement by Cohort



* Students excluded from test reporting, because of special education or bilingual exemptions, might take the test and not have their scores included in public reporting statistics, or they might not take the test at all. Students who did not take the test cannot be added into cohort averages. Postpolicy, many more excluded students took the test than in the prepolicy cohorts. Therefore, a larger percentage of students in special education and bilingual programs are added into the postpolicy averages of students excluded from test reporting. It is surprising that average test scores among students excluded from test reporting improved, since the inclusion of more of the students with bilingual/special education exemptions would have been expected to depress average scores. In addition, about 2 percent more students were excluded from test reporting postpolicy because of higher rates of classification into special education. This may have slightly elevated the postpolicy scores for both included and excluded students, especially in the 1998-99 cohort. Excluding a similar proportion of the lowest-performing students in the 1992-93 cohort would raise that average from -.104 to -.034.

Student Achievement. Achievement improved substantially with the postpolicy cohorts. Figure 3 displays average underlying eighth-grade achievement across the seven cohorts. Some students in the special-education and bilingual programs were not required to take the test. If they took the test, their scores were excluded from public reporting (i.e., not included in calculations of school statistics). Postpolicy, these students were not held to the promotion standards in the same way as other students.[2] Achievement is represented separately for students excluded from test reporting in Figure 3 because they were not subject to the same testing environment as other students, and because estimates of their average achievement could be biased by a number of factors.[3]

By the 1998 cohort, achievement was more than one-half of a standard deviation higher than it was in prepolicy cohorts, among included students. In other words, the test score of an included student who was at the 50th percentile in the 1998 cohort would have been at the 70th percentile in the 1992 cohort.[4] Achievement was also much higher among students whose scores were excluded from test reporting, even though a much larger percentage of these students were tested and were used in calculating the averages for excluded students, postpolicy (see Figure 3).

The rise in student achievement is consistent with the policy effects anticipated by proponents of high-stakes testing. However, the degree to which the policy acted as a catalyst for rising achievement is debatable. Much of the rise in achievement could be attributed to other factors, described below under "Factors unrelated to high-stakes testing." This study does not attempt to discern the degree to which improvements in achievement were due to the promotion gate versus other factors. Regardless of its source, rising achievement should have reduced dropout rates during the postpolicy period.

Factors unrelated to high-stakes testing. During this period, a number of changes occurred in CPS that were unrelated to high-stakes testing and that may have affected dropout rates and student achievement.

Some of these factors can be measured and statistically controlled. These include demographic shifts in the population of students served by CPS, as well as overall economic improvements across the city.[5] In addition, a number of new high schools opened near the end of the decade, and these schools tended to have lower dropout rates than the overall system. The availability of these schools should have helped to lower dropout rates.[6] When demographic changes and school enrollment patterns are statistically controlled, the decline in dropout rates appears slightly smaller—adjusted from a decline of 4.5 percent with the 1998-99 cohort to a decline of 4.1 percent.[7]

The opening of new schools and the eighth-grade promotion gate were two components of district efforts at high-school redesign. These components aimed at improving the quality of students entering the high schools. Students with high elementary-school achievement were slightly more likely to remain in CPS for high school after the new schools opened, and this should have reduced overall dropout rates and achievement slightly. Other components of high-school redesign also may have affected dropout rates.[8] Beginning in 1997, CPS mandated that students take more challenging coursework in high school, and restructured schools to improve personalization to help students meet the demands of these higher standards (e.g., introducing advisories, junior academies, and schools within schools). While there is no evidence that school restructuring had a substantial effect, students were more likely to accumulate the credits needed for graduation after they were required to do so.[9] This likely helped to improve graduation rates. CPS also began focusing resources and penalties on very-low-performing schools. School-level accountability was enacted in the 1995-96 school year, in which schools faced probation, remediation, and reconstitution if too few of their students' test scores met national norms. Over half of the high schools in CPS experienced these consequences, and while these consequences brought additional resources into struggling high schools, they also brought substantial stress and uncertainty.[10] Students in the lowest-achieving schools may have been affected by these consequences for their schools.

There also were other changes in CPS district policy that occurred before efforts at high-school redesign and implementation of the promotion gates that should have affected student achievement and dropout rates during the latter part of the 1990s. In the early 1990s, Chicago's elementary schools underwent substantial changes due to unprecedented decentralization reform enacted in 1988. Following the introduction of the reform act, in the early 1990s, most elementary schools showed significant improvements on the ITBS. Students who were in the primary grades during this time of improving test scores would have moved into eighth grade during the latter part of the decade, around the same time that the promotion gates were put into place. Therefore, at least some of the rise in achievement and the decline in dropout rates are likely due to earlier school reforms.[11]

. . . at least some of the rise in achievement and the decline in dropout rates are likely due to earlier school reforms.

Furthermore, in 1995, Mayor Daley took over control of the schools and appointed a CEO of schools. For the first time in years, there was stability in CPS leadership—conflict with the teachers union subsided, and there was an influx of money for substantial capital improvements. These changes also should have improved conditions for student learning.

Finally, broader societal changes may have reduced dropout rates. Chicago was typical of the national trend in declining births to teenage mothers over this period.[12] Additionally, the introduction of welfare reforms that required school attendance should have helped to reduce dropout rates among teenage mothers.

These societal factors, coupled with the school-system changes discussed earlier, require caution when attributing dropout-rate and achievement improvements to the implementation of the eighth-grade promotion gate. While the overall trends in dropout rates displayed in Figure 1 initially suggest that the policy caused no adverse effects, these effects may be concealed by the numerous other changes that were occurring in the schools. There was an array of factors that potentially had some effect on dropout rates during the 1990s. To understand the total effect of the policy, this analysis picks apart the changes in dropout rates that could be specifically attributed to the policy: the effect of rising retention and the effect of rising achievement.

John Booz

CHAPTER 2

Did Retention at the Gate Affect Students' Likelihood of Dropping Out?

Retention at the promotion gate increased the likelihood of dropping out by age 17 by about 8 percentage points (from 31 percent to 39 percent for a typical retained student), and increased the likelihood of dropping out by age 19 by about 13 percentage points (from 44 to 57 percent).

Teacher-initiated retention has been tied repeatedly to a high likelihood of dropping out, but it is not known if district-mandated retention affects students' likelihood of dropping out. If experiencing retention at the promotion gate increased a student's likelihood of dropping out, we would expect to see more low-achieving students drop out after high-stakes testing was introduced, since so few of these students were retained prior to the policy and so many were retained after the policy. We would not expect, however, to see any difference in the pre- to postpolicy dropout rates of students whose achievement suggests they were not at risk of retention.

Yet, the simultaneous increases in achievement and retention rates confound the analysis of the effects of either on dropout rates. If achievement had not increased dramatically with the policy, it would be possible to calculate the retention effect by comparing very-low-achieving postpolicy students to those with similar achievement prepolicy. However, because of the increase in achievement, such groupings would not be equivalent. Scores that would place students in the bottom quartile of the 1992 cohort would place them in the bottom 5 percent of the 1998 cohort—a very different group of students. It is also not possible to look at the effect of retention just by controlling for student achievement. After policy implementation, retention and achievement became strongly correlated.[1] Therefore, making a direct comparison of each to dropout rates becomes problematic.[2]

However, among nonretained eighth-grade students, the relationship between *underlying* achievement[3] and dropping out was virtually the same prepolicy and postpolicy.[4] It was the same because there was no retention effect among nonretained students, beyond any motivating effect on achievement, and because underlying test scores do not contain the testing effects found in raw scores.[5] In other words, because these students were not retained by the gate, the relationship between underlying achievement and dropping out remained unchanged postpolicy. This does not mean that the gate had no effect on nonretained students. The gate could have affected their motivation, and thus, their achievement, resulting in a lower likelihood of dropping out. This only means that the *relationship* between underlying achievement and dropout rates did not change. Therefore, by using information from prepolicy students together with information from postpolicy students, achievement can be controlled across cohorts and the postpolicy retention effect can be properly estimated.[6]

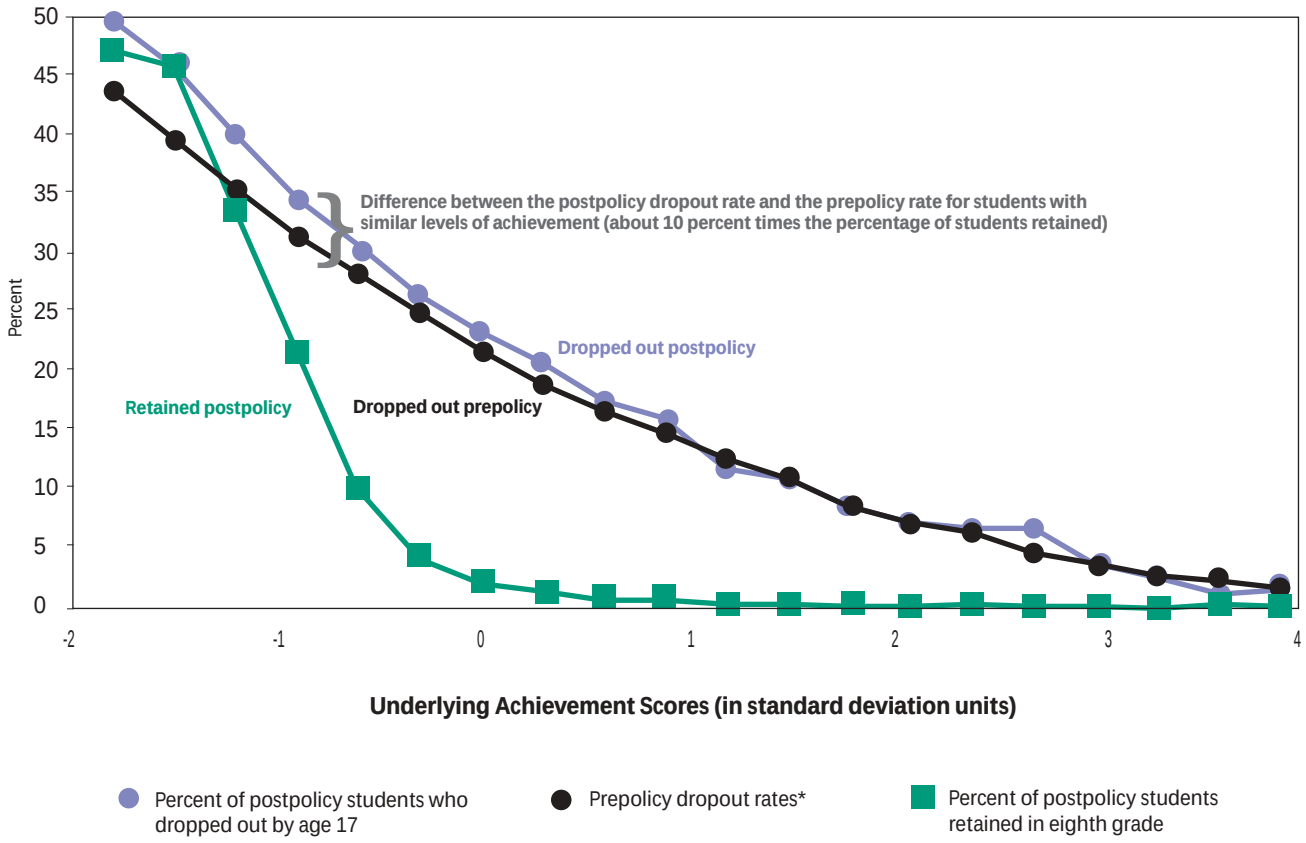
Figure 4 shows the percentage of students who dropped out by their underlying achievement level for cohorts prior to and after the institution of the eighth-grade gate, as well as the percentage of students at each achievement level retained in postpolicy cohorts. This graph provides strong evidence that retention under Chicago's high-stakes testing policy affected students' likelihood of dropping out. As shown, virtually no students with above-average, underlying-eighth-grade achievement (those with a score above one) were retained in eighth grade. For these students, there is very little difference in prepolicy versus postpolicy dropout rates—the lines representing dropout rates for postpolicy and prepolicy students are almost on top of each other. This is not true, however, among low-achieving students. As seen in Figure 4, the lower the students' test scores, the higher the probability of retention and the greater the discrepancy between prepolicy and postpolicy dropout rates. The discrepancy between prepolicy and postpolicy rates gets larger as the probability of retention gets larger; for each 10 percent increase in retention, the discrepancy between prepolicy and postpolicy dropout rates is about 1 percent larger. For example, almost half of the students with the lowest underlying-eighth-grade test scores (two standard deviations below the mean) were retained postpolicy, and almost 50 percent of these students dropped out by age 17, compared to 44 percent of students with the same underlying achievement prior to the policy. Note, however, that while this provides substantial evidence of a retention effect on dropout rates, it does not show the entire policy effect since the distribution of students' latent achievement was higher postpolicy than prepolicy. That is, while very-low-achieving students' likelihood of dropping out was higher postpolicy than prepolicy, students were less likely to have very low achievement.

Statistical models that estimate retention effects using the prepolicy achievement-dropout relationship show a significant postpolicy eighth-grade retention effect on dropping out of 9.9 percentage points; postpolicy, 31.1 percent of students who had achievement typical for retained students, but were not retained, dropped out of school by age 17.[7] In comparison, 41.0 percent of retained students with the same demographic and achievement characteristics dropped out. (See the Appendix for details.)

Prepolicy, the relationship of eighth-grade retention to dropping out was much stronger. The same statistical models that estimated the postpolicy effect show that retained students in prepolicy cohorts had dropout rates that were 25.3 percentage points larger than nonretained prepolicy students with similar demographic and achievement characteristics (55.7 percent, compared to 30.4 percent).[8] However, prepolicy students retained in eighth grade were a very select group of students (less than 2 percent of the total). Postpolicy students retained in eighth grade were less likely to drop out than prepolicy students retained in eighth grade, but many more postpolicy students were retained in eighth grade (see Figure 2). Therefore, while the postpolicy retention effect on dropout rates seems to be much smaller than the prepolicy effect—either

Figure 4

Retention and Dropout Rates at Age 17 by Achievement, Postpolicy Students Included in Test Reporting



* The prepolicy line is a regression line showing the prepolicy relationship between achievement and dropping out. The prepolicy dropout rates are adjusted for changes in the demographic composition of students postpolicy by applying regression coefficients from models with only prepolicy students to pospolicy students, and plotting their probability of dropping out based on the prepolicy relationships. Only students included in test reporting and in 8th grade at age 13 were used for these calculations to improve precision in estimation. Similar trends are seen if students old-for-grade at age 13 are included, but there is more uncertainty in measurement.

because of the different context of retention or because spurious factors are minimized postpolicy, or both—its total impact on systemwide dropout rates was larger than prepolicy retention.

While the vast majority of postpolicy students were retained because of low test scores, some were retained for other reasons, as in prepolicy years. These students' potential for dropping out was likely more similar to that of prepolicy students held back in eighth grade. Therefore, the postpolicy retention effect must be adjusted to estimate the effect of retention at the promotion gate. If it is assumed that the same percentage of postpolicy students would have been held back without high-stakes testing as were held back in prepolicy years,[9] and that the dropout rate for these students was the same as prepolicy (because they would have been retained for similar reasons), then retention from high-stakes testing is estimated to have increased the probability of dropping out from 30.8 percent to 38.8 percent for a typical retained student.[10]

While 8 percentage points may seem small, it is a 26 percent increase in the likelihood of dropping out by age 17. Furthermore, if students are followed until age 19, the effect is proportionately larger. Modeling dropout rates by age 19, without the last two postpolicy cohorts, produces an estimate that dropout rates among students retained by the promotion gate were 13 percentage points higher than among similar nonretained students—57 percent, compared to 44 percent, which is a 29 percent increase in the likelihood of dropping out.

John Booz

CHAPTER 3

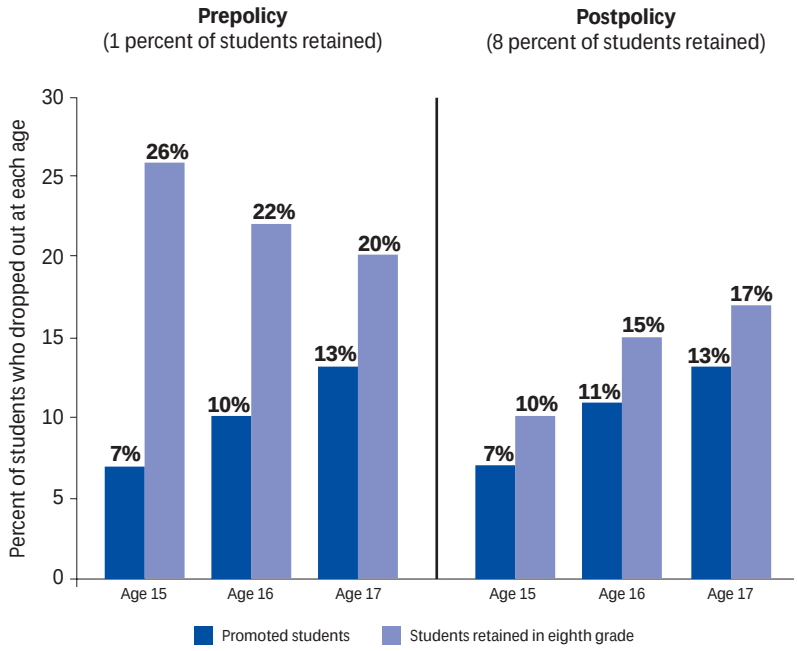
Did Retention at the Gate Lead Students to Drop Out at an Earlier Age Than They Would Have without the Gate?

Students held back by the promotion gate were not especially likely to drop out at early ages; instead, their likelihood of dropping out was elevated for the remainder of their time in school.

To determine whether retention from high-stakes testing affected the timing of dropping out, as well as the likelihood of dropping out, a specialized statistical analysis was performed to predict the hazard of dropping out at each age, from 15 to 17.[1] This analysis adjusted dropout rates for changes in achievement, demographic characteristics, and school enrollment, so that the retention-dropout relationship could be compared between prepolicy and postpolicy students (see the Appendix for details). Figure 5 displays dropout rates for students with average demographic and achievement characteristics, calculated from the statistical model. The dropout rates are smaller than those shown earlier because these rates are the percentage of students who dropped out *at* each age, instead of by age 17. For example, among nonretained students, 7 percent dropped out by age 15, 10 to 11 percent of the remaining students dropped out at age 16, and 13 percent of the remaining students dropped out at age 17.[2]

Figure 5

Dropout Rates at Each Age by Eighth-Grade Retention: Prepolicy and Postpolicy Students Included in Test Reporting



Note: Dropout rates have been calculated from event history models with variables controlling for student race, gender, SES, and achievement. Only students included in test reporting who were in eighth grade at age 13 were included in the estimations. Models with old-for-grade students show the same trend.

John Booz

Consistent with studies of teacher-initiated retention, prepolicy students who were retained in eighth grade were not only much more likely to drop out of school than other students, but they were also more likely to drop out early. Only about 1 percent of prepolicy students included in test reporting and in eighth grade at age 13 were held back in eighth grade. But 26 percent of those students dropped out by age 15, compared to only 7 percent of nonretained prepolicy students. Prepolicy, the relationship between eighth-grade retention and the hazard of dropping out was significantly smaller at older ages. In contrast, postpolicy, eighth-grade retention was not associated with a greater likelihood of dropping out at age 15 or 16 than it was at age 17. Instead, postpolicy students' likelihood of dropping out was elevated throughout their remaining years in school. At each age, they were about 3 to 4 percent more likely to drop out than other postpolicy students with similar demographic and achievement characteristics. It should be noted, however, that while postpolicy retained students did not drop out at earlier ages, they did drop out at earlier grades. Because their entry into high school was delayed, students had less time to accumulate high-school credits before they dropped out.[3]

CHAPTER 4

Did Simultaneous Improvements in Student Achievement Lead More Students to Stay in School?

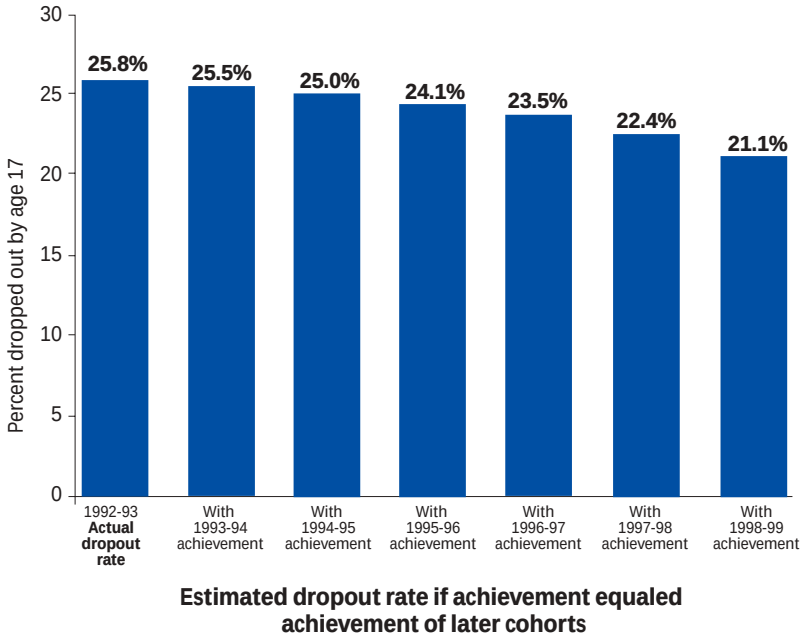
Rising achievement was associated with a decline in dropout rates of 1.3 percentage points averaged over the four postpolicy cohorts, and almost 4 percentage points by the fourth postpolicy cohort. This decline in dropout rates among the 90 percent of students not retained by the gate more than balanced the increase in dropout rates among the 10 percent of students retained.

Retention rates rose with the policy, and postpolicy retention appears to have increased students' risk of dropping out. Yet, overall dropout rates did not rise with implementation of the policy. The primary reason for this was a countervailing trend of increasing achievement. While there has been debate about whether the rise in achievement seen after 1995 reflects real learning versus testing effects, the rise in the *underlying* achievement measure appears substantive, given that its relationship to dropout rates did not change after policy implementation. Whether it is *attributable* to the policy is a different question.

Figure 6 demonstrates the relationship between achievement and dropout rates by taking the 1992 cohort (both retained and promoted students), and estimating what that cohort's dropout rate would have been had its achievement been the same as seen in subsequent cohorts.[1] The effect of rising achievement on dropout rates is not large—about 1.3 percent fewer students dropped out, on average, over the four postpolicy cohorts than would be expected if achievement remained the same. The 1998

Figure 6

1992-93 Cohort Dropout Rate and Expected Rate if Achievement Equaled the Achievement of Subsequent Cohorts
Students Included in Test Reporting and in Eighth Grade at Age 13



Note: There may be selection bias in these calculations. They are based on the sample of students included in testing and in eighth grade at age 13. Four percent more students were excluded from testing or were old-for-grade with the 1998-99 cohort. This is equivalent to 6 percent of the 67 percent of included on-grade students in the 1992-93 cohort. If the bottom 6 percent of scores are removed from the 1992-93 cohort average, the dropout rate would be 24 percent for the 1992-93 cohort, which would be the maximum possible effect of the bias.

cohort, which had the highest achievement, should have had dropout rates that were 3 to 4 percentage points higher if achievement had remained at 1992 levels. However, while the achievement effect on dropout rates was not large, the slight decline in dropout rates among the 90 percent of students who were not retained by the gate more than balanced the increase in dropout rates among the 10 percent of retained students. As a result, overall dropout rates did not rise with implementation of the promotion gate.

What Was the Net Effect of Rising Retention Rates and Rising Achievement for Retained and Nonretained Students?

Improvements in achievement were seen across all students, including those with low achievement who were at risk of retention by the promotion gate. Among retained students, the increase in achievement should have lowered dropout rates by as much as 3.3 percentage points across four cohorts, from 44.0 percent to 40.7 percent.[2] Improvements in achievement were largest in the later cohorts, so the estimates of achievement effects on dropout rates among retained students range from a decline of 1.2 percentage points with the 1995-96 cohort, to as much as 6.1 percentage points with the 1998-99 cohort. These estimates may be high because of selection bias, however, since more students were excluded from test reporting postpolicy. If adjusted for the maximum potential selection effect, the estimate for the 1998-99 cohort drops to 3.5 percentage points.[3] With either estimate, for students retained under high-stakes testing, the adverse effects on dropping out from retention (averaging about 8 percentage points of an increased probability of dropping out) outweighed any potentially beneficial impact of improving achievement. Among promoted students, postpolicy improvements in achievement should have resulted in a decline in dropout rates of 1.8 percentage points (24.3 percent compared to 22.4 percent), averaged across the postpolicy cohorts. The improvements in achievement were largest in the later cohorts, so estimates of the achievement effect on dropout rates among promoted students range from a decline of 0.7 percentage points with the 1995-96 cohort, to as much as 3.4 percentage points with the 1998-99 cohort.[4]

CHAPTER 5

Were Dropout Trends Different for Subgroups of Students?

Students already old-for-grade at age 13 were the most adversely affected by the policy—one-quarter were held back by the gate, and 78 percent of these retained students dropped out of school by age 19. Racial disparities in dropout rates also grew postpolicy, and this was partially attributable to the promotion gate.

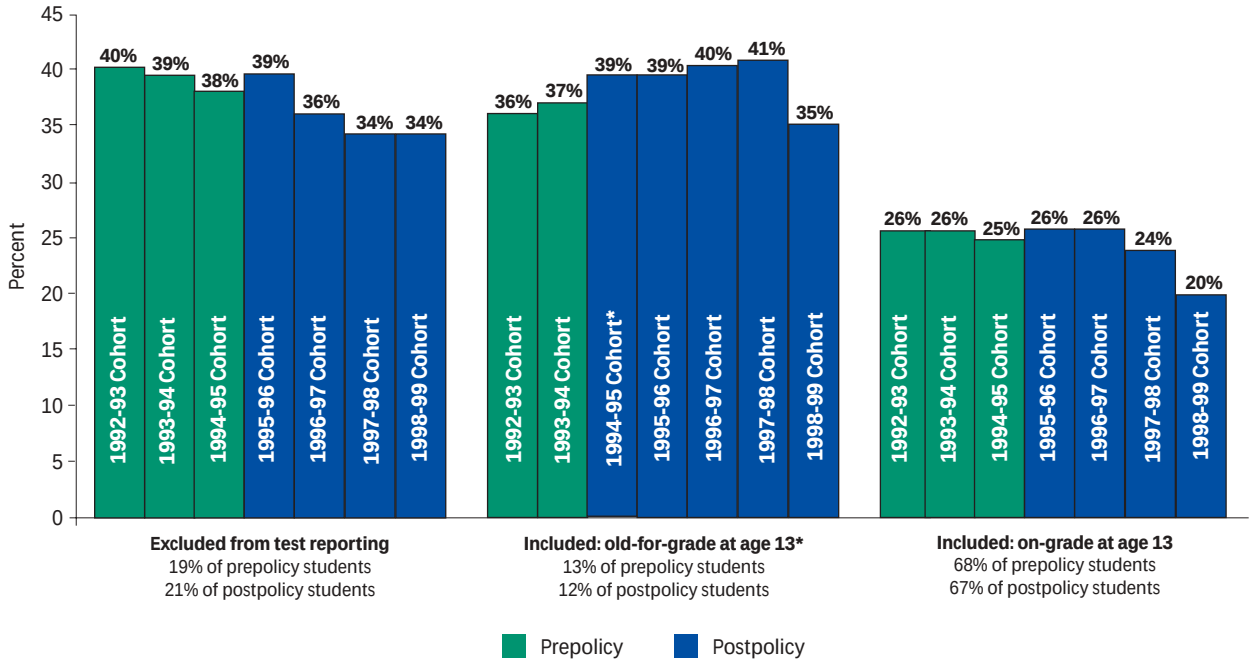
Not all students faced the same likelihood of being retained by the promotion gate, and improvements in achievement were not the same among all groups. Therefore, trends in dropout rates varied across subgroups of students. Only low-achieving students who were included in test reporting were at substantial risk of being held back by the policy, and for these students the odds of dropping out grew. For other students, the likelihood of dropping out decreased, on average, as achievement improved.

Students Old-For-Grade before Reaching Eighth Grade

In each cohort, about 15 percent of the students included in test reporting had not yet made it to eighth grade by the time they were 13-years-old. Most of these students had been retained in grade at an earlier point in school, so they were in sixth or seventh grade at age 13. Postpolicy, these students were much more likely to be retained by the promotion gate than students who had reached eighth grade by age 13. They were at high risk of failing the gate because their achievement, on average, was very low. In addition, some students already old-for-grade at age 13 never even had the opportunity to pass the promotion gate because they were moved into transition centers before reaching the eighth grade. These students were too old to remain in elementary school and they were no longer allowed to

Figure 7

Dropout Rates at Age 17
by Inclusion in Test Reporting and Grade at Age 13



* All of the old-for-grade students in the 1994-95 cohort were subject to the eighth-grade promotion gate because they had not yet made it to eighth grade before the 1995-96 school year. Some of the members of the 1993-94 cohort were also affected by the gate.

move directly into high school. In the transition centers, students received intensive remedial study of reading and math, but took few courses that would allow them to accumulate credits for high school graduation. Prior to the promotion-gate policy, students who were too old to remain in elementary school moved directly into high school. Across the four postpolicy cohorts, 24 percent of students old-for-grade at age 13 were in some way delayed from entering high school because of the promotion gate—either they were retained in eighth grade, or they enrolled in a transition center instead of a high school. In comparison, less than 1 percent of the old-for-grade students in the 1992-93 cohort were held back in eighth grade.

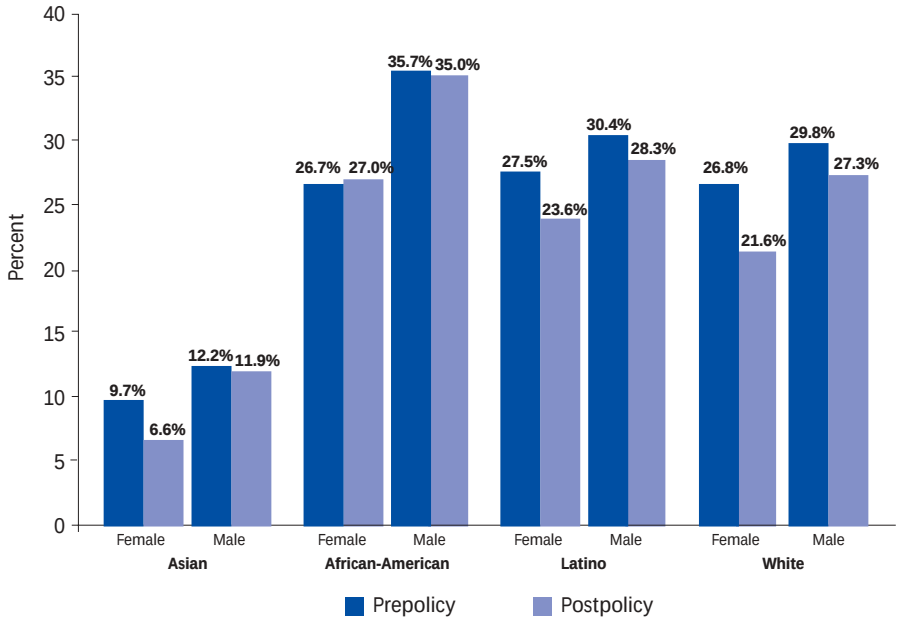
Because postpolicy retention rates were very high among students already old-for-grade at age 13 (24 percent, compared to 8 percent of on-grade students), their dropout rates were higher postpolicy than prepolicy, despite higher achievement, on average, in postpolicy cohorts.[1] Figure 7 shows dropout rates by cohort for students already old-for-grade at age 13. Among these students, dropout rates rose steadily across cohorts and then declined to prepolicy levels with the 1998-99 cohort. This decline may be a selection artifact as more students were excluded from test reporting in this cohort than in others. This cohort was the first to face the sixth-grade promotion gate, and there is evidence that some students who failed the sixth-grade gate were reclassified into special education and excluded from test reporting by the spring of their eighth-grade year.[2]

The increase in the likelihood of dropping out among students already old-for-grade at age 13 is especially distressing because these students were already very likely to drop out before being held back an additional year. By age 19, 78 percent of the students old-for-grade at age 13 who were retained by the promotion gate had dropped out of school.

Students Excluded from Test Reporting

Figure 8

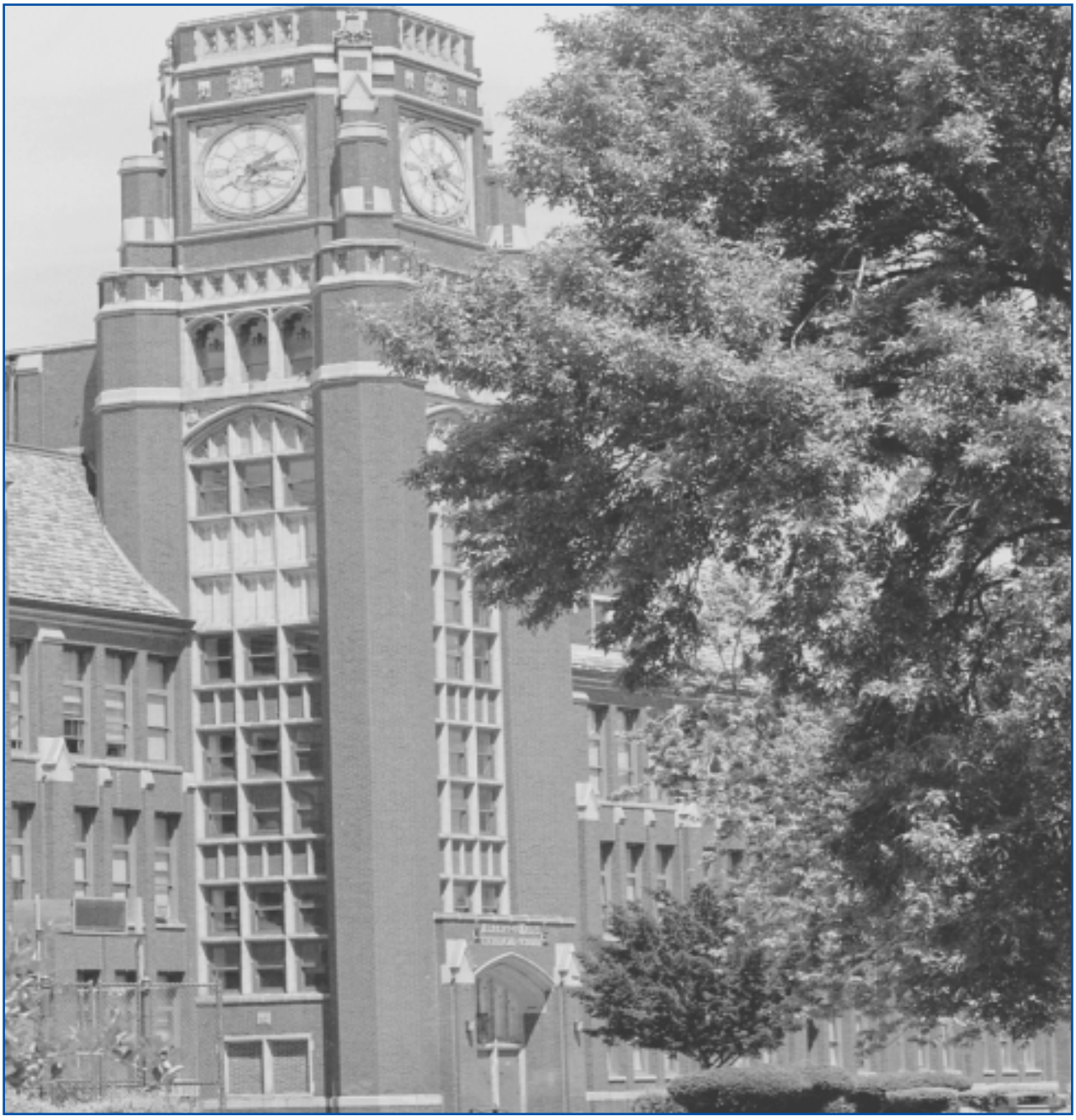
Pre- and PostPolicy Dropout Rates by Age 17 by Race and Gender
All Students



Students excluded from test reporting were also more likely to be held back in eighth grade after implementation of the promotion gate. However, this increase was much smaller than among included students—increasing from 5 to 7 percent in the first year of the policy, and then declining each year to below prepolicy levels by the fourth year of the policy.[3] At the same time, while test scores for excluded students cannot be precisely compared across cohorts because of substantial shifts in test taking among excluded students, the available evidence suggests that achievement improved among excluded students as it did among students included in test reporting. As a result, after an initial increase in dropout rates among excluded students, dropout rates among excluded students declined with each postpolicy cohort (see Figure 7). The decline was smallest with the 1998-99 cohort, but this may be an artifact of the selection effects discussed above; more students with extremely low achievement were excluded from test reporting in this cohort because of higher placement into special education after the sixth-grade promotion gate.

Differences by Race and Gender

Prior to implementation of the promotion gate, African-American students were more likely to drop out than students of other races/ethnicities in Chicago, and boys were more likely to drop out than girls. These differences grew larger postpolicy. Figure 8 shows pre- and postpolicy dropout rates by race and gender, without any adjustments for achievement, retention, school enrollment, or demographic characteristics. While the majority of students in CPS are African-American, the postpolicy decline in dropout rates only occurred among students of other racial-ethnic groups. Among non-African-American students, girls showed more improvements in dropout rates than boys. Nagaoka and Roderick (2004) provide an analysis of retention rates by race and gender, and these retention rate differences explain some of the variability in dropout-rate trends, particularly differences by race.[4] There were a number of other factors that also likely contributed to the differences by race and gender, including changes in school-level accountability, which primarily affected predominantly African-American high schools, and changes in welfare policy and teenage birthrates, which primarily affected girls.[5] While the evidence does not suggest that the promotion-gate policy was responsible for all of the race and gender differences in dropout-rate trends, the policy did exacerbate the differences that existed prepolicy.



INTERPRETIVE SUMMARY

The implementation of an eighth-grade promotion gate in Chicago was not accompanied by a massive rise in overall dropout rates as feared by opponents to the policy. However, overall dropout rates also did not decline substantially despite considerable increases in student achievement. Retention at the promotion gate increased low-achieving students' likelihood of dropping out. But a slight decline in dropout rates among the vast majority of students who were not retained by the gate balanced the adverse effect of retention so that, overall, dropout rates did not rise.

The effect of high-stakes testing-based retention on students' risk of dropping out contrasts sharply with the results from teacher-initiated retention studies. While prepolicy retention showed a very strong association with dropping out similar to that found in studies on teacher-initiated retention, the effect of retention due to the gate was one-third as much. Likewise, while prepolicy eighth-grade retention was found to be strongly associated with dropping out early, as has been found in studies of teacher-initiated retention, retention resulting from the promotion gate was not. These discrepancies may result from the different context of retention under high-stakes testing. For example, because large numbers of students were retained, some of the psychological effects of retention may have been minimized. Teachers also may have been more responsive to the educational needs of large groups of retained students than they would have

been to individuals. Alternatively, the discrepancy may result because postpolicy retention was primarily based on one criterion—low achievement, and that factor could be statistically controlled when measuring the retention effect. Teacher-initiated retention often occurs because a student is behaving in ways indicative that they are already disengaging from school (e.g., poor attendance or declining academic performance).[1] The prepolicy relationship, and those found in studies of teacher-initiated retention, are likely inflated by spurious factors.

. . . for some students [the retention policy] was the deciding factor that eventually led them to drop out.

Still, while this analysis found smaller retention effects than have been found with teacher-initiated retention, and the majority of students were not hurt by the policy, for some students it was the deciding factor that eventually led them to drop out. It likely had an immediate effect on students who were already fairly disengaged from school,[2] and subsequent effects on other students as they struggled to graduate in later years. CPS high school course failure rates are high,[3] so a year delay makes it even more difficult for students to graduate by age 18 or even 19. As students weigh the likelihood of graduating with other factors (e.g., if they view the school environment as threatening[4] or adult concerns become pressing[5]) it may not seem worthwhile to continue. In addition, because retention delayed students' entry into high school, students who were held back by the gate had less time to accumulate high school credits before dropping out. Therefore, although not studied here, recovery may be more difficult for those postpolicy dropouts who eventually attempt to return to school.

Using the threat of retention to motivate student learning has negative consequences for students who do not pass the promotion criteria. It penalizes the lowest-achieving students in an attempt to motivate all students to perform better. And is the cost equal to the gain? If all of the growth in achievement is attributed to high-stakes testing, then the gains balance the costs—overall dropout rates declined despite higher rates of retention. However, even if all of the achievement growth were attributed to the promotion gate, its effect on dropout rates was fairly modest. While the rise in eighth-grade achievement was very large, this rise in achievement was associated with a decline in dropout rates of only 1.3 percentage points, averaged over the four postpolicy cohorts, or as much as 4 percent by the final cohort. While the relationship between achievement and dropping out is strong, it is far from deterministic. In Chicago, even high-achieving students drop out at high rates.[6] Improving achievement alone, without accompanying school reform will not drastically lower overall dropout rates. Furthermore, if rising achievement resulted mostly from other policies or practices, then the net effect of high-stakes testing on dropout rates was adverse.

It is very likely that some or all of the gains in students' test scores were due to factors other than high-stakes testing. Estimates of achievement used here were adjusted for one-time testing effects, such as motivation to perform well at the gate.[7] However, estimates were not adjusted for real improvements in student learning that occurred because of policies other than high-stakes testing. A number of other policies could have affected student achievement over this period, including Chicago's initiative for school accountability that coincided with implementation of the promotion gates and school improvements that occurred with the decentralization of Chicago's schools in 1988 and mayoral control in 1995. The rise in achievement among students who did not face the promotion-gate cutoff (i.e., those excluded from test reporting) itself suggests that achievement likely rose for reasons other than the threat of retention. Furthermore, in Chicago, high-stakes testing did not rely on the threat of retention alone to improve achievement. Beginning in the 1996-97 school year, and expanding each of the fol-

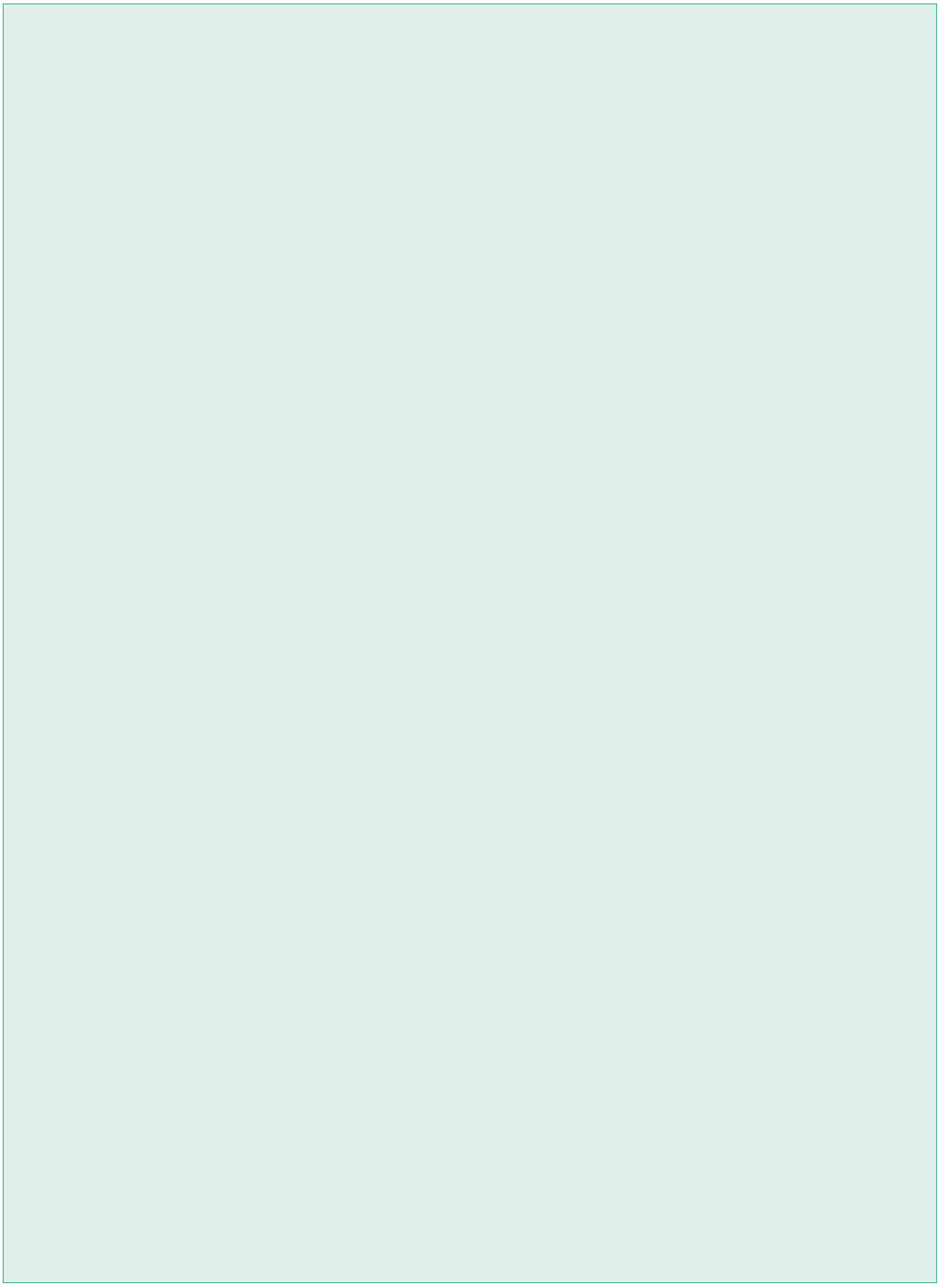
lowing years, Chicago offered extensive support to low-achieving students through after-school and summer-school programs. Only those cohorts that had access to these programs, the third and fourth cohorts through the gate, showed a decline in dropout rates.[8]

Regardless of the source of the achievement effect, for students retained by the promotion gate, retention had a larger consequence than improving achievement on their odds of remaining in school. As a result, students with very low achievement were even less likely to graduate than before the policy was implemented. Students already old-for-grade before reaching the eighth-grade gate were particularly likely to be affected adversely. Racial gaps in school completion also grew after implementation of the gate as dropout rates improved among all but African-American students. Even if all of the increases seen in achievement were attributed to the promotion gate, it could not be viewed as an effective tool for reducing dropout rates among those groups of students most likely to drop out.

In addition, the very high dropout rates among students already old-for-grade who failed the eighth-grade test suggest that the combined effects of the gates at third, sixth, and eighth grade may be more adverse than that of the single gate in the eighth grade. More time is needed before any effects of the third- and sixth-grade promotion gates on dropout rates can be discerned, since students subject to these gates are still too young to be likely to drop out. However, students who failed the gates in the early grades are at high risk to fail subsequent gates. Nagaoka and Roderick (2004) found no evidence that retention had a beneficial effect on their achievement growth, yet the multiple delays in grade progression will make it difficult for these students to eventually graduate.

While overall dropout rates decreased with the promotion gate, the gate had adverse effects on the most vulnerable students, and, at best, only modest beneficial effects on other students' likelihood of completing high school. This is not strong support for the policy.

Chicago has begun to address some of the problems of the eighth-grade gate. In the 2003-04 school year, CPS moved some of the former transition centers into high schools, as schools-within-schools, so that students who are more than two years behind grade level can begin to accumulate credits toward graduation, and so that they do not have to make multiple school transitions. This seems like a promising step, although it affects only students who were already old-for-grade before failing the gate. A similar strategy might be considered for students who were not already behind in school so that they are not pushed behind by repeating the eighth grade. In their report on achievement effects of promotion gates, Nagaoka and Roderick (2004) discuss some alternatives to promotion gates, including focused interventions to substitute for retention,[9] and earlier diagnosis and intervention of learning problems. Until it is known that the threat of retention itself was responsible for improvements in student motivation and achievement, rather than other aspects of the policy and other changes in system policies, such alternatives should be considered.



APPENDIX

Statistical Models and Outcomes

The HLM models of retention effects are shown below and in Table 1. While not shown here, the achievement coefficients are the same as from models of pre- and post-policy *nonretained* students with only demographic and achievement variables. These preliminary models provided evidence that the relationship between latent achievement and dropout was the same across cohorts, once retained students were removed. Analysis of residuals from the preliminary models showed no difference in the underlying achievement-dropout relationship between prepolicy and postpolicy students, even among students with very low achievement, among students who were not retained in eighth grade. This provided substantial evidence that the relationship between underlying achievement and dropout was not affected by the policy.[1] Coefficients from the model displayed in Table 1 were compared to those of the preliminary models, and found to be the same, providing confirmation that the retention and achievement coefficients were not misestimated because of collinearity between those variables postpolicy. Further details on these analyses are available from the Consortium on Chicago School Research at: www.consortium-chicago.org.

Analyses were run using HGLM.[2] The model presented in Table 1 is:

$$
\begin{aligned}
\eta(_{dropout})_{ij} &= \beta_{0j} + \beta_{1j}(Male)_{ij} + \beta_{2j}(Native\ American)_{ij} + \beta_{3j}(Asian)_{ij} \\
&\quad + \beta_{4j}(African\ American)_{ij} + \beta_{5j}(Latino)_{ij} + \beta_{6j}(Poverty)_{ij} + \beta_{7j}(Social\ Status)_{ij} \\
&\quad + \beta_{8j}(Special\ Education)_{ij} + \beta_{9j}(8th\ Grade\ \ Underlying\ Achievement)_{ij} \\
&\quad + \beta_{10j}(Achievement^2)_{ij} + \beta_{11j}(Achievement^3)_{ij} + \beta_{12j}(Retained\ In\ 8th\ Grade)_{ij} \\
&\quad + \beta_{13j}(Post\ Policy\ Cohort)_{ij} + \beta_{14j}(Postpolicy \times Retained\ In\ 8th\ Grade)_{ij} \\
&\quad + \beta_{15j}(Postpolicy \times Special\ Education)_{ij} \\
\beta_{0j} &= \gamma_{00} + u_{0j} \\
\beta_{1j} &= \gamma_{10} \\
&\vdots \\
\beta_{15j} &= \gamma_{15\ 0}
\end{aligned}
$$

Where $Y_{ij} \mid \varphi_{ij} \sim B(1, \varphi_{ij})$

$$\eta_{ij} = \log\left[\varphi_{ij} / (1 - \varphi_{ij})\right]$$

Table 1: Model Predicting Log Odds of Dropping Out by Age 17, Students Who Reached Eighth Grade by Age 13

Parameter	Coeff.	(s.e.)
Intercept	-1.179	(.063) ***
Male	.277	(.017) ***
Native American	.486	(.160) **
Asian	-1.093	(.051) ***
African-American	-.529	(.062) ***
Latino	-.545	(.040) ***
Poverty	.041	(.016) **
Social status	-.029	(.011) **
Special Education Student	-.322	(.085) ***
Underlying Eighth Grade Achievement	-.306	(.032) ***
Underlying Achievement Squared	.022	(.014)
Underlying Achievement Cubed	-.025	(.006) ***
Retained in Eighth Grade	**1.054**	**(.122) *** **
Postpolicy	**.030**	**(.028)**
Postpolicy x Retained in Eighth Grade	**-.622**	**(.136) *** **
Postpolicy x Special Education Student	-.059	(.115)
Level-2 Variance Explained		44.4%

*p<.05, **p<.01, ***p<.001

All control variables except those for retention and post-policy were centered around zero prior to analysis, so that the intercept would represent the typical student. Metric variables were standardized. Only students included in the testing program and on-grade at age 13 were included in this analysis.

Changes in the timing of dropping out were measured with a discrete-time event history analysis, performed within HLM. The models, and resulting coefficients, are shown below and in Table 2. The model for Table 2 is:

$$\begin{aligned}
\eta(_{dropout})_{ij} = {} & \beta_{0j} + \beta_{1j}(Age\ 15)_{ij} + \beta_{2j}(Age\ 16)_{ij} + \beta_{3j}(Male)_{ij} + \beta_{4j}(Native\ American)_{ij} \\
& + \beta_{5j}(Asian)_{ij} + \beta_{6j}(African\ American)_{ij} + \beta_{7j}(Latino)_{ij} + \beta_{8j}(Poverty)_{ij} \\
& + \beta_{9j}(Social\ Status)_{ij} + \beta_{10j}(Special\ Education)_{ij} \\
& + \beta_{11j}(8th\ Grade\ Underlying\ Achievement)_{ij} + \beta_{12j}(Achievement^2)_{ij} \\
& + \beta_{13j}(Achievement^3)_{ij} + \beta_{14j}(Age\ 15\ x\ Achievement)_{ij} \\
& + \beta_{15j}(Age\ 15\ x\ Achievement\ Squared)_{ij} + \beta_{16j}(Age\ 15\ x\ Achievement^3)_{ij} \\
& + \beta_{17j}(Age\ 16\ x\ Achievement)_{ij} + \beta_{18j}(Age\ 16\ x\ Achievement^2)_{ij} \\
& + \beta_{19j}(Age\ 16\ x\ Achievement^3)_{ij} + \beta_{20j}(Retained\ In\ 8th\ Grade)_{ij} \\
& + \beta_{21j}(Age\ 15\ \times\ Retained)_{ij} + \beta_{22j}(Age\ 16\ \times\ Retained)_{ij} \\
& + \beta_{23j}(Post\ Policy\ Cohort)_{ij} + \beta_{24j}(Age\ 15\ \times\ Postpolicy)_{ij} + \beta_{25j}(Age\ 16\ \times\ Postpolicy)_{ij} \\
& + \beta_{26j}(Postpolicy\ \times\ Retained)_{ij} + \beta_{27j}(Age\ 15\ \times\ Postpolicy\ \times\ Retained)_{ij} \\
& + \beta_{28j}(Age\ 16\ \times\ Postpolicy\ \times\ Retained)_{ij} + \beta_{29j}(Postpolicy\ \times\ Special\ Education)_{ij} \\
\beta_{0j} = {} & \gamma_{00} + u_{0j} \\
\beta_{1j} = {} & \gamma_{10} \\
& \vdots \\
\beta_{29j} = {} & \gamma_{29\ 0}
\end{aligned}$$

Where $\eta_{ij} = \log[\varphi_{ij} /(1 - \varphi_{ij})]$

$$Y_{ij} \mid \varphi_{ij} \sim B(1, \varphi_{ij})$$

Table 2: Event History Models Predicting Dropping Out by Age

Predictor	Coeff.	Standard error
Intercept	-1.900	(.075) ***
Age 15	-.682	(.053) ***
Age 16	-.339	(.032) ***
Male	.247	(.016) ***
Native American	.436	(.117) ***
Asian	-1.000	(.041) ***
African-American	-.459	(.051) ***
Latino	-.481	(.033) ***
Poverty	.037	(.014) **
Social status	-.028	(.010) **
Special education student	-.169	(.069) *
Achievement	-.327	(.035) ***
Achievement squared	.025	(.017)
Achievement cubed	-.018	(.007) *
Age 15 x achievement	.269	(.053) ***
Age 15 x achievement squared	-.040	(.031)
Age 15 x achievement cubed	-.020	(.011)
Age 16 x achievement	-.030	(.029)
Age 16 achievement squared	-.030	(.016)
Age 16 achievement cubed	.000	(.008)
Retained in Eighth Grade	**.520**	**(.148) ***
Age 15 x retained	**.994**	**(.239) *** **
Age 16 x retained	**.455**	**(.169) ** **
Postpolicy	-.007	(.029)
Age 15 x postpolicy	-.085	(.045)
Age 16 x postpolicy	.153	(.033) ***
Postpolicy x retained	**-.187**	**(.164)**
Age 15 x postpolicy retained	**-.825**	**(.229) ** **
Age 16 x postpolicy x retained	**-.459**	**(.185) ** **
Postpolicy x special education	-.140	(.094)
Level-2 Variance Explained		35.9%

*p<.05, **p<.01, ***p<.001

All control variables except those for age, retention, and post-policy were centered around zero prior to analysis, so that the intercept would represent the typical student. Metric variables were standardized. Only students included in the testing program (i.e., subject to the eighth grade promotional gate if in post-policy cohorts) were included in this analysis.

[1] It is possible that those low-achieving postpolicy students who were not retained were less likely to drop out for some unmeasured reason, and so do not accurately represent the actual achievement-dropout relationship postpolicy. However, all available evidence suggests that this is not the case. First, nonretained postpolicy students were not a small, select group, but the majority of students. Even at the lowest levels of latent achievement, only about half of postpolicy students were retained in eighth grade. Second, an analysis of students who were "waived" through the gate (i.e., promoted despite failing the cutoff) showed that these students were not less likely than other students with the same latent achievement to drop out, suggesting that there was not a "waiver" bias in estimation. Third, the consistency of the latent achievement-dropout relationship among nonretained students across cohorts makes sense, given that the policy should not have affected these students' dropout rates, beyond any effect it might have had on their achievement.

[2] See chapter 10 in Raudenbush and Bryk (2002) for details on this type of model.

ENDNOTES

Introduction

[1] For details on the policy, see the sidebar, "How the Policy Works" and also Roderick, et al. (1999).

[2] These numbers include all students who repeated eighth grade or enrolled in a transition center, including double-retainees and students who entered CPS through a transition center.

[3] e.g., Alexander, Entwisle, and Kabbani (2001); Barro and Kolstak (1987); Grissom and Shepard (1989); Roderick (1994); and Rumberger (1995).

[4] Goldschmidt and Wang (1999); Roderick (1994).

[5] Allensworth and Easton (2001).

[6] Achievement has been shown consistently to be strongly related to dropout rates (e.g., Alexander, Entwisle, and Kabbani [2001]; Grissom and Shepard [1989]; and Pallas [1987]).

[7] Summaries of the debate over the implications of high-stakes testing on dropout rates are available from the National Research Council (2001) and Rabiniwitz, Zimmerman, and Sherman (2001). The conclusion of both is that more research is needed.

[8] There are studies that have looked at the effects of 12th-grade graduation exams on dropping out (Griffin and Heidorn [1996]) and Kreitzer, Madaus, and Haney [1989]. However, these extras affect students at the end of their academic career and should have different effects than those that determine grade promotion.

Chapter I

[1] Special–education classification increased substantially with enactment of the promotion gate, and the grades in which students were classified shifted. Details on this phenomenon are available in Gladden and Miller (2002).

[2] Promotion standards for students in special education were determined by their IEP. Administrative records show that retention rates among excluded students did not increase substantially with implementation of the eighth-grade gate. They rose slightly (from 5 to 7 percent) in the first year, and then declined each following year, reaching below prepolicy levels by the 1998-99 cohort.

[3] Averages for excluded students may have been substantially affected by selection bias; many more students excluded from reporting were tested in the later cohorts, so postpolicy averages include many students who would not have been tested in previous years. Furthermore, these students' scores were excluded from reporting precisely because they were not meant to be compared to those of other students. Both included and excluded averages are also affected by the increasing rates, postpolicy, of

exclusion from test reporting, particularly with the 1998 cohort. In the 1998 cohort (the cohort with the highest exclusion rates), 3 percent more students were excluded from testing than were excluded prepolicy. To find the maximum possible discrepancy due to the selection effect, the 1992-93 cohort average could be calculated without the lowest-achieving 4 percent of students. (Three percent of the entire cohort would be 4 percent of the 80 percent of students included in test reporting.) This would shift the 1992 average for included students from -.104 to -.034.

[4] The standard deviation of underlying achievement for the 1992 cohort alone is very similar to that across all cohorts (.985).

[5] Over the seven cohorts, relatively more students were Latino and fewer were African-American or white. Additionally, students' average economic status increased. These changes should have lowered dropout rates. On average, students with economically-advantaged backgrounds tend to have lower dropout rates than students from economically-disadvantaged backgrounds. Overall, dropout rates in Chicago are highest among African-American students, followed by Latino students, white students, and then Asian students. Once economic status and school enrollment are controlled, white students have the highest dropout rates, followed by Latino students, African-American students, and Asian students.

[6] There are substantial school effects on dropout rates (Bryk and Thum [1989]; Goldschmidt and Wang [1999]).

[7] Statistical controls were done through hierarchical linear models (HLM), which adjusted dropout rates for changes in the composition of students based on race, gender, and socioeconomic status as measured by economic indicators from census data of students' residential block groups. See the Appendix for a description of the basic models.

[8] For details on high-school redesign plans, see Chicago Public Schools (1997).

[9] For details on high-school redesign and an analysis of its effectiveness, see Hess (2002). For an analysis of trends in credit accumulation and graduation rates, see Miller, Allensworth, and Kochanek (2002).

[10] Hess and Cytrynbaum (2002) found that high schools on probation showed no significant improvements in instruction or improvements in personalism. To the contrary, the focus on reading that came about from probation status resulted in less attention paid to other curricular topics while not improving substantially the reading ability of high-school students.

[11] See Bryk (2003) for elaboration on this theory.

[12] Kidscount (2003).

Chapter 2

[1] The prepolicy relationship between achievement and retention was small (-.05) compared to postpolicy (-.32), among students included in test reporting.

[2] The issue is complicated by the nonlinear relationship between achievement and the probability of dropping out.

[3] "Underlying achievement" is the term used in this report for the measure of latent achievement obtained from HLM models that adjusted students' eighth-grade score for exceptionally low or high performance, based on their scores from earlier grades.

[4] The achievement-dropout relationship was virtually the same among *nonretained* students pre- and postpolicy (within one standard error), with linear terms of -.323 (.039) and -.297 (.033), respectively. When retained students are included, the relationship appears stronger postpolicy because the retention effect is incorporated into the achievement-dropout relationship.

[5] It might have been that those students who remained with low achievement postpolicy were also those that were most likely to drop out (e.g., because they were those most disengaged from school). However, if this were the case, the relationship between latent achievement and dropping out should have been stronger postpolicy, especially among students with very low achievement, and it was not. Analysis of residuals from models with only achievement and demographic variables, using students not retained by the gate, shows no difference between prepolicy and postpolicy students in the relationship of latent achievement with dropping out, even among students at the lowest levels of achievement. It was also possible that those postpolicy students with very low achievement who were not retained were also those who were less likely to drop out because of some unmeasured characteristic. This would be especially true if they were a small subset of the total number of postpolicy students. However, they are not a small subset—at even the very lowest levels of achievement, only about half of postpolicy students were retained in eighth grade. Overall, 90 percent of postpolicy students were not retained. Furthermore, if postpolicy students who were not retained were a special subset of students who were less likely to drop out than average because of some unmeasured characteristic, they should have been less likely to drop out than prepolicy students with the same latent achievement. But another analysis showed that students waived through the gate were as likely to drop out as prepolicy students with the same latent achievement.

[6] Students not yet in eighth grade when 13-years-old were taken out of these estimates because many had irregular testing and grade-progression patterns which introduce error into measurement of the retention effect. Separate analyses of old-for-grade students show a similar retention effect to that of on-grade students.

[7] This is estimated from the statistical model, with achievement one standard deviation below the mean, which is typical for retained students.

[8] This difference in retention effects may result because retention resulting from high-stakes testing was qualitatively very different than teacher-initiated retention. It also may result because much of the prepolicy relationship is actually attributable to spurious factors. Many of the same factors that cause students to be held back in school may also contribute to their decisions to drop out (e.g., lack of engagement, low school attendance). It is unlikely that studies of teacher-initiated re-

tention have sufficiently controlled for these spurious factors. One advantage in this study is that postpolicy retention decisions were primarily based on one criterion (test scores), which could be statistically controlled.

[9] This is 0.88 percent of the cohort included in testing and on-grade at age 13.

[10] Postpolicy, 7.8 percent of students included in testing and in eighth grade at age 13 were retained in eighth grade, but 0.88 percent of the total cohort (which is 11.3 percent of the 7.8 percent held back) would likely have been retained without the high-stakes testing policy. If 11.3 percent of the postpolicy *retained* students had the same likelihood of dropping out as students retained prepolicy, then the retention effect due to high-stakes testing is: [.433 – .113 (1.054)] / .887.

Chapter 3

[1] This was done through a discrete-time-event history analysis using hierarchical linear models. As with the previous models, the analyses included only students who were not old-for-grade at age 13 and had test scores included in test reporting. Separate analyses for students who were old-for-grade at age 13 are available from the author.

[2] Students who dropped out between age 13 and 15 were classified as 15-year-old dropouts because their numbers were so small.

[3] Postpolicy, 21 percent of students who dropped out by age 17 never made it to high school. This compares to 15 percent of prepolicy dropouts.

Chapter 4

[1] These calculations were made by comparing differences in the latent achievement of all students with test scores across the cohorts, adjusted for changes in demographic composition that could have affected achievement. On average, latent scores were .280 standard deviations higher postpolicy than prepolicy, adjusted for changes in demographic characteristics (.286 without adjustments). The achievement coefficients come from models of nonretained students that have only coefficients representing demographic variables and achievement, but they are also equivalent to the coefficients in the full model in the Appendix. Scores were compared among students with the same class rank so that differences in achievement gains by low- and high-achieving students could be discerned. Selection bias was an issue in these calculations because only students included in test reporting and in eighth grade at age 13 were included as they were based on test-score data analysis. Therefore, the figure includes a note with an estimate of the maximum extent to which this bias could exist.

[2] The estimate of the achievement effect varies across students according to: 1) their overall probability of dropping out given other factors; and 2) the extent to which achievement improved across the cohorts among students with similar class rank. (Because the gate should have had the largest effects on achievement among students at risk of failing the gate, achievement growth was compared among students with similar class rank.) This estimate is for students on-grade at age 13. For students old-for-grade at age 13 and retained by the gate, the combined effect of retention and achievement may be more adverse because many of these students had extremely low achievement, and achievement gains were smallest among very-low- and very-high-performing students.

[3] The 3.5 percent figure assumes that the lowest-achieving prepolicy students would have been excluded if part of the 1998-99 cohort.

[4] Adjusted for selection bias, the 1998-99 estimate is as small as 2.7 percentage points.

Chapter 5

[1] Underlying achievement scores showed the same improvements among students old-for-grade at age 13 as among students in eighth grade at age 13, with the exception of the 1998 cohort. The 1998 cohort of old-for-grade students showed a smaller increase in achievement over the 1997 cohort than seen among on-grade students. This was likely an artifact of the sixth-grade promotion gate, which made the lowest-achieving students in the cohort old-for-grade by age 13.

[2] See Miller and Gladden (2002) for details on increases in special-education placements at the promotion gates.

[3] See Miller and Gladden (2002) for a description of these trends.

[4] Differences in eighth-grade retention rates and achievement growth by race explain about one-third of the difference between African-American and white students in postpolicy improvements in dropout rates. The remaining differences likely result from school differences (see the next endnote).

[5] The high dropout rates among African-American students are explained largely by higher rates of attendance in schools with very high dropout rates. Because Chicago's schools are highly segregated by race, it is difficult to disentangle school effects from race effects. However, schools that were predominantly (over 85 percent) African-American were much more likely than other schools to be put on probation during the period studied, and this may have been related to the lack of improvement in dropout rates among African-American students. Schools on probation received assistance to improve their test scores, but faculty experienced substantial stress and uncertainty, which may have affected the climate of the schools in ways that were not conducive to retaining students. Further research is needed on this. Gender differences in dropout trends are largely not explained by different retention rates or changes in achievement, particularly among white and Asian students. These gender differences may partially reflect the decline in births to teenage mothers that has occurred over the last several years, and changes in welfare reform that required school attendance.

Interpretive Summary

[1] In particular, the very strong relationship between middle-grade retention and dropping out (e.g., seen in Alexander, Entwisle, and Kabbani [2001] and Roderick [1994]) likely results from being commonly related to disengagement from school. Students retained in the middle-school years tend not to have been the lowest-performing students in the early grades. Instead, their performance declined over the elementary years (see Frank and Bryk [1991]), suggesting stressors in students' lives that may have made them more likely to disengage from school.

[2] Research on teacher-initiated retention suggests it can have negative psychological consequences, such as lowering students' self-efficacy, and increasing feelings of failure and negative attitudes towards school that might encourage students to leave school (e.g., Byrnes [1989]; Campbell [1993]).

[3] Only about half of first-time high-school students in CPS receive enough credits in their first year to be on-track to graduate in four years (Miller, Allensworth, and Kochanek [2002]).

[4] Conflicts with other students and with teachers and concerns about safety are related to students' decisions to drop out (Caterall [1998]; DeLuca and Rosenbaum [2000]). Many students in CPS schools report poor relationships with teachers and problems with safety in school (Sporte 2003).

[5] For example, family or financial responsibilities

[6] For example, 13 percent of the prepolicy students with test scores in the top quarter of their class dropped out of school by age 17, and 20 percent dropped out by age 19.

[7] Jacob (2002) estimated that most of the improvements in test scores after high-stakes testing in Chicago were a result of motivational and test-specific factors. Likewise, in their study of the TAAS in Texas, Klein et al. (2001) found that gains on the high-stakes test were not mirrored in the NAEP, suggesting that basing judgments of student achievement on high-stakes tests may be misleading. Because underlying test scores were used in these analyses, rather than actual eighth-grade scores, motivational factors may have been minimized in these analyses. Still, some of the improvements in test scores were likely due to better preparation for the tests.

[8] Unfortunately, there is little evidence on the long-term effects of these programs with eighth graders. In general, eighth-grade students had very positive reports about their summer-school experiences, and they showed substantial short-term gains. Evidence at the third- and sixth-grade levels suggests that short-term gains in summer school were not sustained over two years, but eighth-grade two-year gains could not be calculated because students no longer take the ITBS after eighth grade. The short-term summer school eighth-grade gains were larger than those seen in the earlier grades (Roderick, Engel, and Nagaoka, [2003]).

[9] For example, they note that school systems in Boston and North Carolina have used high-stakes tests to identify students who are not progressing and focus interventions on them, rather than retaining them.

Appendix

[1] It is possible that those low-achieving postpolicy students who were not retained were less likely to drop out for some unmeasured reason, and so do not accurately represent the actual achievement-dropout relationship postpolicy. However, all available evidence suggests that this is not the case. First, nonretained postpolicy students were not a small, select group, but the majority of students. Even at the lowest levels of latent achievement, only about half of postpolicy students were retained in eighth grade. Second, an analysis of students who were "waived" through the gate (i.e., promoted despite failing the cutoff) showed that these students were not less likely than other students with the same latent achievement to drop out, suggesting that there was not a "waiver" bias in estimation. Third, the consistency of the latent achievement-dropout relationship among nonretained students across cohorts makes sense, given that the policy should not have affected these students' dropout rates, beyond any effect it might have had on their achievement.

[2] See chapter 10 in Raudenbush and Bryk (2002) for details on this type of model.

WORKS CITED

Alexander, Karl L., Doris R. Entwisle, and Carrie S. Horsey. 1997. From first grade forward: Early foundations of high school dropout. *Sociology of Education*, 70(2):87-107.

Alexander, Karl L., Doris R. Entwisle, and Nader S. Kabbani. 2001. The dropout process in life course perspective: Early risk factors at home and school. *Teachers College Record*, 103(5):760-822.

Allensworth, Elaine M. and John Easton. 2001. *Calculating a cohort drop-out rate for the Chicago Public Schools*. Chicago: Consortium on Chicago School Research.

Allensworth, Elaine M. and Shazia Rafiullah Miller. 2002. *Declining high school enrollment: An exploration of causes*. Chicago: Consortium on Chicago School Research.

The Annie E. Casey Foundation. 2003. Teen childbearing in America's largest cities, A KidsCount working paper. http://www.aecf.org/kidscount/childbear/cb-changes.htm.

Bryk, Anthony S. 2003. No child left behind, Chicago style. In *No child left behind? The politics and practice of school accountability,* eds. Paul E. Peterson and Martin R. West. Washington DC: Brookings Institution.

Bryk, Anthony S. and Yeow Meng Thum. 1989. The effects of high school organization on dropping out: An exploratory investigation. *Amercian Educational Research Journal,* 26(3):353-83.

Bryk, Anthony S., Yeow Meng Thum, John Q. Easton, and Stuart Luppescu. 1998. *Academic productivity of Chicago public elementary schools: A technical report.* Chicago: Consortium on Chicago School Research.

Catterall, James S. 1998. Risk and resilience in student transition to high schools. *American Journal of Education*, 106:302-33.

Chicago Public Schools. 1997. Design for high schools.

Dworkin, Anthony Gary, Jon Lorence, Laurence A. Toenjes, Antwanette N. Hill, Noe Perez, and Michael Thomas. 1999. Elementary school retention and social promotion in Texas: An assessment of students who failed the reading section of the TAAS. The Sociology of Education Research Group, University of Houston.

Finn, Jeremy D. 1989. Withdrawing from school. *Review of Educational Research*, 59(2):117-42.

Gladden, Robert M. and Shazia Rafiullah Miller. 2002. *Changing special education enrollments: Causes and distribution among schools*. Chicago: Consortium on Chicago School Research.

Gleason and Dynarski. 2002. Do we know whom to serve? Issues in using risk factors to identify dropouts. *Journal of Education for Students Placed at Risk*, 7(1):25-42.

Goldschmidt, Pete and Jia Wang. 1999. When can schools affect dropout behavior? A longitudinal multilevel analysis. *American Educational Research Journal*, 36(4):715-38.

Griffin, Bryan W. and Mark H. Heidorn. 1996. An examination of the relationship between minimum competency test performance and dropping out of high school. *Education Evaluation and Policy Analysis*, 18(3):243-52.

Grissom, James B. and Lorrie A. Shepard. 1989. Repeating and dropping out of school. In *Flunking grades: Research and policies on retention*, eds. Lorrie A. Shepard and M. L. Smith. New York: The Falmer Press, 34-63.

Hess, G. Alfred and Solomon Cytrynbaum. 2002. The effort to redesign Chicago high schools: Effects on schools and achievement. In *Reforming Chicago's high schools: Research perspectives on school and system level change*, ed. Valerie Lee. Consortium on Chicago School Research, 19-49.

Holmes, Charles Thomas and Janice Saturday. 2000. Promoting the end of retention. *Journal of Curriculum and Supervision*, 15(4):300-14.

Holmes, Charles Thomas. 1989. Grade level retention effects: A meta-analysis of research studies. In *Flunking grades: Research and policies on retention*, eds. Lorrie A. Shepard and M. L. Smith. New York: Falmer Press, 16-63.

House, Ernest R. 1998. *The predictable failure of Chicago's student retention program*. Boulder, CO: University of Colorado School of Education.

Jacob, Brian A. 2002. Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools, working paper 8968. Cambridge, MA: National Bureau of Economic Research.

Kaufman, Philip, Martha Naomi Alt, and Christopher D. Chapman. Dropout rates in the United States: 2000. 2001. NCES 2002-114. Washington DC: U.S. Department of Education, National Center for Education Statistics.

Klein, Stephen P., Laura S. Hamilton, Daniel F. McCaffrey, and Brian M. Stecher. 2000. *What do test scores in Texas tell us?* Rand Education Issue Paper, http://www.rand.org/publications/IP/IP202/

Kreitzer, Amelia E., George F. Madaus, and Walt Haney. 1989. Competency testing and dropouts. In *Dropouts from school: Issues, dilemmas, and solutions*, eds. L. Weis, E. Farrar and H. G. Petrie. Albany, NY: State University of New York Press, 129-52.

Miller, Shazia Rafiullah, Elaine M. Allensworth, and Julie R, Kochanek. 2002. *Student performance: Course taking, test scores and outcomes*. Chicago: Consortium on Chicago School Research.

Nagaoka, Jenny and Melissa Roderick. 2004. *Ending social promotion: The effects of retention*. Chicago: Consortium on Chicago School Research.

National Research Council. 2001. Understanding dropouts: Statistics, strategies, and high-stakes testing, eds. Alexandra Beatty, Ulric Neisser, William T. Trent, and Jay P. Heubert. Board on Testing and Assessment, Center for Education. Washington, DC: National Academy Press.

Pallas, Aaron M. 1987. *School dropouts in the United States*. Center for Education Statistics, Washington, DC.

Rabinowitz, Stanley, Joy Zimmerman, and Kerry Sherman. 2001. *Do high stakes tests drive up student dropout rates? Myths versus reality*. San Francisco, West Ed.

Raudenbush, Stephen W. and Anthony S. Bryk. 2002. *Hierarchical linear models: Applications and data analysis methods*, 2nd Edition. Thousand Oaks, CA: Sage Publications.

Reynolds, Arthur J. 1992. Grade retention and school adjustment: An explanatory analysis. *Educational Evaluation and Policy Analysis*, 14(2):101-21.

Roderick, Melissa. 1994. Grade retention and school dropout: Investigating the association. *American Educational Research Journal* 31(4):729-59.

Roderick, Melissa, Anthony S. Bryk, Brian A. Jacob, John Q. Easton, and Elaine M. Allensworth. 1999. *Ending social promotion: Results from the first two years*. Chicago: Consortium on Chicago School Research.

Roderick, Melissa and Mimi Engel. 2001. The grasshopper and the ant: Motivational responses of low achieving students to high-stakes testing. *Journal of Educational Evaluation and Policy Analysis*, 23(3): 197-227.

Roderick, Melissa, Brian A. Jacob, and Anthony S. Bryk. 2002. High stakes testing in Chicago: Effects on achievement in promotional gate grades. *Educational Evaluation and Policy Analysis*. 24: 333-58.

Roderick, Melissa, Mimi Engel, and Jenny Nagaoka. 2003. *Ending social promotion: Results from Summer Bridge*. Chicago: Consortium on Chicago School Research.

Rosenkranz, Todd. 2002. *2001 CPS test trend review: Iowa Tests of Basic Skills*. Chicago: Consortium on Chicago School Research.

Rumberger, Russel W. 1995. Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal*, 32(3):583-625.

Sporte, Sue. 2003. *Key Measures of School Development*. Chicago: Consortium on Chicago School Research.

ABOUT THE AUTHOR

Elaine Allensworth is the Associate Director for Statistical Analysis and Archives at the Consortium on Chicago School Research. Her research focuses on the structural factors that affect school development, as well as policy effects on high school student outcomes. She is currently analyzing system- and school-level factors that affect trends in dropout rates, and is beginning work on students' postsecondary outcomes. Elaine is part of the research team that is testing and elaborating the Consortium's Theory of Essential Supports. She holds a Ph.D. in Sociology from Michigan State University.

Consortium on Chicago School Research

Mission

The Consortium on Chicago School Research aims to conduct research of high technical quality that can inform and assess policy and practice in the Chicago Public Schools. By broadly engaging local leadership in our work, and presenting our findings to diverse audiences, we seek to expand communication between researchers, policy makers, and practitioners. The Consortium encourages the use of research in policy action, but does not argue for particular policies or programs. Rather, we believe that good policy is most likely to result from a genuine competition of ideas informed by the best evidence that can be obtained.

Founded in 1990, the Consortium is located at the University of Chicago.

Consortium on Chicago School Research
1313 East 60th Street, Chicago, IL 60637
773-702-3364 fax -773-702-2010
www.consortium-chicago.org

ISBN 0-9726035-3-0