Assessing the Academic Mobility on Test and Non-Test Outcomes for Students from Third Grade Through High School

Dan Goldhaber
Stephanie Liddle
Suvekshya Gautam

February 2026

WORKING PAPER No. 330-0226-2

CALDER
National Center for Analysis of Longitudinal Data in Education Research

AIR®

Assessing the Academic Mobility on Test and Non-Test Outcomes for Students from Third Grade Through High School

Dan Goldhaber
American Institutes for Research / CALDER
University of Washington / CEDR

Stephanie Liddle
University of Washington / CEDR

Suvekshya Gautam
University of Washington / CEDR

Contents

Acknowledgments

The research presented here uses confidential data from the Education Research and Data Center (ERDC) located within the Washington Office of Financial Management (OFM). ERDC's data system is a statewide longitudinal data system that includes de-identified data about people's preschool, educational, and workforce experiences. The views expressed here are those of the authors and do not necessarily represent those of OFM or other data contributors. Any errors are attributable to the authors. We are grateful to the Bill and Melinda Gates Foundation for funding this work. We also wish to acknowledge Katherine Baird for her work on an earlier version and useful comments from Roddy Theobald and Micheal DeArmond.

CALDER working papers have not undergone final formal review and should be cited as working papers. They are intended to encourage discussion and suggestions for revision before final publication. Any errors are attributable to the authors.

CALDER • American Institutes for Research
1400 Crystal Drive 10th Floor, Arlington, VA 22202
202-403-5796 • www.caldercenter.org

Assessing the Academic Mobility on Test and Non-Test Outcomes for Students from Third Grade Through High School

Dan Goldhaber, Stephanie Liddle, & Suvekshya Gautam

CALDER Working Paper No. 330-0226-2
February 2026

Abstract

While achievement gaps have been widely studied, less attention has been given to how these gaps relate to students' patterns of *academic mobility*—that is, how students' relative academic standing changes from early to later grades. Using rich administrative data from Washington state, we apply an academic mobility framework to examine both test-based and non-test outcomes for student subgroups defined by race/ethnicity and gender, tracking their progress from 3rd grade through high school. Our findings align with prior research: after controlling for other student characteristics, higher performance on 3rd or 8th grade assessments strongly predicts better high school outcomes, including test scores, GPAs, college-level course-taking in math and ELA, graduation rates, and lower rates of absenteeism and disciplinary incidents. We also find that early math achievement is a stronger predictor of later outcomes than early ELA performance. Although some initially low-performing students demonstrate upward mobility, 3rd grade test scores remain powerful predictors of long-term academic trajectories. Students who begin behind tend to stay behind. Though magnitudes of academic mobility vary significantly across outcomes and subgroups. Asian students, especially Asian females, were the most upwardly mobile subgroup while Black males were often, though not always, the least mobile. Gender gaps also vary in size nearly always favoring females. Our results underscore the need for earlier and more intensive interventions to change the educational trajectories of students who struggle academically.

1. Introduction

Understanding the factors explaining variation in students' academic performance has been a focus of policy and research going all the way back to the *Coleman Report* (Coleman et al., 1966). Finding what drives academic achievement and achievement trajectories is key to informing policy and practice decisions designed to address student needs. Thus, tracing the typical academic progression of student groups is important to both identify which students may benefit from interventions as well as to discern the points in a student's academic career when interventions are most likely to be beneficial.

It is well-known that measures of students' academic achievement at one point in school are highly predictive of both later academic performance and post-schooling outcomes (e.g., Chetty et al., 2014; Goldhaber and Ozek, 2019; Gray-Lobe et al., 2023). Indeed, various policies and interventions are predicated on the predictive power of early academic skills and the notion that key skills should be achieved by certain grades to help ensure later schooling success. It has, for instance, become a norm to assert that reading by the 3rd grade is a necessary skill (Fiester, 2010). Consistent with this, several states require students to achieve test benchmarks in the 3rd grade or risk being retained (ExcelinEd, 2024; Jacob et al., 2017; LiCalsi et al., 2019). Some states and localities have also developed early warning systems and diagnostic tools to help alert practitioners that individual students are at risk of not meeting specific skills or benchmarks such as high school graduation (e.g., Allensworth, 2013; Berne, 2025; Curtin et al., 2012).

The persistence of gaps in academic achievement across student subgroups is also well-known. At a national level, results from the 2024 National Assessment of Educational Progress (NAEP) show that racial/ethnic gaps in 4th grade performance favored Asian students over White students by 0.29 standard deviations in math and 0.24 standard deviations in reading. White students outperformed students from other non-Asian racial/ethnic groups by between 0.59

(Hispanic) and 0.85 (American Indian) standard deviations in math and between 0.53 (Hispanic) and 0.73 (American Indian) standard deviations in reading. Gender gaps were smaller favoring females in reading (by 0.20 standard deviations) and males in math (by 0.15 standard deviations).[1]

While there is extensive research linking academic and demographic factors to K-12 differential schooling outcomes, most of it relies on linkages across short grade spans (e.g., elementary to middle, or middle to high school), focuses mainly on test outcomes, and compares student subgroups along a single dimension, e.g., race/ethnicity, gender, or socioeconomic status. Fewer studies have applied an intersectionality framework (see Crenshaw, 1989) to understanding how students' characteristics may intersect to influence achievement gaps.

Furthermore, despite extensive investigations into achievement gaps, little attention has been paid to how they relate to patterns of students' "academic mobility" throughout their schooling—how a student's standing on academic measures relative to one's peers changes as they move from early to later grades (McDonough, 2015). One recent exception is work by Austin et al. (2023). Their examination, which includes data from students in seven different states, found student rankings—at least in terms of test scores and high school graduation—to be highly persistent between 3rd grade and high school, reflecting relatively little upward academic mobility.

In this paper, we analyze academic success through an academic mobility framework akin to Austin et al. (2023). This framework focuses on how students' rankings in various academic measures evolve over their school careers. While we follow Austin et al. in defining academic mobility, our analysis extends beyond test scores and graduation to include a broader

[1] For more details, see: https://nces.ed.gov/nationsreportcard/.

range of non-test academic outcomes for initially low-achieving 3rd-grade students. Tracking this group is crucial, as early achievement gaps often widen over time, increasing the risk of lower high school performance and dropout (Duncan et al., 2007; Reardon, 2011). These disparities, often tied to socioeconomic factors, highlight the need to study mobility patterns to ensure equitable opportunities (Heckman, 2006). We also examine how academic mobility varies by race and gender and explore the progression of academic mobility from grade to grade. Understanding these patterns can inform targeted interventions to close achievement gaps and support long-term student success (Alexander, Entwisle, & Olson, 2007).

We utilize rich administrative data from Washington State to assess how demographic characteristics as well as 3rd and 8th grade test achievement predict the academic success of public school students through high school. This paper contributes to the literature in several ways. First, we consider a wide variety of non-test high school outcomes that, to our knowledge, have yet to be addressed within an academic mobility framework, such as: college-level course-taking, absenteeism, disciplinary incidences, and grade point average (GPA). This is important given new research suggesting that non-test factors in school predict longer-run outcomes (Backes et al., 2023; Jackson et al., 2020; Jackson, 2018). Second, since we include a census of 3rd grade public school students in an entire state, we can examine outcomes for subgroups of students more extensively than prior research, making this paper one among a limited number of studies that analyze high school outcomes for intersectionalites of student subgroups, defined by the combination of a student's race/ethnicity and gender. Third, we examine the *progression* of academic mobility from grade to grade and whether there are points along students' educational careers (such as the transition from elementary to middle school) where achievement gaps tend to emerge. More specifically, we address the following research questions (RQs):

1. How predictive are students' earlier-grade (3^{rd} or 8^{th}) test scores of their test and non-test high school outcomes?
2. To what extent do patterns of academic mobility between 3rd grade test scores and high school outcomes vary by student subgroups?
3. What is the progression of academic mobility from grade to grade and are there particular grade(s) where the trajectories of student subgroups diverge?

We have several main findings. First, consistent with previous literature, we find that—controlling for other student characteristics—students who performed better on 3^{rd} or 8^{th} grade tests, had higher high school test scores and GPAs, took more college-level math and ELA courses, were more likely to graduate, and had fewer absences and disciplinary incidents. We also find that early achievement in math (from 3^{rd} or 8^{th} grade) was generally more predictive of high school test and non-test outcomes than was early achievement in ELA.

Second, compared to models that use 8^{th} grade tests as predictors, models with 3^{rd} grade tests offer a surprisingly close approximation of high school outcomes with correlations across such models of at least 0.75.

Third, academic mobility varies considerably across outcomes for students with the same achievement level in 3^{rd} grade. Low-achieving 3^{rd} graders from all subgroups were, to some extent, upwardly mobile when looking at high school test scores, GPA, and graduation. For instance, on average, students who scored at the 25^{th} percentile of their 3^{rd} grade test score distribution moved up to the 31^{st} percentile of that performance distribution by high school. But those same students experienced varying degrees of downward mobility in college-level course-taking, absenteeism, and (for males) disciplinary incidents.

Fourth, the magnitudes of academic mobility for the various outcomes vary significantly across subgroups. Asian students were the most upwardly mobile racial/ethnic subgroup—an advantage generally even more pronounced for Asian females. In contrast, Black males were often—though not always—the least mobile. Gender gaps also varied in size but almost always favored females.

Fifth, students showed a general upward *progression* in test-based academic mobility from grade to grade (3rd grade through high school). The ethno-racial gap in the progression of test-based academic mobility between the most (Asian) and least (Black or Hispanic) upwardly mobile groups were relatively small in 4th grade—about 6 percentage points—but widened as students progressed through school. The widening of these gaps appears largely driven by the near constant increasing upward mobility of Asian students as they progressed from grade to grade.

2. Background on the Factors Predicting Progression and Success Through Public Schools

There is a significant academic literature that focuses on the predictors of long-run K-12 academic and postsecondary success. Much of this examines the extent to which demographic factors are predictive of academic outcomes as students progress through school. A primary measure has been test scores. Indeed, testing students annually and using the results to identify and communicate students' needs and to inform policy decisions has been one of the primary strategies to battle educational inequities. While the use of tests as a measure of identifying learning gaps is contentious (e.g., Forte, 2021; Koretz, 2017; Strauss, 2015), research finds that tests are an important predictor of future academic and labor market success (see Goldhaber & Ozek (2019) for a review).

Gaps in test scores between different student groups have long been a focus of social science research. Several recent studies have shown that racial/ethnic gaps in student achievement have been declining throughout the last decade or so, though the magnitude of change in these educational gaps varies for different races and for different outcomes. For instance, the White-Hispanic gap for test scores has narrowed more than the White-Black gap and the evidence for other non-test school outcomes are less pronounced (Reardon & Portilla, 2016). Studies have also shown that the magnitude of changes in gaps vary by other factors such as segregation, lagged school and home inputs, cultural and social factors, and parental characteristics (Card & Rothstein, 2007; Domina et al., 2017; Duncan & Magnuson, 2005; Gamoran, 2001; Hanushek & Rivkin, 2009; Kao & Thompson, 2003; Reardon, 2015; Todd & Wolpin, 2007). The degree to which schools mitigate or amplify overall inequality is a topic of significant discussion, as highlighted by varying viewpoints (Domina et al., 2017; Dumont & Ready, 2020; Northrop, 2017; von Hippel et al., 2018).[2] Martínez and Serna (2018) examine the evolution of gender disparities in mathematics achievement throughout elementary school and find that gender gaps in math scores are minimal at the start of schooling but expand rapidly as students' progress. Between first and fifth grade, these gaps increase by 60.8%.

There is also significant academic focus on inequality in educational achievement and the degree to which gaps in student achievement change as students move from grade to grade. Research indicates that these disparities mainly emerge during the early stages of childhood, preceding the commencement of formal education (von Hippel and Hamrock, 2019), and persist throughout the students' academic progression (Clotfelter et al., 2009; Finkelstein and Fong,

[2] Teacher quality and within-school sorting of teachers and students may cause some of this mitigation or amplification (Goldhaber et al., 2022; Northrop, 2017), and there is some evidence that gaps decrease over time if schools and teachers cater to the needs of disadvantaged students by offering targeted instruction (Downey et al., 2004; von Hippel et al., 2018).

2008). The findings across both reading and math scores (Betts et al., 2003; Clotfelter et al., 2009; Reardon, 2011) and non-test outcomes (Gamoran, 1987; Kelly, 2009; Lee, 2002) document persistence in student achievement gaps.

The decades-long trend toward the reversal of the gender gap in educational outcomes—now largely (though not always) favoring females—is well known (see Reeves (2022) for a review of this trend and recent literature). There is considerable evidence for this change in terms of tertiary schooling outcomes such as college enrollment and completion (see Conger & Long, 2010; Diprete & Buchmann, 2013). Yet there is evidence that gender gaps appear in grade school as early as kindergarten (Chatterji, 2006; Tach & Farkas, 2006) and persist into middle (Quinn & Cooc, 2015; Reardon, 2019; Robinson & Lubienski, 2011) and high school (Fortin & Phipps, 2015; Reardon, et al. 2019).

Though less studied, researchers have also explored the link between various student characteristics and non-test outcomes. Sorensen (2019), for example, predicts high school dropout or graduation based on early academic, behavioral and background indicators. Goldhaber, Wolff, and Daly (2021) show that 3rd grade test scores are strong predictors of high school graduation and college-level course-taking. A large body of research has documented stark disparities in both absenteeism (e.g., Gottfried, 2014; Malkus, 2024) and disciplinary incidents (Gregory, Skiba, and Noguera 2010; Losen et al. 2015; Skiba et al. 2011) across racial, socioeconomic, and gender groups. Furthermore, related work suggests that some degree of subjectivity or bias exists when educators are deciding what disciplinary actions to take against students as Black and Hispanic students are delt harsher punishments than White students when involved in the same incident (Barrett et al., 2021; Liu et al., 2024; Shi & Zhu, 2022). However, while these behaviors are often treated as early warnings of later outcomes, such as disengagement and drop out, far less is known about how early academic performance itself may shape students' future attendance (except see Skedgell & Kearney (2018)) and disciplinary outcomes.

Regardless of whether prior research has focused on measures of academic skill (i.e., tests) or other non-test aspects of schooling success, most studies have examined overall differences between student subgroups along a single dimension, e.g., race/ethnicity, socioeconomic status, or gender. A small but growing number of studies have applied an intersectionality framework to academic achievement to examine how students' background factors may intersect to influence K-12 outcomes (see Becares & Priest, 2015; Coffey & Tyner, 2023; Conwell, 2021; Gordon & Cui, 2018; McGraw et al. 2006).

Our work adds to the literature by following multiple student cohorts throughout their entire 3–12 trajectories, allowing us to examine how early academic performance relates to a broad set of test and non-test outcomes over time. We place particular emphasis on academic mobility among initially low-performing students, examining what the trajectories look like for different subgroups within this achievement category and how these patterns relate to later non-test outcomes such as college-level course-taking behavior, absenteeism, disciplinary incidents and GPA—capturing dimensions of student progress beyond test performance.

Closest to what we do is prior work by Austin et al. (2023) who advance the concept of academic mobility to describe the extent to which students' *ranks* in the distribution of academic performance *change* during their school careers. For instance, they look at students' initial performance levels on 3rd grade tests to predict longer-term outcomes like students' ranks in test performance distributions in 8th grade and high school as well as high school graduation. They find considerable heterogeneity in the academic mobility of students of different race/ethnicities and by economic (free or reduced-price lunch) status, and across the districts in which students are enrolled.

To our knowledge, there are no studies that look at academic mobility across any non-test outcome other than high school graduation. Nor do any studies examine academic mobility within an intersectionality framework or look at grade-to-grade changes as students progress through their primary school careers. We build upon prior research by examining students' academic mobility for intersectionalities of two demographic characteristics (race/ethnicity + gender) as they relate to both test and non-test high school outcomes (beyond just high school graduation) and—for some outcomes—the progression of these patterns from grade to grade.

3. Data, Measures, and Analytical Sample

The data for this study come from two data systems maintained by the Office of the Superintendent of Public Instruction (OSPI): the Core Student Record System (CSRS) and the Comprehensive Education Data and Research System (CEDARS). Individual student records in CSRS and CEDARS are longitudinally linkable across years based on a unique student record ID that the state maintains.

3.1 Core Student Record System (CSRS) and Comprehensive Education Data and Research System (CEDARS) Data

The CSRS and CEDARS data systems include annual information on student demographics (race/ethnicity and gender); student classifications (free or reduced-price lunch eligibility, special education/learning disability/highly capable statuses and limited English proficiency); as well as various educational outcomes. CSRS includes this information in earlier school years from 2004-05 to 2008-09 and CEDARS for school years thereafter.

These systems include test outcomes on statewide standardized tests for students in 3^{rd} - 8^{th} and 10^{th} or 11^{th} grade in math and English language arts (ELA).[3] We keep students' earliest subject-specific scores in each grade from 3^{rd} to 8^{th} grade and high school which we standardize within subject, year and grade to have a mean of 0 and a standard deviation of 1. Our main predictors in this paper are the subject-specific 3^{rd}/8^{th} grade math and ELA standardized test scores as well as the average of students' standardized math and ELA scores.

For selected years, the CEDARS data also have information on other non-test student outcomes. Beginning in the 2009-10 school year, the data track students' course-taking behavior, absences, grade point averages, and graduation outcomes. In the 2012-13 school year, the state changed the measure of absences to capture full-/part-day and excused/unexcused absences. And in the 2013-14 school year, the state began collecting data on student disciplinary actions. From these data holdings we construct several measures of test and non-test outcomes for students during their high school careers (see Table 1 for a summary).

Test scores. We use high school test scores on the math and ELA Smarter Balanced Assessments which have been required of 11th graders since the 2014-15 school year and 10^{th} graders since the 2017-18 school year. We keep students' earliest high school scores regardless of whether they were tested in 10^{th} or 11^{th} grade. As we note above, because tests have changed through the years, we standardize test scores within subject, year and grade to have a mean of 0 and a standard deviation of 1.

---

[3] Math and ELA tests in Washington changed over the period for which we have data. Students in grades 3-8 took the Washington Assessment of Student Learning (WASL) from 2005-06 to 2008-09 and the Measurements of Student Progress (MSP) from 2009-10 to 2013-14. Starting in 2014-15, high schoolers were required to take the Smarter Balanced Assessments (SBAs) in 11^{th} grade until the 2017-18 school year when the state began requiring students to take these tests in 10^{th} grade. (See https://ospi.k12.wa.us/student-success/testing/state-testing for more details.)

College-level course-taking. If a course is designated as Advanced Placement, College in High School, International Baccalaureate, Running Start, or part of the Cambridge Scholars Program we considered this an "college-level" course.[4] We sum the number of unique college-level courses taken by each student during their high school career (grades 9-12) for math and ELA separately.[5] Students enrolled in any grade in high school are included in this measure.[6]

Absences. We calculate a weighted measure of student absenteeism by summing the number of full-day absences (either excused or unexcused) and half the value of the number of part-day absences for each student during their high school career (grades 9-12).[7] Students enrolled in any grade in high school are included in this measure.

Disciplinary Incidences. We sum the number of all reported disciplinary incidences for each student during their high school career (grades 9-12). Such incidences involved disruptive behavior, bullying, drug or alcohol use, possession of a weapon, etc. and resulted in interventions ranging from no action to classroom exclusion, long-term suspension or expulsion.[8] Students enrolled in any grade in high school are included in this measure.

GPA. We use students' cumulative GPA in 12th grade as reported on the state's standardized transcript. Values range from 0.0 to 4.0. Students enrolled in 12th grade are included

[4] This definition does not include Honors courses or other course levels past what are required for HS graduation that do not have a college dual credit designation (i.e., Calculus). As in prior work that finds 3rd grade tests to be predictive of taking a college-level math course in high school (Goldhaber et al., 2021), we use high school course names to identify college-level coursework in both math and ELA.
[5] We count each unique college-level course taken by a student in a given year. A small share of students (3%) retake the same course in different years; for simplicity, we treat each instance in a different year as a separate course when tallying total college-level courses.
[6] In our primary analytic sample, 89% of students were enrolled in all four years of high school, 6% were enrolled in any three years, 3% in any two years, and 2% in a single year.
[7] Since students in the earliest cohort (2006) were in 9th grade in 2012 and the state had not yet started differentiating absences by full- and part-day, we can only use "full day" absences for that cohort and grade.
[8] We can observe these intervention types separately. When we replicate models predicting all disciplinary incidences with models that predict only suspensions or only expulsions we find qualitatively similar results.

in this measure regardless of whether we observed the student in other grades throughout high school.

Graduation. We construct a binary measure of whether students graduated high school on time, i.e., within four years after entering high school.[9] Results from models below are qualitatively similar when we look at whether students graduated within 5 years after entering high school (available from authors upon request).

3.2 Analytical Samples and Sample Statistics

Our primary analytic sample is anchored to eight 3rd grade cohorts of students from 2006 to 2013 for whom we observe 3rd grade demographics as well as 3rd grade test scores in both math and ELA.[10] The earliest cohort entered 3rd grade in 2006, 9th grade in 2012, and 12th grade in 2015. The most recent cohort entered 3rd grade in 2013, 9th grade in 2019, and 12th grade in 2022. Across all eight cohorts, our unrestricted sample includes 581,094 unique students. Part of our analysis looks at the progression of academic mobility from grade to grade and students can appear multiple times in those subsamples which have between 597,052 to 2,745,233 *student-year* observations depending on the outcome.[11]

[9] We use the school withdrawal code in the CEDARS data to identify students who graduated on-time with a regular high school diploma, an associate degree, an International Baccalaureate diploma, a regular high school diploma with modifications identified in a student's Individualized Education Program (IEP); and those who received a General Education Development (GED) certificate.

[10] Our sample does not include 3rd graders in any cohort who took any "alternative" test in those school years (i.e., students with a cognitive disability documented in their Individual Education Plan—2.6% of all 3rd graders across cohorts) or no test at all (due to limited English proficiency, medical exemption, absence, incompletion, or refusal—7.4% of all 3rd graders across cohorts). Of the remaining 3rd graders (n=589,579), we further limit the sample by excluding students for whom we observe test scores in only one subject (9.9%) and students who have missing demographic characteristics (0.2%). With these restrictions, our sample consist of students who are, on average, very slightly better performing—by 0.4% and 0.1% of a standard deviation in math and ELA, respectively—than a sample that would also include single-subject test-takers in 3rd grade.

[11] We look at progression in academic mobility from grade to grade for four outcomes that we can observe in multiple grades from 3rd grade to high school: math test scores (421,829 students; 2,098,169 student-years), ELA test scores (422,153 students; 2,151,164 student-years), absences (359,601 students; 2,745,223 student-years), and disciplinary incidences (73,290 students; 597,052 student-years). We cannot do this for other outcomes because they only exist at the high school level.

In Table 2 we present summary statistics for various student characteristics in 3rd grade, broken out by race/ethnicity and gender.[12] Panel A focuses on student classifications and Panel B on average subject-specific test scores. Overall, approximately 40% of our sample consists of students of color, nearly half of whom are Hispanic. Results highlight the differences in means between the overall sample and each subgroup. For example, in Panel A we see that compared to all students, White and Asian students were less likely to be eligible for free or reduced-price lunch, males were more likely to be identified for special education services, and Black and Hispanic students were less likely to be identified for the Highly Capable Program (the advanced learning designation in Washington).

Turning to Panel B, we see, for example, that both Black and Hispanic students scored nearly half a standard deviation below the overall student test average in math, whereas White and Asian students scored well above it (by 15% and 39% of a standard deviation, respectively). In other words, on average, the math test score gaps between Black or Hispanic students and their Asian or White peers exceed half a standard deviation. Gender disparities are also evident, particularly in ELA, where female students outperformed male students by 12% of a standard deviation; the gender differences in math were negligible.[13]

---

[12] We group students into five racial/ethnic categories. Students in the "Other" racial/ethnic category include those who identify as American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, and/or multiracial students. If in the text, we refer to students of these races/ethnicities simply as "Other" it is purely for efficiency in writing.

[13] The regression results in Appendix Table A.2 show that these patterns persist even after accounting for various student characteristics in a multivariate setting. Controlling for other factors, Asian students outperformed White students (the referent category) by approximately a quarter of a standard deviation in 3rd grade math and 9% of a standard deviation in ELA. In contrast, Black, Hispanic, and students of other races/ethnicities scored lower, with math gaps ranging from 18% (Other) to 41% (Black) of a standard deviation and ELA gaps ranging from 16% (Other) to 33% (Black). Female students performed slightly worse than males in math (2% of a standard deviation) but much better in ELA (17% of a standard deviation). Additionally, students eligible for free or reduced-price lunch, those receiving special education services, students with learning disabilities, and English learners all scored between a quarter and over half a standard deviation lower in both subjects. Meanwhile, students identified as highly capable scored nearly a full standard deviation higher than their peers (in math).

We do not observe all students in both 3rd grade and high school as students may exit the Washington state public school system or leave the state before we observe their high school outcomes. And, as the descriptions of data measures in Section 3.1 make clear, there are some high school outcomes that we do not observe for entire cohorts. This is because of changes to statewide testing policy, the COVID-19 pandemic and different starting dates for the collection of specific data elements. Table 1 provides more details on all the outcome variable definitions, years of data availability and variable construction and Appendix Table A.1 shows the specific 3rd grade cohorts for which high school outcomes could be observed. High school test scores, for example, are available for all cohorts except for the 2006 and 2013 cohorts. Due to a change in the statewide testing regime, students in the 2006 cohort were not required to take the SBA (which began testing 11th graders in the 2014-15 school year, i.e., the year those students entered 12th grade). And students in the 2013 cohort were not required to take the SBA in high school (as 10th graders) due to the COVID-19 pandemic. Likewise, the 2006 and 2007 cohorts only contribute two and three years, respectively, to our measures of disciplinary incidences in high school because when the state started collecting disciplinary data in 2013-14, students in the 2006 cohort were already in 11th grade and students in the 2007 cohort were in 10th grade.

Of the 581,094 3rd graders in our unrestricted sample, we observe different subsamples of students with at least one of the following high school outcomes: math test scores (N=269,494), ELA test scores (N=320,509), college-level math and ELA course-taking (N=521,391), absences (N=521,391), disciplinary incidences (N=517,319), high school GPA (N=467,162), and high school graduation (N=464,855). Students who we observe in 3rd grade, but for whom we *do not* observe high school outcomes are notably different from the unrestricted sample of 3rd graders. We illustrate this in Table 3 which shows the *differences* in means of 3rd grade student characteristics between samples of *3rd graders* we observe with a given high school outcome and those 3rd graders we do not observe with a given high school outcome. So, for instance, a positive value on 3rd grade tests indicates that those students who are in the sample in both periods have higher 3rd grade test scores than those who are only observed in the 3rd grade. Likewise, higher values on a demographic characteristic indicate that demographic is more likely to be present in the sample of students we observe in both 3rd grade and high school.

Across outcomes, there are a number of statistically significant differences in student demographics (Panel A) and classifications (Panel B) between the two subsamples. Notably, there are more Hispanic students (by 5 percentage points) and fewer Black students (by 3 percentage points) in our observable subsamples for college-level course-taking (Columns 3 & 4), absences (Column 5) and disciplinary incidences (Column 6). And our observable subsamples for GPA (Column 7) and on-time graduation (Column 8) include fewer students who are eligible for free or reduced-price lunch by 10 and 8 percentage points, respectively.

In terms of 3rd grade test performance (Panel C), students we can observe with high school outcomes generally have higher 3rd grade tests scores than students for whom we do not observe high school outcomes. For example, students we observe with high school ELA tests (Column 2) scored 5.5 and 4.7 percentage points higher on their 3rd grade math and ELA tests, respectively, compared to students for whom we do not observe high school ELA tests. These findings are generally consistent prior research that found more mobile students (i.e., observed in 3rd grade but not high school) tended to be more disadvantaged (Goldhaber et al., 2022).

In Table 4, we show summary statistics for test and non-test *high school outcomes* for the analytic samples we evaluate. Consistent with evidence on high school test (Betts et al., 2003; Clotfelter et al., 2009; Fortin & Phipps, 2015; Reardon, 2011; Reardon et al., 2019) and non-test

(Gamoran, 1987; Gregory, Skiba, & Noguera, 2010; Gottfried, 2014; Kelly, 2009; Lee, 2002; Losen et al., 2015; Malkus, 2024; Sorenson, 2019; Skiba et al., 2011) gaps in educational outcomes, we observe large differences across student demographics in outcomes. For instance, Panel A shows that, on average, Asian students had standardized test scores in math that are about 0.5 standard deviations higher than White students and about a full standard deviation higher than Black and Hispanic students. We see similar gaps in non-test outcomes. Panel B shows, for example, Asian students took nearly three times as many college-level math courses and more than twice and many college-level ELA courses than White students. Compared to Black and Hispanic students, Asian students had about half as many absences and were more likely to graduate on-time by at least 13 percentage points. Gender differences also exist and almost always favor females, who, for example, had higher GPAs by 0.3 points and fewer than half as many disciplinary incidences than male students.

4. Analytic Approach

To assess the extent to which students' earlier-grade (3rd or 8th) characteristics predict various test and non-test outcomes in high school (RQ 1) we employ multivariate regressions of various types depending on the outcome of interest. More specifically, to measure the relationship between 3rd (or 8th) grade test scores and high school test scores or non-test outcomes we model some function of an outcome using equation (1):

$$f(Y_i) = \alpha_0 + \beta T_i + \tau X_i + \varepsilon_i \tag{1}$$

where Y_i is a high school outcome of interest (i.e., test scores, GPA, absenteeism, etc.), T_i is the 3rd (or 8th) grade test score, and β is a prime coefficient of interest as it measures the relationship between a specific high school outcome variable and earlier-grade test scores. X_i is a vector of control variables that includes student demographics (i.e., race/ethnicity and gender) and student classifications in 3rd grade (i.e., free or reduced-price lunch eligibility, identification for special education or highly-capable services, disability status, and whether students received English-learner services). In addition to OLS for linear outcomes, we use logistic regression to predict the binary outcome of whether a student graduated on time from high school and negative binomial Poisson regressions for count outcomes (i.e., the number of college-level courses, absences, and disciplinary incidences).[14]

We then examine the extent to which patterns of academic mobility between 3rd grade test scores and high school outcomes vary by intersectionalities of students' race/ethnicity and gender (RQ 2).[15] To do this, and akin to Austin et al. (2023), we rank students' 3rd grade performance as well as their various high school outcomes by calculating percentiles of each measure.[16] Then we use linear regressions to predict students' *rankings* on high school outcomes with their *rankings* of 3rd grade test performance. We estimate these rank-to-rank models for each high school outcome for our entire sample and for each student subgroup as summarized by equation (2):

$$O_{is}^{r} = \alpha_s + \beta_s R_{is} + \varepsilon_{is} \tag{2}$$

where O_{is}^{r} is a ranked high school outcome for student *i* in subgroup *s*, $R_{is} = r$ is the rank of their 3rd grade test score, and β_s captures the relationship, by subgroup, between the early- and later-year ranked measure. When predicting subject-specific high school tests (i.e., math or ELA)

[14] For each of these count outcomes the variances are larger than the means suggesting overdispersion and the need for a negative binomial regression to model that overdispersion.

[15] We categorize our data into five racial/ethnic groups (Asian, Black, Hispanic, Other, and White) and two genders (female and male) for a total of 10 subgroups based on intersectionalities. In the analyses that follow, we use White students as the referent among racial/ethnic groups and males as the referent group for genders. For parsimony's sake we do not look at academic mobility by students' socioeconomic status as well. However, we include eligibility for free and reduced-price lunch as a covariate in all our regression models for each high school outcome.

[16] To obtain rankings for the binary outcome of high school graduation, we calculate percentiles of the average probability of graduating from each percentile of the 3rd grade test score distribution.

we use subject-specific 3rd grade scores. For all other outcomes we use the averaged standardized scores of students' 3rd grade math and ELA tests.

Similar to Austin et al., we focus on the academic mobility of initially low-achieving students. And while there is nothing inherently meaningful about the 25th percentile, using it facilitates comparisons relative to Austin et al. Thus, as an additional step, we calculate the predicted outcomes of various ranked high school outcomes for students who scored in the 25th percentile on their 3rd grade tests, as summarized in equation (3):

$$\hat{O}_s^{25} = \alpha_s + \beta_s * 25 \tag{3}$$

where $\hat{O}_s^{25}$ is the predicted average high school percentile ranking for initially low-achieving 3rd graders in a given subgroup s.[17]

For some analyses (RQ 3), we examine the *progression* of academic mobility from grade to grade, by which we mean the successive movement in an outcome distribution relative to one's peers. We do this by estimating versions of equations (2) and (3)—separately, by subgroup and grade—that use students' 3rd grade test percentile rankings, R_{is}, to predict their percentile rankings in each grade from 4th grade through 12th grade. Then we calculate the predicted outcomes at each grade level for initially low-performing 3rd graders. This approach enables us to capture the progression in academic mobility from one grade to the next—which ultimately determines students' academic mobility from 3rd grade to high school—and identify subgroup differentials in mobility trajectories.

Note that values of $\hat{O}_s^{25}$ above 25 indicate upward mobility and values below 25 indicate downward mobility. This is the case for all but two of the outcomes we examine: absenteeism

[17] For every model we also estimated versions of (2) and (3) with controls for students' (3rd grade) classifications. Results are qualitatively similar to what we report here and are available from the authors upon request.

and disciplinary incidences. For these outcomes, the interpretation of $\hat{O}_s^{25}$ is flipped in that *higher* values indicate a more negative outcome, or *downward* mobility.

As discussed in Austin et al., the initial percentile ranks are based on 3rd grade scores from high-stakes state tests, which, while reliable, are not error-free (Boyd et al., 2013; Lockwood & McCaffrey, 2014). Measurement error arises from two main sources: (1) a limited number of test items causing score variability, and (2) random factors that affect performance, such as testing conditions. These errors in students' initial scores will result in mean reversion, and, if left unaccounted for, will lead us to overstate academic mobility. We follow the error correction mechanisms used in Austin et al. First, to set students' initial rank, we average their 3rd grade math and ELA scores thereby reducing the error variance of the two noisy measures.[18] Second, we make an additional correction to remove error due to the sampling variance in the items that appear on the test. This correction relies on test reliability ratios provided by OSPI, which we incorporate into our models using a standard errors-in-variables (EIV) regression framework.[19] This framework adjusts for attenuation bias in our estimates by accounting for measurement error, which in turn reduces the bias in the parameter estimates. Specifically, the error variance is subtracted from the total variance in the initial ranks to calculate the error-adjusted parameter. A complication arises because in models predicting non-test outcomes we use the average of entry ranks, though reliability ratios are for individual tests.[20] So, following

[18] Austin et al. set students' initial rank by averaging their math and ELA test *ranks,* not scores. When we do this as well, results are nearly identical to what we report below.
[19] OSPI maintains annual technical reports on all current and historical statewide tests. These reports include grade- and subject-specific reliability estimates based on measures of internal consistency that are calculated using Cronbach's coefficient alpha. Across all the years in our sample, the reliability of 3rd grade test scores ranged between 0.84 and 0.89 in ELA and either 0.87 or 0.88 in math.
[20] When using only 3rd grade math or ELA scores as predictors we apply the reliability estimates for math and ELA individually.

Wang and Stanley (1970), and as suggested by Austin et al., we calculate the reliability of the average performance across math and ELA tests as shown in equation (4):

$$r_c = \frac{(0.25rm + 0.25re + 0.50\theta m,e)}{(0.50 + 0.50\theta m,e)} \tag{4}$$

We average the annual reliability estimates reported by OSPI across all cohorts to obtain reliability weights of 0.88 for math tests and 0.86 for ELA tests. And using equation (4) we calculate a reliability weight of 0.92 for the average of both tests.

5. Findings

In this section we present findings about the achievement of students as they progress through Washington state's public schools. We begin in Section 5.1 by establishing baseline evidence about the extent to which early-grade student characteristics and achievement levels—i.e., 3rd grade test achievement, 8th grade test achievement, and 3rd grade characteristics—predict high school outcomes. Then, in Section 5.2 we describe the academic mobility for initially low-achieving students (those who scored at the 25th percentile of the averaged 3rd grade test distribution) from 3rd grade to high school for various outcomes and the degree to which mobility varies across subgroups. Finally, in Section 5.3 we look at the low-achieving students' progression in academic mobility as they move from grade to grade.

5.1 Earlier-Grade Achievement and Characteristics as Predictors of High School Outcomes

We report how earlier-grade (3rd or 8th) achievement and student characteristics predict various high school outcomes in Table 5 (for test outcomes) and Table 6 (for non-test outcomes). In Table 5 we estimate the results using math and/or ELA test scores from either 3rd or 8th grade. For parsimony's sake, Table 6 only reports results from models that include both 3rd grade math and ELA tests. But, as we discuss below, additional model specifications for non-test outcomes are included in the appendix.

For test and non-test outcomes, coefficients on student characteristics are broadly similar whether 3rd or 8th grade tests are used as controls. But, as expected, because 8th grade tests are closer in time to high school outcomes—and because test score patterns shift across student demographics from 3rd to 8th grade (discussed below)—the magnitudes of the coefficients for student characteristics are smaller in models with 8th grade tests than in those with 3rd grade tests.

High School Test Score Outcomes

The results for test scores (in Table 5) show, unsurprisingly, that students who performed better on 3rd grade math and ELA tests also scored better in those subjects on their high school tests. Recall that because tests scores are normalized, we can interpret the coefficients on the base year (3rd or 8th) test score in standard deviation units. In models with only *same-subject* tests, a one standard deviation increase in 3rd grade tests scores is associated with about one half of a standard deviation increase in high school test scores for both math (Column 1; $\hat{\beta}_{math}$=0.544) and ELA (Column 5; $\hat{\beta}_{ela}$=0.504). When instead 8th grade same-subject tests are used as predictors (Column 2 for math and Column 6 for ELA), the coefficients are roughly 40% larger.[21]

To better understand the implications of using 3rd grade tests versus 8th grade tests to predict high school outcomes, we compare same-subject model specifications that predict high school test scores with either earlier-grade test in both subjects. Compared to models with 3rd grade predictors, models with 8th grade predictors have larger values of R^2 by 43% for math

[21] In our main specifications we use 3rd grade student classifications (e.g., special education status) because 8th grade classifications are likely endogenous since they may depend on third grade achievement. But, if instead we include 8th grade classifications results are qualitatively similar.

(compare Columns 1 and 2) 32% for ELA (compare Columns 5 and 6). And while the 3rd and 8th grade coefficients differ, the predictions for high school test scores generated from the different specifications are similar, i.e., the correlations of the predicted values of high school test outcomes using the 3rd and 8th grade specifications are about 0.75 for both subjects.

As expected, when we estimate models that include both same- and cross-subject tests for 3rd grade (Columns 3 for math and 7 for ELA) or 8th grade (Columns 4 for math and 8 for ELA), we see within-subject effects (e.g., the coefficient of early-grade math tests on high school math tests) are larger than cross-subject effects (i.e., the coefficient of early-grade math tests on high school ELA tests or vice versa), although both still predict high school test outcomes. Yet, interestingly, 3rd grade math scores are better predictors of high school ELA scores (Column 7; $\hat{\beta}_{math}$=0.263) than 3rd grade ELA scores are of high school math scores (Column 3; $\hat{\beta}_{ela}$=0.194). The same is true of models with 8th grade test scores (Columns 4 & 8). As we discuss below, math tests are nearly always more predictive of non-test outcomes than ELA tests—an interesting finding, since 3rd grade retention policies typically focus on reading proficiency, the weaker of the two predictors.

Consistent with prior evidence, we observe large differences in high school test achievement by student subgroup, even after controlling for same- and cross-subject test scores. For example, in models with 3rd grade test scores as predictors (Columns 3 & 7), Asian students out-performed White students (the referent category) on high school tests by 37% of a standard deviation in math and 25% of the standard deviation in ELA. All non-Asian student groups scored lower than White students by as much as 14% of a standard deviation in math and 11% of a standard deviation in ELA.

And—as was the case above with early-grade test scores—the magnitudes of these ethno-racial disparities are diminished in models that include 8th grade test scores as predictors (Columns 4 & 8).[22] The reason for this is shown in Figure 1.[23] For each subgroup, these "performance pyramids" show the percentage of students in each averaged test score decile in 3rd grade (lighter, left-hand bars in each panel) and in 8th grade (darker, right-hand bars in each panel).[24] If the students in a subgroup were equally distributed across test score deciles (i.e., 10% in each decile) in 3rd and 8th grade, we would see perfect a perfect rectangle (as in Panel A). The subgroup that most closely approximates that shape is Other (Panel E). For all other subgroups, we see pyramid-like shapes. Black (Panel C), Hispanic (Panel D) and FRPL students (Panel G) are overrepresented in the lower deciles across grades. In contrast, Asian (Panel B), Non-FRPL (Panel H) and, to a much lesser extent, White students (Panel F), are overrepresented in higher deciles, inverting the pyramid.

The distribution of students across test score deciles in the performance pyramids varies noticeably from 3rd to 8th grade for some subgroups, as represented by its left-right asymmetry. For instance, from 3rd to 8th grade, the proportion of Asian students in the top decile expands from 18 to 25% whereas the bottom performance decile of Other expands from about 12 to 14%. It is these shifts in test performance as students progress through school that explain the attenuation in the racial/ethnic coefficients when we use 8th grade tests as opposed to 3rd grade

[22] This pattern of attenuation in coefficients also holds for nearly all student classifications. For example, in 3rd grade models (Columns 3 and 7), students eligible for free or reduced-price lunch performed worse than their more socioeconomically advantaged peers, by about one quarter of a standard deviation in both subjects. In 8th grade models (Columns 4 and 8), these coefficients were about half as large.
[23] Note that while we do not discuss students who are and are not eligible for free or reduced-price lunch (FRPL) in this paper, we include those subgroups here to serve as points of reference because their performance over time is highly correlated with the performance of some racial/ethnic subgroups.
[24] These performance pyramids include all students for whom we observe tests in either 3rd or 8th grade, but findings are qualitatively similar when we restrict the sample to only 3rd graders we can also observe with test scores in 8th grade. Results are also qualitatively similar when we look math and ELA scores separately instead of averaged test scores.

tests a covariate; the coefficient for students' racial/ethnic categories in the 3rd grade models reflects everything that happens to students connected to their category between 3rd and 8th grade, including their improved or diminished test performance.

High School Non-Test Outcomes

Next, we examine the predictive power of students' 3rd grade test scores and characteristics as they relate to non-test outcomes in high school including: college-level course-taking, absenteeism, disciplinary incidents, GPA, and on-time graduation.[25] In Table 6, we report the average marginal effects from models that include both 3rd grade math and ELA tests as controls. In the appendix, we report additional specifications that include only 3rd grade same-subject tests as controls (Table A.3 for math and Table A.4 for ELA) and specifications that use 8th grade predictors from both subjects (Table A.5). The findings from the supplemental models in the appendix are qualitatively similar to what we describe below, except where noted.

Consistent with our findings for test outcomes, the coefficients for student classifications in models predicting non-test outcomes are generally smaller in magnitude when models include 8th grade controls instead of 3rd grade controls.[26] And, as was the case for test outcomes, the predicted values from non-test model specifications that include 3rd grade tests as predictors (Table 6) and those that include 8th grade tests (Appendix Table A.5) are highly correlated, ranging from 0.76 for on-time graduation to 0.82 for disciplinary incidents. Likewise, values of pseudo R^2 are higher in models with 8th grade predictors by between 26% (for college-level ELA course-taking) and 64% (for absences.) Though, notably, pseudo R^2 values from models for non-

[25] As we described in Section 4, our model specifications vary depending on the outcome. We estimate non-linear models with negative binomial (for count outcomes) or logistic (for graduation) regressions. For outcomes with many observed zeros (i.e., college-level course-taking and disciplinary incidences), we also estimated zero-inflated Poisson regressions (not shown but available from authors upon request) which closely mirror the results from the negative binomial regressions.

[26] Though this is not the case for students identified for special education who generally have larger coefficients in models with 8th grade predictors.

test outcomes—regardless of whether 3rd or 8th grade tests are used as predictors—are much smaller than R^2 values from model predicting test outcomes.

Turning to Table 6, students who did better on tests in the 3rd grade had better (or fewer "negative" outcomes in the case of absences and disciplinary incidences[27]) non-test high school outcomes. For instance, compared to students at the 25th percentile, students with average (50th percentile) 3rd grade math and ELA test scores were predicted to take about 0.12 more college-level courses in either subject, have 7.4 fewer absences and 0.18 fewer disciplinary incidences throughout high school. They were also predicted to have higher 12th grade GPAs by 0.20 points (or 22% of standard deviation) and be more likely to graduate high school in four years by 4.7 percentage points (as a point of reference, the average on-time graduation rate is 86% across all cohorts).

Notably—as was the case with test outcomes—3rd grade math tests are generally more predictive of non-test high school outcomes than 3rd grade ELA tests. Indeed, the only case in which 3rd grade ELA tests are more predictive than math tests is for the number of college-level ELA courses taken in high school (Column 2). But even for this outcome, the difference in magnitudes between the math and ELA coefficients is small.

As was the case with high school tests, there are large demographic differences across non-test outcomes. Looking across ethno-racial subgroups, we see that, relative to White students, Asians were predicted to have significantly better high school outcomes across the board—taking, for example, 0.35 more college-level math courses (Column 1), having 23.4 fewer absences (Column 3) and being 9 percent more likely to graduate on time (Column 6). Black students were also predicted to take more college-level courses than their White peers

[27] The results are consistent if we look at suspensions and expulsions separately. Notably, 64% of all disciplinary incidents resulted in a suspension. Only 3% resulted in an expulsion.

(Columns 1 and 2). But Black students were predicted to have 12 more absences (Column 3), 0.25 more disciplinary incidents, and lower GPAs by 0.12 points (Column 5). Compared to males, females were predicted to take 0.15 more college-level ELA courses (Column 2) and have higher GPAs by 0.26 points (Column 4)—but they were predicted to have 3.4 more absences than males (Column 3).

Unsurprisingly, there are also very large differences in high school outcomes for students with different 3rd grade classifications (free or reduced-price lunch, learning disability status, etc.). At one extreme, students eligible for free or reduced-price lunch face the most challenging prospects across nearly all measures, with predictions of 24.7 more absences, 0.40 more disciplinary incidents, and GPAs that are 0.32 points lower than their economically advantaged peers. At the other extreme, students identified as highly capable are found to have the best non-test high school outcomes; they are predicted to take 0.19 more college-level math and 0.17 more college-level ELA courses, have 7 fewer absences, and be 3 percent more likely to graduate on time relative to those not identified as highly capable.[28]

Students with limited English proficiency follow a similar but less pronounced pattern of disadvantage. However, the picture is more complex for students receiving special education services. While these students are predicted to take fewer college-level courses and be less likely to graduate on time, they show better attendance and higher GPAs than their non-special education peers.

5.2 Assessing Academic Mobility Across Subgroups From 3rd Grade to High School

Building on Austin et al.'s work, we examine the extent to which initially low-performing students (i.e., those who scored in the 25th percentile on their 3rd grade tests) moved

[28] In models using 8th grade controls (Appendix Table A.5), students identified as highly capable were predicted to have significantly more absences (Column 3) and lower GPAs (Column 5).

up or down (or not at all) in distributions of test and non-test outcomes, i.e., their academic mobility from 3rd grade to high school. We do this by plotting values of $\hat{O}_s^{25}$ derived from linear rank-to-rank EIV regressions as shown in equations (2-3), i.e., the predicted outcomes of various ranked high school outcomes for students who scored in the 25th percentile on their 3rd grade tests.[29] In Appendix Table A.6 we report the beta coefficients and intercepts from EIV regressions for each outcome and student subgroup used to derive $\hat{O}_s^{25}$.

To facilitate comparisons across subgroups, in Figures 2-5 we plot estimates of $\hat{O}_s^{25}$ (and their 95% confidence intervals) for selected high school outcomes by student subgroup. In all the figures, Panel A shows results for females (circle symbols) and Panel B shows results for males (diamond symbols). Racial/ethnic groups are color-coded: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). We also plot the estimates of $\hat{O}_{f,m}^{25}$ for all students *by gender* (solid black lines in each panel). A dashed horizontal line marks the 25th percentile of each outcome variable, when scaling permits. Generally, estimates above this line reflect upward mobility and those below it reflect downward mobility. However, for absenteeism and disciplinary incidents—where higher values are undesirable—the interpretation is reversed. Estimates that cross the line indicate no statistically significant mobility.

Academic Mobility in Test Scores

Consistent with Austin et al., we find that by high school low-achieving 3rd graders experienced modest upward test-based academic mobility (signified by $\hat{O}_s^{25}$ values greater than 25). Our estimates for all students predict increases of roughly 6 percentage points for math

[29] We also estimate more flexible specifications that allow for non-linearity between ranked 3rd grade test scores and ranked high school outcomes akin to student growth percentile models. When comparing estimates of $\hat{O}_s^{25}$ from EIV regressions to the average high school percentile of students who performed at the 25th percentile in 3rd grade, the correlation across all subgroups is at least 0.95 for each outcome, except for on-time graduation for which the correlation is at least 0.90. These results are available from the authors upon request.

($\hat{O}^{25}_{all}$=30.8) and ELA ($\hat{O}^{25}_{all}$=31.4) tests.[30] Notably, while each student subgroup experienced some upward mobility in their test scores, the magnitudes varied significantly across subgroups. For example, Figure 2 shows that in math, Asian females ($\hat{O}^{25}_{af}$=42.1) and Asian males ($\hat{O}^{25}_{am}$=40.5) are estimated to have improved (moved up in the ranked test score distribution) by 17 and 16 percentage points, respectively. Meanwhile, Black males, by contrast, experienced far less upward mobility as they are estimated to have moved up by less than one percentage point ($\hat{O}^{25}_{bm}$=25.9)—over 4 percentage points lower than the average mobility for males. Results for high school ELA tests are qualitatively similar and are shown in Appendix Table A.1.

Gender gaps in test score mobility exist within each racial/ethnic subgroup for both math and ELA tests and consistently favor females. Gender gaps were the largest among Black students in both math (3.1 percentage points) and ELA (4.6 percentage points). At the other extreme, the gaps were as small as 1.1 percentage points in math for students from Other races/ethnicities. Although females were, on average, more upwardly mobile than males (compare solid black lines across panels), Asian males (purple diamond symbol) were much more upwardly mobile than the average female.

Academic Mobility in Non-Test Outcomes

Subgroup mobility patterns are more varied for non-test outcomes. For GPA and graduation, patterns mirror those of test scores—we see upward mobility for each subgroup.[31] For example, all students were upwardly mobile for GPA (Figure 3), with an estimated average increase of 12.1 percentage points ($\hat{O}^{25}_{all}$=37.1); females ($\hat{O}^{25}_{f}$=41.5) were predicted to be

[30] Our estimates of test-based academic mobility closely mirror Austin et al.'s who report (in their Table 4) that $\hat{O}^{25}_{all}$=29.68 in averaged high school test scores for students in Washington state.
[31] Our estimate of academic mobility in on-time graduation for all students ($\hat{O}^{25}_{all} = 77.8$) is generally consistent with Austin et al. who report (in their Table 4) $\hat{O}^{25}_{all}$ =71.15 for students in Washington state, and may be larger due to our including more recent cohorts with higher rates of graduation.

significantly more upwardly mobile than males ($\hat{O}_{m}^{25}$=33.4) by 8 percentage points (compare difference between solid black horizontal lines across panels). Indeed, all female subgroups were predicted to be more upwardly mobile than every non-Asian male subgroup. As was the case for test scores, the magnitudes of academic mobility vary significantly across subgroups, with gaps as large as 23 percentage points (when comparing Asian females ($\hat{O}_{af}^{25}$=52.8) to Black males ($\hat{O}_{bm}^{25}$=29.5)). Qualitatively similar results for on-time graduation are reported in Appendix Figure A.2.

By contrast, we see downward mobility for each student subgroup for two outcomes: college-level math course-taking (a negative trend, indicating fewer college-level courses) and absenteeism (a positive trend, indicating worse attendance). For college-level math course-taking (Figure 4), on average, all students were downwardly mobile by 17.2 percentage points—though some subgroups were less downwardly mobile than others. Asians females ($\hat{O}_{af}^{25}$=18.5) and males ($\hat{O}_{am}^{25}$=18.6), for example, were downwardly mobile by about 6.5 percentage points while White females ($\hat{O}_{wf}^{25}$=5.4) and males ($\hat{O}_{wm}^{25}$=5.5) were downwardly mobile by three times that amount (about 19.5 percentage points). Qualitatively similar results for absenteeism are reported in Appendix Figure A.3.

For the remaining two non-test outcomes—i.e., college-level ELA course-taking and disciplinary incidences—the direction of academic mobility (either upward or downward) varies across student groups. Consider disciplinary incidences (Figure 5). Because higher incidence rates represent worse outcomes, the interpretation is reversed from the other outcomes. Here, estimates above the dashed horizontal line indicate downward mobility (worsening outcomes), while estimates below the line indicate upward mobility (improved outcomes). All females except Black females ($\hat{O}_{bf}^{25}$=26.5), experienced upward mobility ranging from 1.6 (Other females) to 16.4 percentage points (Asian females). In contrast, most males, except for Asian males ($\hat{O}^{25}_{am}$=15.8), showed downward mobility ranging from 3.3 percentage points (White males) to 14.9 percentage points (Black males).[32] Qualitatively similar results for ELA course-taking are reported in Appendix Figure A.4.

Gender gaps within racial ethnic groups for non-test outcomes generally follow the pattern we observed for test outcomes, varying in size across racial/ethnic groups. The gender gap is largest among Black students for nearly every non-test outcome (except GPA and absences). In cases where there are statistically significant gender differences, females were nearly always more upwardly mobile than males. The one outcome where this is not true is absenteeism. Here, Hispanic, Other, and White females were significantly more downwardly mobile (i.e., were absent more often) than their male counterparts. The gaps are small, however: less than 2.2 percentage points for each subgroup.

Across all outcomes, Asian students consistently have the most favorable academic mobility patterns compared to other racial/ethnic groups. Within this high-performing group, Asians females show even more upwardly mobility than Asian males for most outcomes, except college-level math course-taking and absences where confidence intervals overlap. By contrast, no single subgroup consistently ranks lowest in academic mobility. Black males show the least upwardly mobility for half of the outcomes (i.e., math and ELA test scores, disciplinary incidences and GPA). White students of both genders were the least upwardly mobile in college-level math course-taking, while White males were the least upwardly mobile in college-level

[32] When looking at suspensions and expulsions separately, the mobility across subgroups follows the same patterns as overall disciplinary incidences. However, it is important to note here that overall rate of expulsion is lower for all racial/ethnic groups as only 3% of the disciplinary incidences lead to an expulsion and values of $\hat{O}^{25}_{s}$ are never above 7 for any subgroup.

ELA course-taking. Finally, Hispanic females were the least upwardly mobile for attendance outcomes.

5.3 Assessing Grade-to-Grade Progression in Academic Mobility

In the previous section we looked at students' academic mobility between two points: 3rd grade to high school. But what is the evolution of academic mobility between 3rd grade and high school? As we discuss in Section 4, we address this question by focusing on the outcomes we can observe in multiple grades from elementary to high school: test scores (from grades 4 to 8 and again in either 10th or 11th grade), absences (grades 4 to 12), and disciplinary incidences (grades 4 to 12).[33] We estimate versions of equations (2) and (3)—separately, by subgroup and grade—that use students' 3rd grade test percentile rankings, R_{is}, to predict their subject-specific test percentile rankings in each tested grade from 4th grade through high school and their percentile rankings in terms of absenteeism or disciplinary incidences from grades 4 to 12. Recall that this approach enables us to look at grade-to-grade *progression* in academic mobility. It also allows us to identify subgroup differentials in such progression.

In Figure 6 we show progression in academic mobility from grade to grade in math test scores for low-achieving 3rd graders (i.e., values of $\hat{O}_s^{25}$).[34] Estimates are shown for females in Panel A (circle symbols) and males in Panel (diamond symbols). Racial/ethnic groups are color-coded as follows: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). Estimates for all students, by gender, are shown via the thicker black trendlines of each panel. A dashed horizontal line is plotted at the 25th percentile of the grade-level test score outcome.

[33] We observe six cohorts (2007-2012) for math and ELA test outcomes, five cohorts for absences (2009-2013) and one cohort for disciplinary incidences (2013). As we note above in Section 3.2, we cannot investigate the other non-test outcomes (i.e., college-level course-taking, GPA, graduation) in this manner because those data only exist at the high school level.

[34] Results are qualitatively similar for ELA tests and are reported in Appendix Figure A.5.

Estimates above this line reflect upward academic progression from 3rd grade to each subsequent grade; estimates below this line reflect downward academic progression.

Several findings are of note. First, the upward slope of the black trendlines in either panel indicate that, on average, initially low-performing students tended to gain in the math test score distribution as they advanced from grade to grade. This is particularly true for females who have steeper slopes from grade to grade. Second, there are places in the distribution where the gains look more dramatic, e.g., from 7th to 8th grade for all female subgroups and for Asian males. Third, and arguably the most striking finding, is the separation of Asian students from other ethno-racial group in terms of their making upward progress. A gap exists in 4th grade and grows, for the most part, from grade to grade. Indeed, in 4th grade, across all subgroups, the gap between the *most* upwardly progressing group (Asian) and the *least* upwardly progressing group (Black) was 4.9 percentage points. By high school, that gap was over 3 times as large at 18 percentage points. Notably, the biggest grade-to-grade divergence was between 5th and 6th grade—a widening of 4.1 percentage points—driven by a significant increase in progress among Asian students.

Turning to gender, recall that when looking at test score mobility from 3rd grade to high school in Section 5.2, we found that females were more upwardly mobile than males. This pattern generally holds across students' schooling careers. That is, in any given grade, when there are statistically significant gender gaps in mobility, such gaps almost always favor females. And although, in 5th grade, both Hispanic and White males were more upwardly progressing in ELA tests (Appendix Figure A.5) than their female counterparts, it was only by maximum of 1.2 and 1.8 percentage points, respectively. And by 6th grade the gender gap reversed to favoring females by up to 3.2 (Hispanic) and 3.5 (White) percentage points—a gap that persisted into high school.

We now turn to progression in mobility for two non-test outcomes: absenteeism (Figure 7) and disciplinary incidences (Figure 8) for females (Panels A) and males (Panels B).[35] To understand these figures, recall that because these are negative schooling outcomes, the interpretation of $\hat{O}_s^{25}$ is flipped— *higher* values (e.g., being absent more often) represent more negative outcomes (relative to peers). Thus, in Figure 7, the upward slopes for all students (thicker, black lines in each panel) show a *downward* progression in terms of absenteeism from elementary into middle school. That trendline flattens out for males in middle school and early high school but continues to rise for females into high school until, by 12th grade, differences between genders are statistically indistinguishable. In terms of ethno-racial differences, two findings are of particular note. Perhaps the most striking is the different trajectory of Asian students who showed significantly less downward progress than other ethno-racial subgroups from grade to grade. Equally notable is Other students' significantly higher downward progress until late high school and White students' significantly lower downward progress throughout the high school years.

Turning to disciplinary incidences (Figure 8), note the downward progression (rise in slopes) for all students (thicker, black lines in each panel) in early grades and into middle school. The trend reverses in high school with a steep drop in later grades. Notably, in every grade (except 11th grade), males were more downwardly mobile than females. The steep drop in 10th and 11th grades is an artifact of this specific cohort (3rd graders in 2013) being composed of

[35] Recall that, unlike test scores, there these outcomes might be considered to be at least somewhat subjective. As we note in Section 2, this is especially noteworthy for disciplinary incidents given the evidence of different standards for judging infractions across student subgroups.

students who were in high school during the COVID-19 pandemic. When the pandemic hit, they were halfway through 10th grade (the 2019-20 school year). In 11th grade (the 2020-21 school year), these students were being schooled remotely and very few disciplinary incidences were recorded. In terms of ethno-racial differences, across grades, Asian students showed the least downward progression compared to the average student (except in 11th grade) while Black students, on the other hand, showed the most downward progression, especially Black males during middle school years.

Such progression in disciplinary incidences looks slightly different for cohorts unaffected by the pandemic. For three such cohorts of 3rd graders whom we observe only from grades 9-12 (i.e., 2008-2010), mobility estimates still drop significantly between 10th and 12th grade for the average student (and for each student subgroup except Asian females). But by 12th grade, the average estimated mobility was higher than for the 2013 cohort at 5.7 for females and 10.8 for males. It may be that the upward progression in discipline throughout high school (i.e., downward trending estimates) is because students who incurred disciplinary incidences in middle school eventually dropped out of school.[36]

Lastly—noting that it is an inherently speculative exercise and beyond the scope of this paper to deeply explore—in light of research showing that achievement levels tend to drop as students transition from elementary to middle school (Marks, 2000; Rockoff & Lockwood, 2011), we take a quick look at the extent to which our results differ depending on whether students' 6th grade year was also the year they transitioned to middle school. Table 7 compares estimates academic mobility (i.e., values of $\hat{O}_s^{25}$) in all tested grades after 3rd grade for samples

[36] While we don't observe dropping out directly, we do see that the more disciplinary incidences students experienced in middle school, the more likely they were to attrite from our sample by 11th grade. For example, students with no disciplinary incidences had a 5% chance of leaving, while students with between 2 and 5 incidences had a 9% chance and students with 6 or more incidences had a 12% chance of leaving.

composed of 6th graders who were still in their elementary schools (Columns 1 and 4) and 6th graders who had moved to middle schools (Columns 2 and 5) for math (Panel A) and ELA tests (Panel B).[37] We find statistically significant differences in estimates of academic mobility between students who transitioned to middle school in 6th grade and those who did not. Sixth graders in middle schools made significantly *less* upward progress in 6th grade (by 4 and 2.3 percentage points in math and ELA, respectively) compared to 6th graders in elementary school. This advantage tends to taper off after 6th grade.

6. Discussion and Conclusions

In this paper we document the extent to which students' race/ethnicity, gender and early measures of academic achievement predict their educational progress and success for a variety of high school test and non-test outcomes: math and ELA test scores, college-level math and ELA course-taking, absenteeism, disciplinary incidents, GPA, and on-time graduation. Not surprisingly, the academic performance of students at one point in their schooling careers is highly predictive of later performance. More surprising, however, is the finding that 3rd grade math tests are *more* predictive of nearly all high school test and non-test outcomes than 3rd grade ELA achievement. While many state's retention and intervention policies are predicated on 3rd grade reading achievement (Berne et al., 2025; ExcelinEd, 2024; Fiester, 2010), our findings suggest that states might want to pay more attention to early measures of numeracy.

And while it is not remarkable that early tests are highly predictive of longer-run outcomes, the extent to which early test measures predict high school outcomes *relative* to those

[37] Of our full analytic sample (581,094 students), 49% (285,160) were enrolled in middle school in 6th grade while 29% (168,730) of them were not, i.e., they were in schools with starting grades lower than 6th grade (e.g., PK-4, K-6, PK-12 schools). The remaining 22% of 6th graders (127,204) are either not observed with test scores in 6th grade (74%), attended schools with no corresponding grade span information (25%), or attended schools with starting grades higher than 6th grade (1%).

in 8^{th} grade is still striking. As noted above, 8th grade tests are clearly better predictors of high school outcomes—across outcomes, values of R^2 rise by at least 26% when 8^{th} grade tests are used as predictors instead of 3^{rd} grade tests. But the two models are highly correlated (at least 0.75) and lead to very similar conclusions about students' future academic performance and behavior. Their similarity has important implications: identifying students at risk as early as 3rd grade could enable earlier, more effective interventions. Given the brain's developmental plasticity and heightened receptiveness to learning in early childhood (Nelson & Sheridan, 2011), early support could not only be more impactful but also more cost-effective in the long run.

When we focus on the academic mobility of initially low-achieving 3^{rd} graders, we find that they experienced modest upward *test-based* academic mobility. Mobility patterns for non-test outcomes were more mixed. Our test-based findings align with Austin et al. for high school tests and graduation: initially low-achieving students performed relatively better over time, gaining in the distributions relative to their peers. Our novel findings for other non-test outcomes—GPA, college-level course-taking, absenteeism and, disciplinary actions—show that 3^{rd} grade tests also strongly predict these outcomes. And that the *direction* of mobility (upward or downward) varies across outcomes. For example, in terms of GPA, initially low-achieving students showed upward mobility. Yet, on average, these students were downwardly mobile in college-level course-taking, absenteeism and, at least for males, disciplinary incidents. Though, as we note above with respect to disciplinary action, an important caveat to all these measures is that they may be influenced by the judgments of school personnel who make decisions about these non-test outcomes.

Students' upward academic mobility in test scores from 3rd grade to high school was the result of their general upward *progress* in the math and ELA test score distributions as they advanced from grade to grade. Progression trends for absenteeism and discipline (the only two non-test outcomes we could examine) were more mixed. And, while it merits further attention, we find that academic mobility in math seems particularly sensitive (more so than ELA) to schools' grade configurations: we find that 6th graders who transitioned to middle school showed lower academic mobility in that year than 6th graders who did not move schools. Whether this means there is something harmful about the act of transitioning schools in the 6th grade specifically, or that any school transition in any grade reduces academic mobility in that year, is unclear. Our sample of students who did not transition to middle school in 6th grade includes students who attended a variety of primary schools with grade spans that end before or after 6th grade.

We also document considerable variation in academic mobility by student subgroups. Asian students consistently had the most advantageous patterns of academic mobility compared to other racial/ethnic groups. By contrast, Black students were the least upwardly mobile group for half of the high school outcomes we examined. Notably, when gender gaps existed, they nearly always favored females.

The subgroup differences in high school outcomes are troubling, as they foreshadow postsecondary success in college and the labor market (Backes et al., 2023; Chetty et al., 2014; Goldhaber & Özek, 2019; Gray-Lobe et al., 2023; Jackson et al., 2020; Jackson, 2018), signaling potential inequity beyond high school. Importantly, these disparities are clear well before students reach high school. While there is ongoing debate how much public schools should be expected to change students' educational trajectories—and we do not know what achievement would look like in the absence of any existing interventions—our findings suggest that five years of elementary and middle school experiences and interventions do little to substantially shift the educational trajectories of students who begin school with low achievement.

References

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American sociological review*, *72*(2), 167-180.

Allensworth, E. (2013). The Use of Ninth-Grade Early Warning Indicators to Improve Chicago Schools, *Journal of Education for Students Placed at Risk (JESPAR),* 18(1), 68-83,

Austin, W., Figlio, D., Goldhaber, D., Hanushek, E. A., Kilbride, T., Koedel, C., Sean Lee, J., Lou, J., Özek, U., Parsons, E., Rivkin, S. G., Sass, T. R., & Strunk, K. O. (2023). Academic Mobility in U.S. Public Schools: Evidence from Nearly 3 Million Students. *Journal of Public Economics*, *228*, Article 105016. https://doi.org/10.1016/j.jpubeco.2023.105016

Backes, B., Cowan, C., Goldhaber, D., & Theobald, R. (2023). How to Measure a Teacher: The Influence of Test and Nontest Value-Added on Long-Run Student Outcomes. CALDER Working Paper No. 270-0423-2

Balfanz, R., & Byrnes, V. (2012). The Importance of Being in School: A Report on Absenteeism in the Nation's Public Schools: Promising Approaches. The Education Digest, 78(2), 4-9.

Barrett, N., McEachin, A., Mills, J. & Valant, J. (2021). Disparities and Discrimination in Student Discipline by Race and Family Income. *Journal of Human Resources*, 56(3), 711-748.

Becares, L. & Priest N. (2015). Understanding the Influence of Race/Ethnicity, Gender, and Class on Inequalities in Academic and Non-Academic Outcomes among Eighth-Grade Students: Findings from an Intersectionality Approach. PLoS ONE 10(10).

Berne, J., Jacob, B. A., Weiland, C., & Strunk, K. (2025). The Impacts of Grade Retention Policy with Minimal Retention. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.5250729

Betts, J. R., Zau, A., & Rice, L. (2003). Determinants of Student Achievement: New Evidence from San Diego. Public Policy Institute of California, San Francisco.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2013). Measuring test measurement error: A general approach. *Journal of Educational and Behavioral Statistics*, *38*(6), 629-663.

Card, D., Rothstein, J. (2007). Racial segregation and the black–white test score gap. *Journal of Public Economics.*91(11–12), 2158–2184. https://doi.org/10.1016/j.jpubeco.2007.03.006

Chatterji, Madhabi., 2006. Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence From the Early Childhood Longitudinal Study (ECLS) Kindergarten to first grade sample. Journal of Educational Psychology 98, 489–507.

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics* 129(4), 1553-1623.

Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2009). The Academic Achievement Gap in Grades 3 To 8. *Review of Economics and Statistics,* 91(2), 398-419.

Coffey, M. & Tyner, A. (2023). Excellence Gaps by Race and Socioeconomic Status. Washington D.C.: Thomas B. Fordham Institute (August 2023). https://fordhaminstitute.org/national/research/excellence-gaps-race-and-socioeconomic-status

Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfield, F., & York, R. L. (1966). Equality of educational opportunity. *Washington D. C.: U. S. Government Printing Office.*

Conger, D., & Long, M. C. (2010). Why are men falling behind? Gender gaps in college performance and persistence. *The ANNALS of the American Academy of Political and Social Science,* 627(1), 184–214.

Conwell, J. (2021). Diverging Disparities: Race, Parental Income, and Children's Math Scores 1960 to 2009. *Sociology of Education*. 94(2) 124-142.

Crenshaw, K. (1989). "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics," University of Chicago Legal Forum: Vol. 1989: Iss. 1, Article 8.

Curtin, J., Hurwitch, B., & Olson, T. (2012). Development and Use of Early Warning Systems. SLDS Spotlight. *National Center for Education Statistics.*

DiPrete, T. A., & Buchmann, C. (2013). *The rise of women: The growing gender gap in education and what it means for American schools*. Russell Sage Foundation.

Domina, T., Penner, A., & Penner, E. (2017). Categorical Inequality: Schools As Sorting Machines. *Annual Review of Sociology*, 43, 311-330. 10.1146/annurev-soc-060116-053354

Downey, D. B., von Hippel, P. T., & Broh, B. A. (2004). Are Schools the Great Equalizer? Cognitive Inequality during the Summer Months and the School Year. *American Sociological Review*, *69*(5), 613–635. http://www.jstor.org/stable/3593031

Dumont, H., & Ready, D. D. (2020). Do Schools Reduce or Exacerbate Inequality? How the Associations Between Student Achievement and Achievement Growth Influence Our Understanding of the Role of Schooling. *American Educational Research Journal*, *57*(2), 728–774. https://doi.org/10.3102/0002831219868182

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., & Japel, C. (2007). School readiness and later achievement. *Developmental psychology*, *43*(6), 1428.

Duncan, G. J., Magnuson, K. A. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *Future Child*, 15(1), 35-54. 10.1353/foc.2005.0004

ExcelinEd (2024), "Comprehensive Early Literacy Policy: A State-by-State Analysis of Fundamental Principles," https://excelined.org/wpcontent/uploads/2021/10/ExcelinEd_PolicyToolkit_EarlyLiteracy_StatebyStateAnalysis_2021.pdf

Fiester, L., (2010). Early Warning! Why Reading by the End of Third Grade Matters. KIDS COUNT Special Report. Baltimore, MD: Annie E. Casey Foundation.

Figlio, D. N. (2005). Names, Expectations and the Black-White Test Score Gap. NBER Working Papers 11195. *National Bureau of Economic Research (NBER).*

Finkelstein, N.D., & Fong, A.B. (2008). *Course-taking patterns and preparation for postsecondary education in California's public university systems among minority youth.* (Issues & Answers Report, REL 2008–No. 035). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. Retrieved from http://ies.ed.gov/ncee/edlabs

Finn, J. D., Gerber, S. B., & Boyd-Zaharias, J. (2005). Small classes in the early grades, academic achievement, and graduating from high school. *Journal of Educational Psychology*, *97*(2), 214.

Forte, D. (2021). *Why Parents, School Leaders, & Advocates Shouldn't Underestimate the Power of Statewide Assessments.* The Education Trust. https://edtrust.org/the-equity 27 line/why-parents-school-leaders-and-advocates-shouldnt-underestimate-the-power-ofstatewide-assessments/

Fortin, N. M., Oreopoulos, P., & Phipps, S. (2015). Leaving boys behind: Gender disparities in high academic achievement. *The Journal of Human Resources*, *50*(3), 549–579.

Fryer, R. G., & Levitt, S. D. (2006). The Black-White Test Score Gap Through Third Grade. *American Law and Economics Review*, *8*(2), 249–281. http://www.jstor.org/stable/42705499

Gamoran, A. (1987). The stratification of high school learning opportunities. *Sociology of Education, 60*(3), 135–155. https://doi.org/10.2307/2112271

Gamoran, A. (2001). American Schooling and Educational Inequality: A Forecast for the 21st Century. *Sociology of Education*, *74*, 135–153. https://doi.org/10.2307/2673258

Goldhaber, D., & Özek, U. (2019). How Much Should We Rely on Student Test Achievement as a Measure of Success? *Educational Researcher*, *48*(7), 479 483.

Goldhaber, D., Koedel, C., Özek, U., & Parsons, E. (2022). Using Longitudinal Student Mobility to Identify At-Risk Students. *AERA Open*, 8. https://doi.org/10.1177/23328584211071090

Goldhaber, D., Theobald, R., & Fumia, D. (2022). The role of teachers and schools in explaining stem outcome gaps. *Social Science Research, 105*, 102709. https://doi.org/10.1016/j.ssresearch.2022.102709

Goldhaber, D., Wolff, M., & Daly, T. (2021). Assessing the Accuracy of Elementary School Test Scores as Predictors of Students' High School Outcomes. CALDER Working Paper 235-0520

Gordon, M & Cui, M. (2018). The Intersection of Race and Community Poverty and Its Effects on Adolescents' Academic Achievement. *Youth & Society*, 50(7) 947-965.

Gottfried, M. A. (2014). Can Neighbor Attributes Predict School Absences? (vol 49, pg 216, 2014). *Urban Education (Beverly Hills, Calif.)*, *49*(3), 216–250.

Gray-Lobe, G., Pathak, P. A., & Walters, C. R. (2023). The long-term effects of universal preschool in Boston. *The Quarterly Journal of Economics*, *138*(1), 363-411.

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The Achievement Gap and the Discipline Gap: Two Sides of the Same Coin? *Educational Researcher*, *39*(1), 59–68.

Hanushek, E. A., & Rivkin, S. G. (2009). Harming the best: How schools affect the black-white achievement gap. *Journal of policy analysis and management*, *28*(3), 366-393.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*(5782), 1900-1902.

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of Political Economy*, *126*(5), 2072-2107.

Jackson, C. K., Porter, S. C., Easton, J. Q., Blanchard, A., & Kiguel, S. (2020). School effects on socioemotional development, school-based arrests, and educational attainment. *American Economic Review: Insights*, 2(4), 491–508. https://doi.org/10.1257/aeri.20200029.

Jacob, B., Dynarski, S., Frank, K., & Schneider, B. (2017). Are expectations alone enough? Estimating the effect of a mandatory college-prep curriculum in Michigan. *Educational Evaluation and Policy Analysis*, *39*(2), 333-360.

Kao, G., & Thompson, J. S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual review of sociology*, *29*(1), 417-442.

Kelly, S. (2009). The black-white gap in mathematics course taking. *Sociology of Education*, *82*(1), 47-69.

Koretz D. (2017). *The testing charade: Pretending to make schools better*. Chicago, IL: University of Chicago Press.

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational researcher*, 31(1), 3-12.

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

Liu, J., Hayes, M. S., & Gershenson, S. (2024). JUE Insight: From referrals to suspensions: New evidence on racial disparities in exclusionary discipline. *Journal of Urban Economics*, *141*, Article 103453. https://doi.org/10.1016/j.jue.2022.103453

Lockwood, J. R., & McCaffrey, D. F. (2014). Correcting for test score measurement error in ANCOVA models for estimating treatment effects. *Journal of Educational and Behavioral Statistics*, *39*(1), 22-52.

Losen, D., Hodson, C., Ee, J., & Martinez, T. (2015). Disturbing Inequities: Exploring the Relationship Between Racial Disparities in Special Education Identification and Discipline. *The Journal of Applied Research on Children*, *5*(2).

Martínez, L., & Serna, N. (2018). Disparities at the entrance door: gender gaps in elementary school. *Educação e Pesquisa*, *44*.

Malkus, N. (2024). Long COVID for public schools: Chronic absenteeism before and after the pandemic. https://www.aei.org/wp-content/uploads/2024/01/Long-COVID-for-Public-Schools.pdf

Marks, H. M. (2000). Student Engagement in Instructional Activity: Patterns in the Elementary, Middle, and High School Years. *American Educational Research Journal*, *37*(1), 153–184.

McDonough, I. K. (2015). Dynamics of the black–white gap in academic. *Economics of Education Review*, *47*, 17-33.

McGraw, R. Lubienski, S. & Strutchens, M. (2006). A Closer Look at Gender in NAEP Mathematics Achievement and Affect Data: Intersections with Achievement, Race/Ethnicity, and Socioeconomic Status. Journal for Research in Mathematics Education, 37(2) 129-150.

Nelson, Charles A., III & Sheridan, Margaret A. (2011). Lessons from Neuroscience Research for Understanding Causal Links between Family and Neighborhood Characteristics and Educational Outcomes.

Duncan, Greg J. & Murnane, Richard J. (Eds.) (pp. 27-46). New York, NY: Russell Sage Foundation.

Northrop, L. (2017). Breaking the cycle: cumulative disadvantage in literacy. *Reading Research Quarterly*, *52*(4), 391-396.

Quinn, D. M., & Cooc, N. (2015). Science Achievement Gaps by Gender and Race/Ethnicity in Elementary and Middle School: Trends and Predictors. *Educational Researcher*, *44*(6), 336–346.

Reardon, S. F. (2011). The widening socioeconomic status achievement gap: New evidence and possible explanations. *Whither opportunity*, 91-116.

Reardon, S. F. (2019). Educational opportunity in early and middle childhood: Using full population administrative data to study variation by place and age. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *5*(2), 40-68.

Reardon, S. F., & Portilla, X. A. (2016). Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry. *Aera Open*, *2*(3), 2332858416657343.

Reardon, S. F., Fahle, E. M., Kalogrides, D., Podolsky, A., & Zárate, R. C. (2019). Gender achievement gaps in US school districts. *American Educational Research Journal*, *56*(6), 2474-2508.

Reardon, S.F. (2015). School Segregation and Racial Academic Achievement Gaps. RSF : Russell Sage Foundation journal of the social sciences, 2016-09, Vol.2 (5), p.34-57; New York: Russell Sage Foundation

Reeves, R. (2022). *Of Boys And Men. Why The Modern Male Is Struggling, Why It Matters, And What To Do About It*. Brookings Institution Press. Washington, D.C.

Robinson, J. P., & Lubienski, S. T. (2011). The Development of Gender Achievement Gaps in Mathematics and Reading During Elementary and Middle School: Examining Direct Cognitive Assessments and Teacher Ratings. *American Educational Research Journal*, *48*(2), 268–302.

Rockoff, J. & Lockwood, B. (2010). Stuck in the Middle: Impacts of Grade Configuration in Public Schools. *Journal of Public Economics*, 94:1051-1061.

Shi, Y., & Zhu, M. (2022). Equal time for equal crime? Racial bias in school discipline. *Economics of Education Review*, *88.*

Sorensen, L. C. (2019). “Big data” in educational administration: An application for predicting school dropout risk. *Educational Administration Quarterly*, 55(3), 404-446.\

Skedgell, K., & Kearney, C. A. (2018). Predictors of school absenteeism severity at multiple levels: A classification and regression tree analysis. *Children and Youth Services Review*, *86*, 236–245.

Skiba, R. J., Horner, R. H., Chung, C.-G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race Is Not Neutral: A National Investigation of African American and Latino Disproportionality in School Discipline. *School Psychology Review*, *40*(1), 85–107.

Strauss, V. (2015, March 1). The important things standardized tests don't measure. *The Washington Post*. https://www.washingtonpost.com/news/answersheet/wp/2015/03/01/the-important-things-standardized-tests-dont-measure/

Tach, L. M., & Farkas, G. (2006). Learning-related behaviors, cognitive skills, and ability grouping when schooling begins. *Social Science Research*, *35*(4), 1048-1079.

Todd, P. E., & Wolpin, K. I. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital*,1(1),91–136.

Von Hippel, P. T., & Hamrock, C. (2019). Do test score gaps grow before, during, or between the school years? Measurement artifacts and what we can know in spite of them. *Sociological Science*, 6, 43–80. https://doi.org/10.15195/v6.a3

Von Hippel, P. T., Workman, J., &Downey, D. B. (2018). Inequality in reading and math skills forms mainly before kindergarten: A replication, and partial correction, of ''Are Schools the Great Equalizer?'' *Sociology of Education*, 91(4), 32335.

Wang, M. W., & Stanley, J. C. (1970). Differential weighting: A review of methods and empirical studies. *Review of Educational Research*, *40*(5), 663-705.

Tables and Figures

Table 1. High School Outcome Definitions and Construction

Outcome	Element	Definition	3rd Grade Cohorts	Outcome years	Details on Variable Construction
Test Scores	Math Score	Math test scores on SBA standardized (to mean 0 and 1 standard deviation) by grade and year	2007-2012	2015-2019	Test scores are standardized by year and grade. We use students' earliest available high school test score from either 10th or 11th grade. We exclude the 2006 cohort because those students were not required to take the SBA in high school. And we exclude the 2013 cohort because those students were not tested due to the COVID-19 pandemic.
	ELA Score	ELA test scores standardized (to mean 0 and 1 standard deviation) by grade and year	2007-2012	2015-2019	
College-level Course Taking	College-level math	Number of college-level math courses taken in grades 9-12 weighted by course term length	2006-2013	2012-2022	College-level courses include: College-level placement, College in HS, International baccalaureate, Running start, and Cambridge scholars' program. Total counts weighted by course term length.
	College-level ELA	Number of college-level ELA courses taken in grades 9-12 weighted by course term length	2006-2013	2012-2022	
Absences	Total absences	Total number of excused and unexcused absences in grades 9-12 calculated by weighted sum of full- and half-day absences	2006-2013	2012-2022	Until the 2012-13 school year absences were not reported separately by full-/part-day or excused/unexcused. Since students in the 2006 cohort were in 9th grade in 2012, we use "full day" absences for that cohort and grade.
Disciplinary Incidences	All disciplinary incidences	Total number of disciplinary incidences in grades 9-12	2006-2013	2014-2022	Students in the 2006 cohort contribute two years of data (11th and 12th grade) and students in the 2007 cohort contribute three years of data (10th – 12th grade) to this measure.
GPA	12th grade GPA	Grade point average measured on a 4.0 scale	2006-2013	2015-2022	Students' cumulative high school GPA in 12th grade.
Graduation	Graduation	Whether or not there is a record of the student graduating high school	2006-2013	2015-2022	On-time graduation is defined as graduating 4 years after entering high school.

Table 2. Summary Statistics of 3rd Grade Classifications and Test Scores by Race/Ethnicity and Gender

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Overall	Race					Gender	
		Asian	Black	Hispanic	Others	White	Male	Female
Panel A: Student Classifications (3rd Grade)								
Free/Reduced Lunch	0.46	0.34	0.71	0.80	0.54	0.34	0.46	0.46
	(0.50)	(0.47)	(0.45)	(0.40)	(0.50)	(0.47)	(0.50)	(0.50)
Special Education	0.11	0.06	0.12	0.10	0.12	0.12	0.14	0.07
	(0.31)	(0.24)	(0.33)	(0.30)	(0.33)	(0.32)	(0.35)	(0.26)
Learning Disability	0.04	0.02	0.05	0.05	0.04	0.04	0.05	0.03
	(0.19)	(0.12)	(0.22)	(0.21)	(0.20)	(0.19)	(0.21)	(0.17)
Limited English	0.10	0.16	0.08	0.37	0.03	0.02	0.11	0.09
	(0.30)	(0.37)	(0.27)	(0.48)	(0.18)	(0.13)	(0.31)	(0.29)
Highly Capable	0.03	0.08	0.01	0.01	0.03	0.04	0.03	0.03
	(0.18)	(0.27)	(0.12)	(0.10)	(0.17)	(0.19)	(0.18)	(0.18)
Panel B: Student Test Scores (3rd Grade)								
Math	0.00	0.39	-0.47	-0.44	-0.13	0.15	-0.01	0.02
	(1.00)	(1.05)	(0.96)	(0.93)	(1.01)	(0.95)	(1.00)	(0.99)
ELA	0.00	0.23	-0.40	-0.45	-0.11	0.16	-0.11	0.12
	(1.00)	(0.96)	(0.96)	(0.96)	(1.01)	(0.96)	(1.01)	(0.98)
No. Obs.	581,094	43,171	27,716	110,418	46,995	352,794	297,320	283,774
Percent of Total Obs.	100	7.43	4.77	19.00	8.09	60.71	51.17	48.83

Note. Panel A reports the proportion of students with various classifications. Panel B reports average test scores. Standard deviations shown in parentheses. Results from one-sample t-tests (not shown) comparing differences in subgroup means (Columns 2-8) to the overall mean of each variable (Column 1) are significant at $p<0.001$ for all subgroups except for differences by gender for highly-capable status where $p<0.01$ and free-reduced lunch eligibility which are not significant at conventional levels.

Table 3. Differences in 3rd Grade Means for 3rd Graders for Whom We Do and Do Not Observe High School Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(9)
High School Outcomes:	High School Math Test	High School ELA Test	College-level Math Courses	College-level ELA Courses	Absences	Disciplinary Incidences	GPA	Graduated On-time
Panel A: Student Demographics (3rd Grade)								
Asian	0.000	0.003***	-0.006***	-0.006***	-0.006***	-0.004***	0.019***	0.017***
Black	-0.014***	-0.014***	-0.032***	-0.032***	-0.032***	-0.032***	-0.023***	-0.026***
Hispanic	0.030***	0.018***	0.046***	0.046***	0.046***	0.045***	0.006***	0.027***
Other	0.001	-0.001	0.003*	0.003*	0.003*	0.003**	-0.010***	-0.011***
White	-0.017***	-0.006***	-0.011***	-0.011***	-0.011***	-0.012***	0.008***	-0.007***
Female	0.001	0.004**	-0.006**	-0.006**	-0.006**	-0.004	0.025***	0.015***
Panel B: Student Classifications (3rd Grade)								
Free/Reduced-Lunch	-0.005***	-0.027***	0.004	0.004	0.004	-0.003	-0.101***	-0.081***
Special Education	-0.005***	-0.010***	-0.001	-0.001	-0.001	-0.002	-0.015***	-0.013***
Learning Disability	-0.003***	-0.005***	0.002**	0.002**	0.002**	0.001	-0.008***	-0.007***
Limited English	0.004***	-0.004***	0.020***	0.020***	0.020***	0.018***	0.001	0.013***
Highly Capable	-0.002***	-0.001	-0.001	-0.001	-0.001	0.000	0.008***	0.006***
Panel C: Test Scores (3rd Grade)								
Math (Mean=1, SD=0)	0.031***	0.055***	0.013**	0.013**	0.013**	0.029***	0.189***	0.154***
ELA (Mean=1, SD=0)	0.024***	0.047***	-0.007	-0.007	-0.007	0.007	0.172***	0.130***
No. of Students Observed	269,494	320,509	521,391	521,391	521,391	517,319	467,162	464,855
No. of Students Not Observed	311,600	260,585	59,703	59,703	59,703	63,775	113,932	116,239

Note. This table shows *differences* in means of student characteristics between different subsamples of 3rd graders we observe with a given high school outcome and subsamples of 3rd graders we do not observe with a given high school outcome. Positive values indicate higher proportions of students in our "observable" subsample compared to our "unobservable" subsample. And vice versa for negative values. We also report the statistical significance of t-tests between the two subsamples: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4. Summary Statistics of High School Outcomes by Race/Ethnicity and Gender

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Overall	Race					Gender	
		Asian	Black	Hispanic	Other	White	Male	Female
Panel A: High School Test Scores								
Math	0.06	0.68	-0.43	-0.37	-0.07	0.19	0.06	0.07
	(0.97)	(0.98)	(0.88)	(0.85)	(0.97)	(0.94)	(1.02)	(0.91)
ELA	0.07	0.54	-0.36	-0.34	-0.06	0.20	-0.04	0.20
	(0.94)	(0.85)	(0.95)	(0.90)	(0.98)	(0.90)	(0.98)	(0.89)
Panel B: High School Non-test Outcomes								
College-level Math Courses	0.36	0.98	0.25	0.20	0.32	0.35	0.36	0.37
	(0.87)	(1.40)	(0.73)	(0.64)	(0.82)	(0.83)	(0.86)	(0.87)
College-level ELA Courses	0.46	0.91	0.51	0.37	0.44	0.43	0.38	0.55
	(0.94)	(1.26)	(0.99)	(0.86)	(0.95)	(0.90)	(0.86)	(1.01)
Absences	74.5	48.94	99.93	93.61	86.11	68.04	73.44	75.61
	(67.84)	(55.52)	(86.09)	(75.40)	(76.39)	(61.35)	(68.20)	(67.44)
Disciplinary Incidences	0.53	0.16	1.14	0.72	0.70	0.45	0.74	0.31
	(1.97)	(0.90)	(2.98)	(2.50)	(2.14)	(1.73)	(2.37)	(1.41)
12th Grade GPA	2.79	3.19	2.44	2.49	2.64	2.88	2.64	2.94
	(0.90)	(0.86)	(0.87)	(0.88)	(0.92)	(0.88)	(0.91)	(0.87)
Graduated On-time	0.86	0.94	0.78	0.81	0.81	0.87	0.83	0.88
	(0.35)	(0.24)	(0.41)	(0.39)	(0.39)	(0.34)	(0.37)	(0.33)

Note. Panel A reports average standardized high school test scores. Panel B reports averages for various high school outcomes. Standard deviations shown in parentheses. Results from one-sample t-tests (not shown) comparing differences in subgroup means (Columns 2-8) to the overall mean of each variable (Column 1) are significant at $p<0.001$ for all subgroups except for differences by gender for college-level math courses where $p < 0.01$.

Table 5. Regressions Predicting High School Test Outcomes with 3rd Grade Predictors

	Panel A: High School Math Tests				Panel B: High School ELA Tests			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Student Test Scores								
3rd Grade Math	0.544***		0.433***				0.263***	
	(0.002)		(0.002)				(0.002)	
3rd Grade ELA			0.194***		0.504***		0.348***	
			(0.002)		(0.001)		(0.002)	
8th Grade Math		0.766***		0.644***				0.284***
		(0.001)		(0.002)				(0.002)
8th Grade ELA				0.188***		0.690***		0.494***
				(0.002)		(0.001)		(0.002)
Student Demographics (3rd Grade)								
Asian	0.362***	0.138***	0.373***	0.133***	0.296***	0.121***	0.247***	0.051***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)
Black	-0.161***	-0.085***	-0.141***	-0.073***	-0.168***	-0.117***	-0.113***	-0.067***
	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Hispanic	-0.091***	-0.063***	-0.077***	-0.054***	-0.081***	-0.060***	-0.056***	-0.039***
	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)
Other Race	-0.064***	-0.033***	-0.054***	-0.026***	-0.080***	-0.062***	-0.060***	-0.045***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)
Female	0.006*	-0.036***	-0.030***	-0.081***	0.125***	0.032***	0.159***	0.072***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Student Classifications (3rd Grade)								
Free/Reduced-Lunch	-0.275***	-0.144***	-0.242***	-0.120***	-0.263***	-0.173***	-0.226***	-0.129***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Special Education	-0.080***	-0.025***	-0.034***	-0.001	-0.115***	-0.087***	-0.071***	-0.054***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)
Learning Disability	-0.186***	-0.111***	-0.107***	-0.075***	-0.132***	-0.151***	-0.099***	-0.107***
	(0.009)	(0.007)	(0.009)	(0.007)	(0.008)	(0.008)	(0.008)	(0.007)
Limited English	0.013*	-0.025***	0.083***	0.013**	-0.033***	-0.081***	0.011*	-0.068***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)
Highly Capable	0.442***	0.233***	0.391***	0.203***	0.380***	0.267***	0.252***	0.137***
	(0.008)	(0.007)	(0.008)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)
N	269,494	239,558	269,494	238,995	320,509	282,561	320,509	281,342
R^2	0.471	0.674	0.490	0.687	0.441	0.583	0.479	0.620

Note. Results are from OLS regressions and can be interpreted in standard deviation units.
Standard errors in parentheses * p < 0.05, ** p < 0.01, *** p < 0.001.

Table 6. Average Marginal Effects from Regressions Predicting High School Non-Test Outcomes with 3rd Grade Predictors

	(1)	(2)	(3)	(4)	(5)	(6)
	No. Adv. Math Courses	No. Adv. ELA Courses	No. of Absences	No. of Disciplinary Incidents	12th Grade GPA	Graduated On-time
Student Test Scores (3rd Grade)						
Math	0.209***	0.113***	-7.243***	-0.165***	0.193***	0.042***
	(0.002)	(0.002)	(0.123)	(0.004)	(0.002)	(0.001)
ELA	0.100***	0.139***	-4.383***	-0.142***	0.127***	0.027***
	(0.002)	(0.002)	(0.125)	(0.004)	(0.002)	(0.001)
Student Demographics (3rd Grade)						
Asian	0.354***	0.322***	-23.439***	-0.495***	0.256***	0.088***
	(0.004)	(0.005)	(0.342)	(0.012)	(0.004)	(0.003)
Black	0.156***	0.324***	12.035***	0.246***	-0.117***	0.004
	(0.007)	(0.007)	(0.427)	(0.011)	(0.006)	(0.002)
Hispanic	0.083***	0.152***	7.310***	0.050***	-0.070***	0.021***
	(0.004)	(0.004)	(0.262)	(0.007)	(0.004)	(0.001)
Other Race	0.063***	0.104***	7.910***	0.115***	-0.091***	-0.017***
	(0.005)	(0.005)	(0.322)	(0.008)	(0.004)	(0.002)
Female	-0.007**	0.154***	3.434***	-0.473***	0.256***	0.040***
	(0.002)	(0.003)	(0.173)	(0.006)	(0.004)	(0.001)
Student Classifications (3rd Grade)						
Free/Reduced-Lunch	-0.239***	-0.181***	24.711***	0.397***	-0.319***	-0.098***
	(0.003)	(0.003)	(0.200)	(0.006)	(0.003)	(0.001)
Special Education	-0.032***	-0.071***	-5.760***	-0.030***	0.087***	-0.011***
	(0.005)	(0.006)	(0.335)	(0.009)	(0.005)	(0.002)
Learning Disability	-0.160***	-0.152***	6.262***	-0.092***	-0.009	0.030***
	(0.012)	(0.012)	(0.541)	(0.014)	(0.008)	(0.003)
Limited English	0.099***	0.104***	-6.064***	-0.223***	0.109***	0.038***
	(0.005)	(0.005)	(0.332)	(0.009)	(0.005)	(0.002)
Highly Capable	0.191***	0.165***	-7.027***	-0.279***	0.199***	0.031***
	(0.005)	(0.006)	(0.493)	(0.017)	(0.007)	(0.004)
N	521,391	521,391	521,391	517,319	467,162	464,855
(Pseudo) R^2	0.103	0.053	0.011	0.047	0.224	0.093

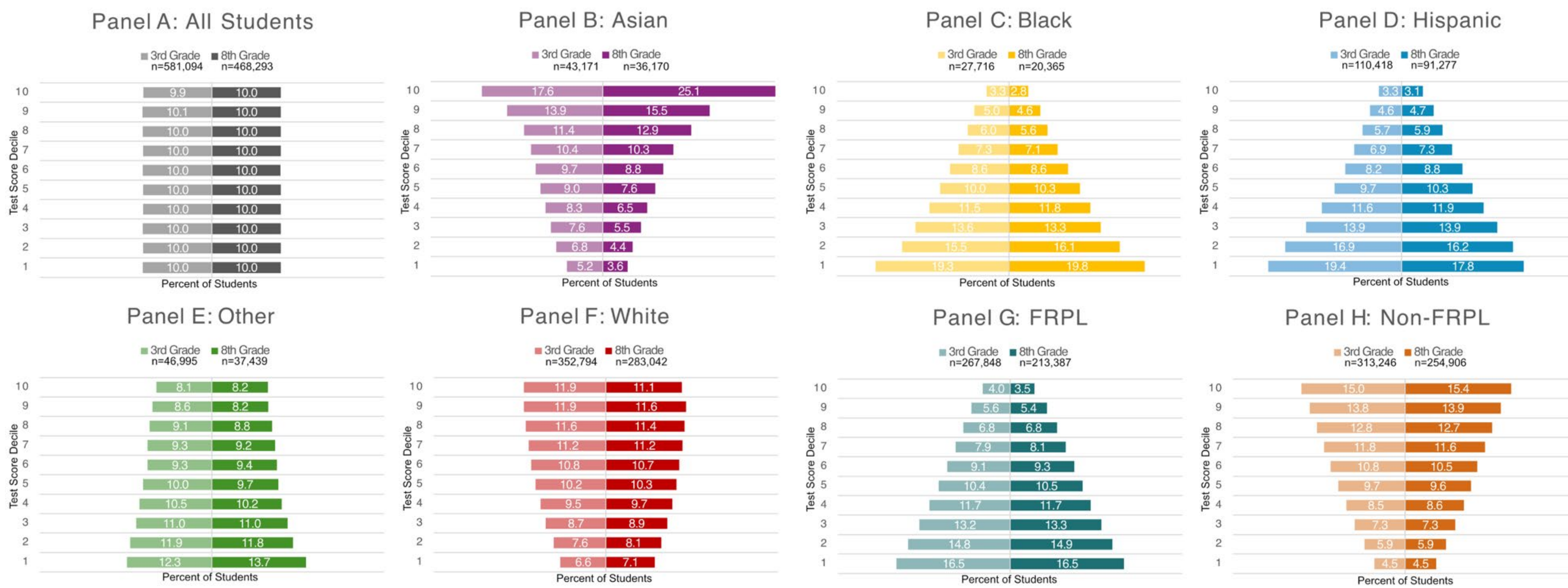
Note. All results are reported as average marginal effects. Model specifications vary depending on the outcome. We estimate non-linear models with negative binomial (Columns 1-4) or logistic (Column 6) regressions. Standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 7. Academic Mobility Estimates for Students Who Did and Did Not Transition to Middle School in 6th Grade

	Panel A: Math			Panel B: ELA		
	(1)	(2)	(3)	(4)	(5)	(6)
Grade	6th Grade in Middle School	6th Grade in Elementary School	Difference	6th Grade in Middle School	6th Grade in Elementary School	Difference
4	27.73	27.71	0.02	28.94	29.24	-0.29***
5	28.47	27.89	0.59***	29.65	29.30	0.36***
6	27.01	31.10	-4.09***	29.18	31.44	-2.26***
7	28.57	29.84	-1.27***	30.28	30.53	-0.25**
8	30.17	30.52	-0.34***	31.37	31.27	0.10
10	30.86	32.42	-1.56***	31.69	32.16	-0.47***

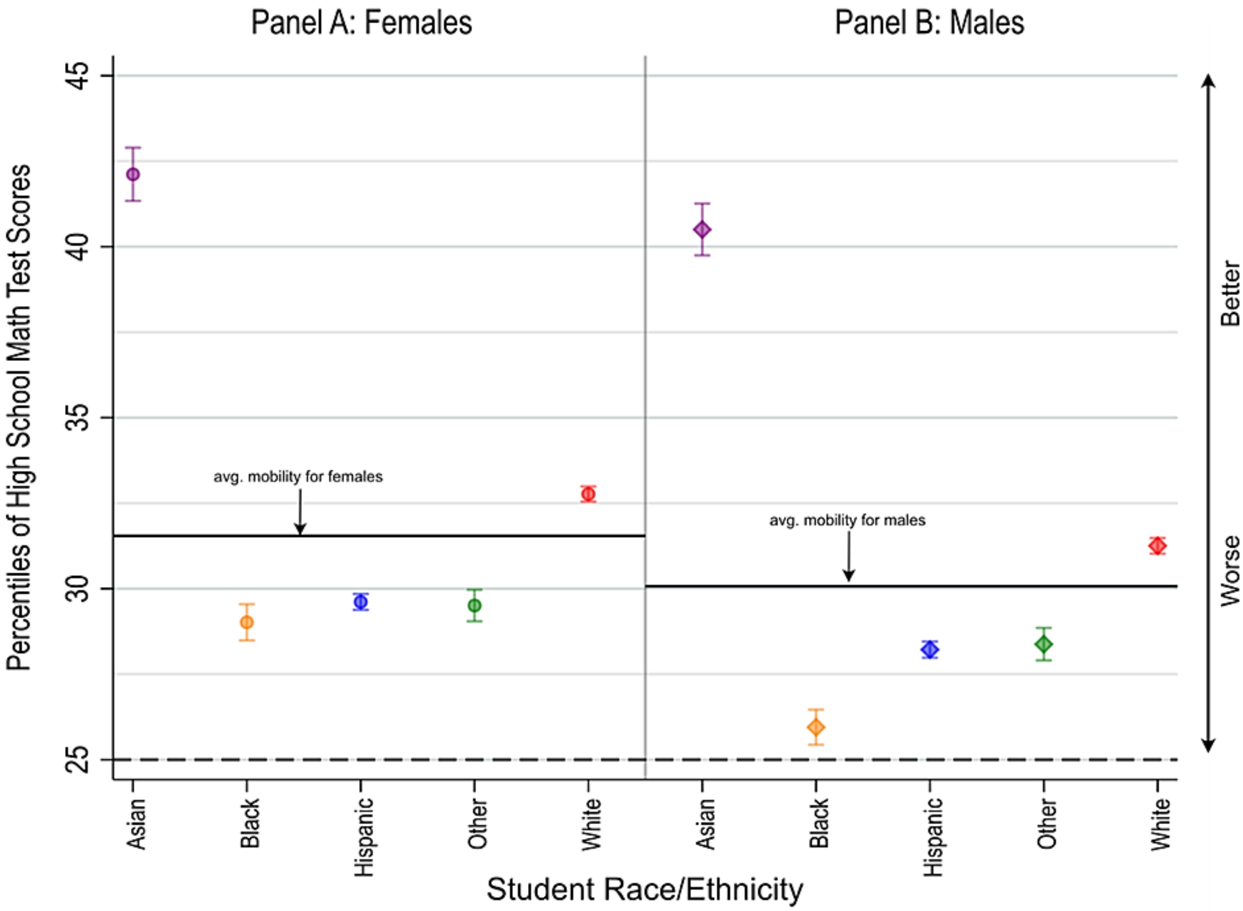
Note. This table shows the average academic mobility of students in each (tested) grade by whether their 6th grade year took place in a middle (Columns 1 & 4) or an elementary school (Columns 2 & 5) and the difference in mobility for those two groups (Columns 3 & 6). Panel A reports mobility for math tests and Panel B for ELA tests. The sample is restricted to students for whom we could observe both 6th grade test scores and school-level information about grade spans: N=453,890 (84% of our full analytic sample). Statistical significance shown as: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure 1. Percentage of Students Across Deciles of Averaged Test Scores in 3rd and 8th Grade, by Student Subgroup



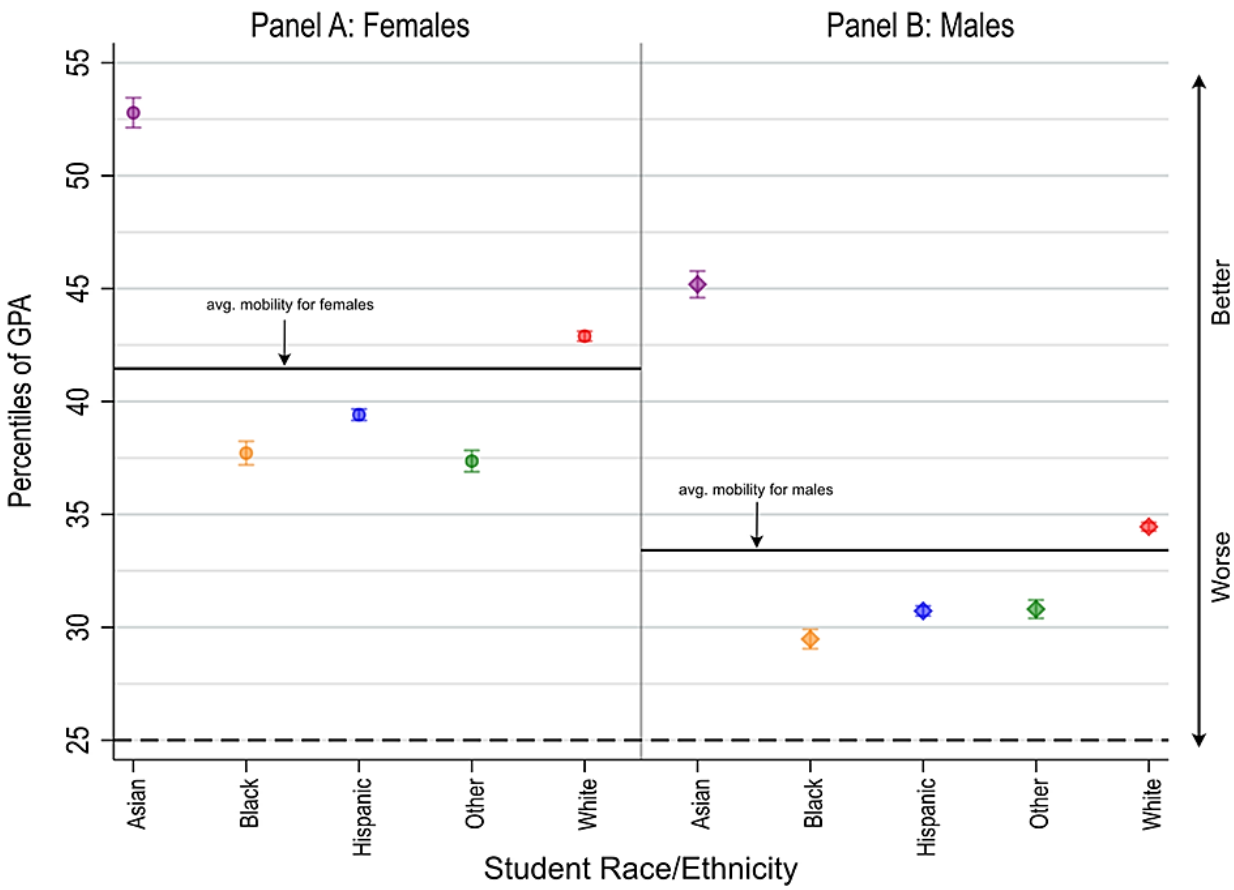
Note. This figure shows "performance pyramids" for each racial/ethnic group and by free/reduced-price lunch (FRPL) eligibility, i.e., the percentage of students who scored in each averaged test score decile (plotted on the y-axis) in 3rd grade (lighter, left-hand bars in each panel) and in 8th grade (darker, right-hand bars in each panel).

Figure 2. Academic Mobility in High School Math Test Scores for Low-Performing 3rd Graders, by Race/Ethnicity and Gender



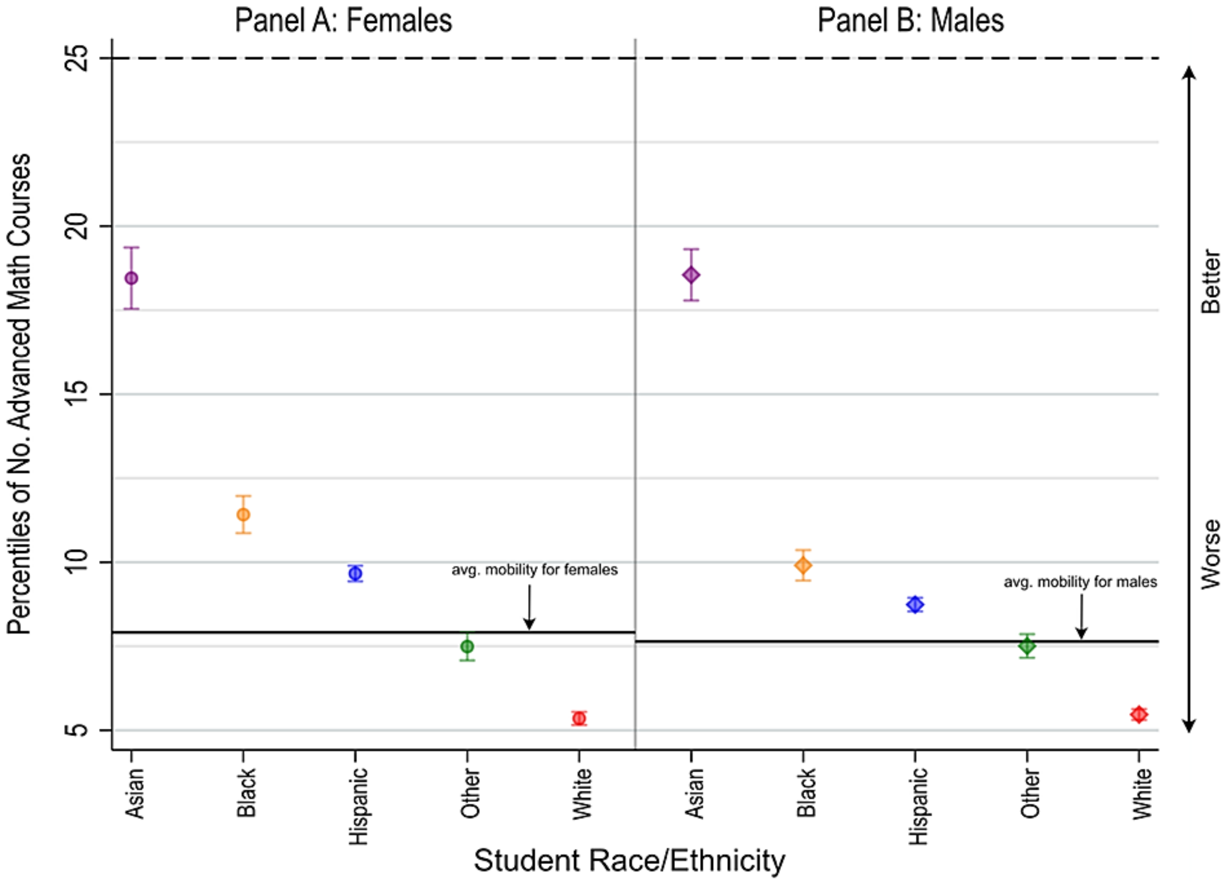
Note. This figure shows estimates of academic mobility in high school math tests for low-achieving 3rd graders (i.e., 3rd graders who scored in the 25th percentile on math tests). Results are shown for each student subgroup. Panel A shows results for females (circle symbols) and Panel B shows results for males (diamond symbols). Racial/ethnic groups are color-coded: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). In each panel, solid horizontal black lines show the average academic mobility for that gender. The dashed horizontal black line is plotted at 25th percentile of the outcome variable. Estimates above the dashed line reflect upward mobility, estimates below it reflect downward mobility and estimates that cross it reflect no statistically significant academic mobility.

Figure 3. Academic Mobility in 12th Grade GPA for Low-Performing 3rd Graders, by Race/Ethnicity and Gender



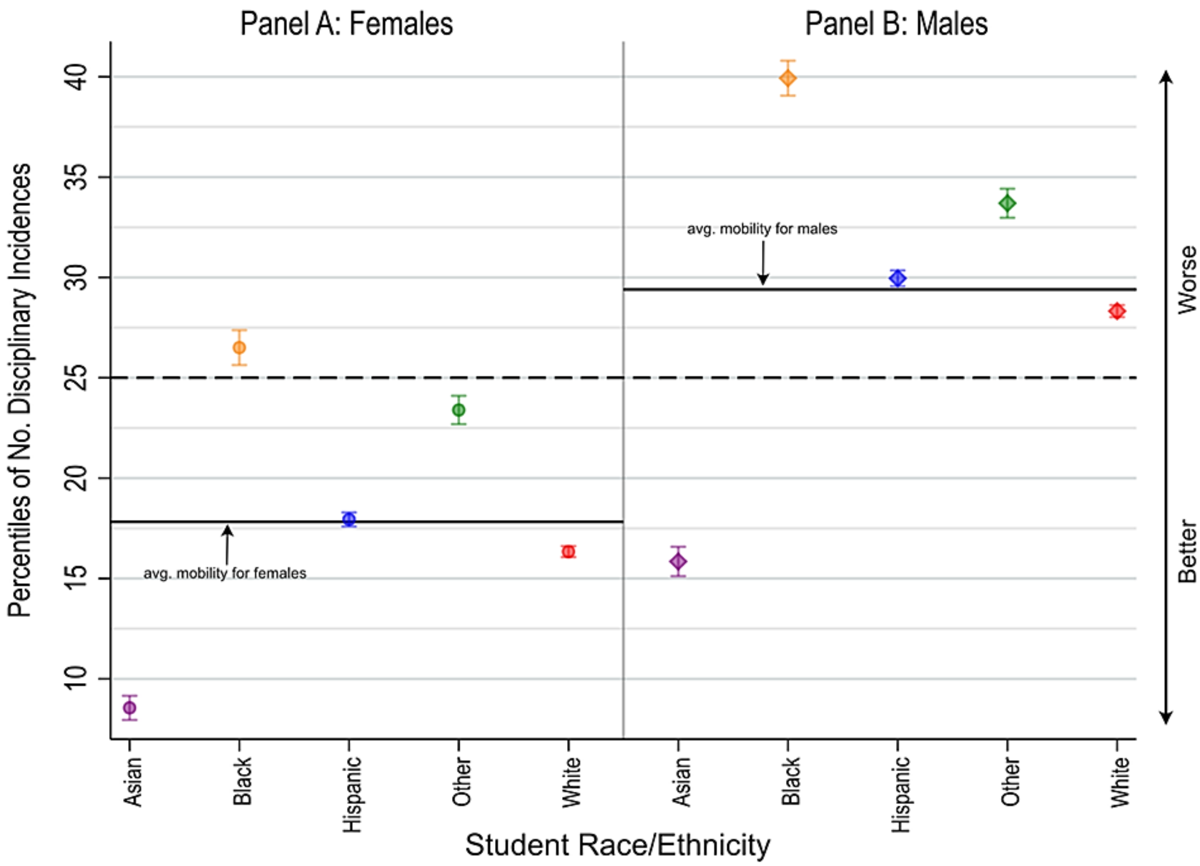
Note. This figure shows estimates of academic mobility in 12th grade GPA for low-achieving 3rd graders (i.e., 3rd graders whose averaged math and ELA scores were at the 25th percentile). Results are shown for each student subgroup. Panel A shows results for females (circles) and Panel B shows results for males (diamonds). Racial/ethnic groups are color-coded: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). In each panel, solid black horizontal black lines show the average academic mobility for that gender. The dashed horizontal black line is plotted at 25th percentile of the outcome variable. Estimates above the dashed line reflect upward mobility, estimates below it reflect downward mobility and estimates that cross it reflect no statistically significant academic mobility.

Figure 4. Academic Mobility in Math Course-Taking for Low-Performing 3rd Graders by Race/Ethnicity and Gender



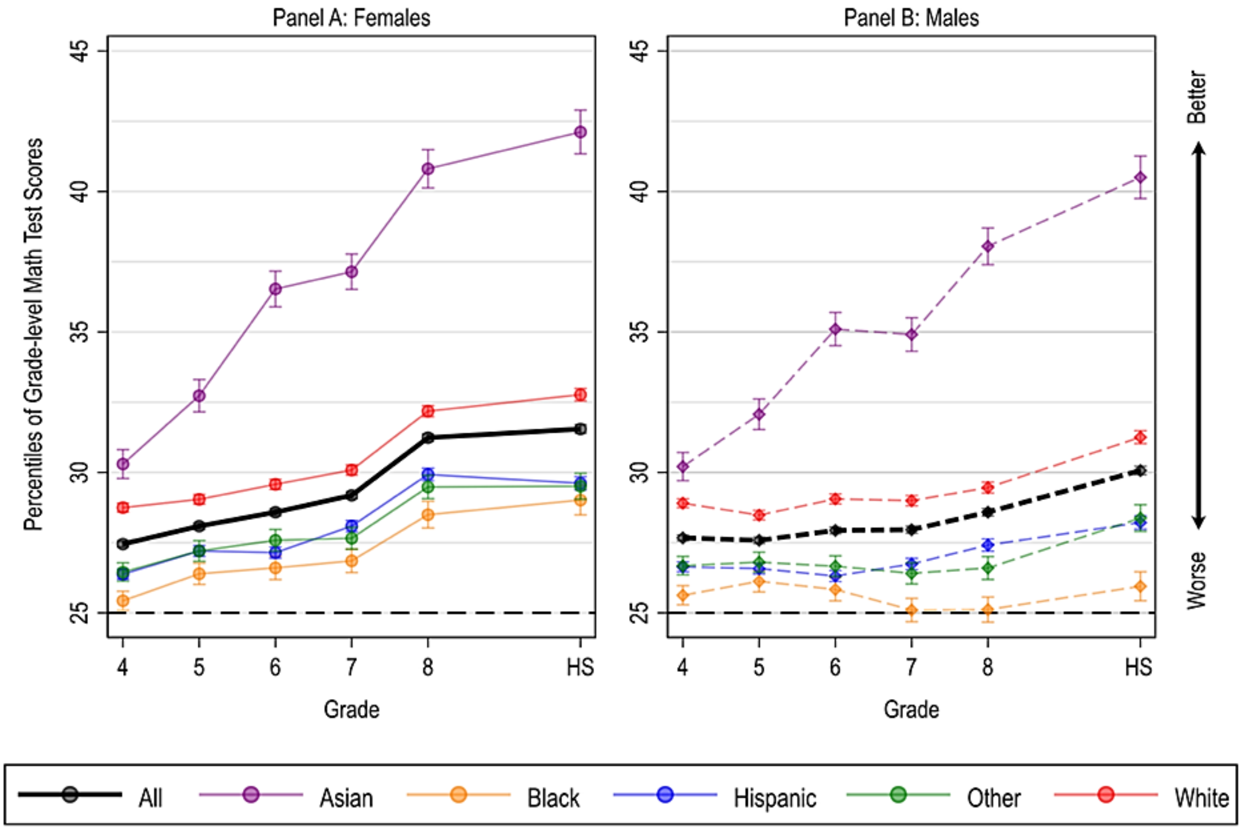
Notes. This figure shows estimates of academic mobility in high school math course-taking for low-achieving 3rd graders (i.e., 3rd graders whose averaged math and ELA scores were at the 25th percentile). Results are shown for each student subgroup. Panel A shows results for females (circles) and Panel B shows results for males (diamonds). Racial/ethnic groups are color-coded: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). In each panel, solid horizontal black lines show the average academic mobility for that gender. The dashed horizontal black line is plotted at 25th percentile of the outcome variable. Estimates above the dashed line reflect upward mobility, estimates below it reflect downward mobility and estimates that cross it reflect no statistically significant academic mobility.

Figure 5. Academic Mobility in Disciplinary Incidences for Low-Performing 3rd Graders by Race/Ethnicity and Gender



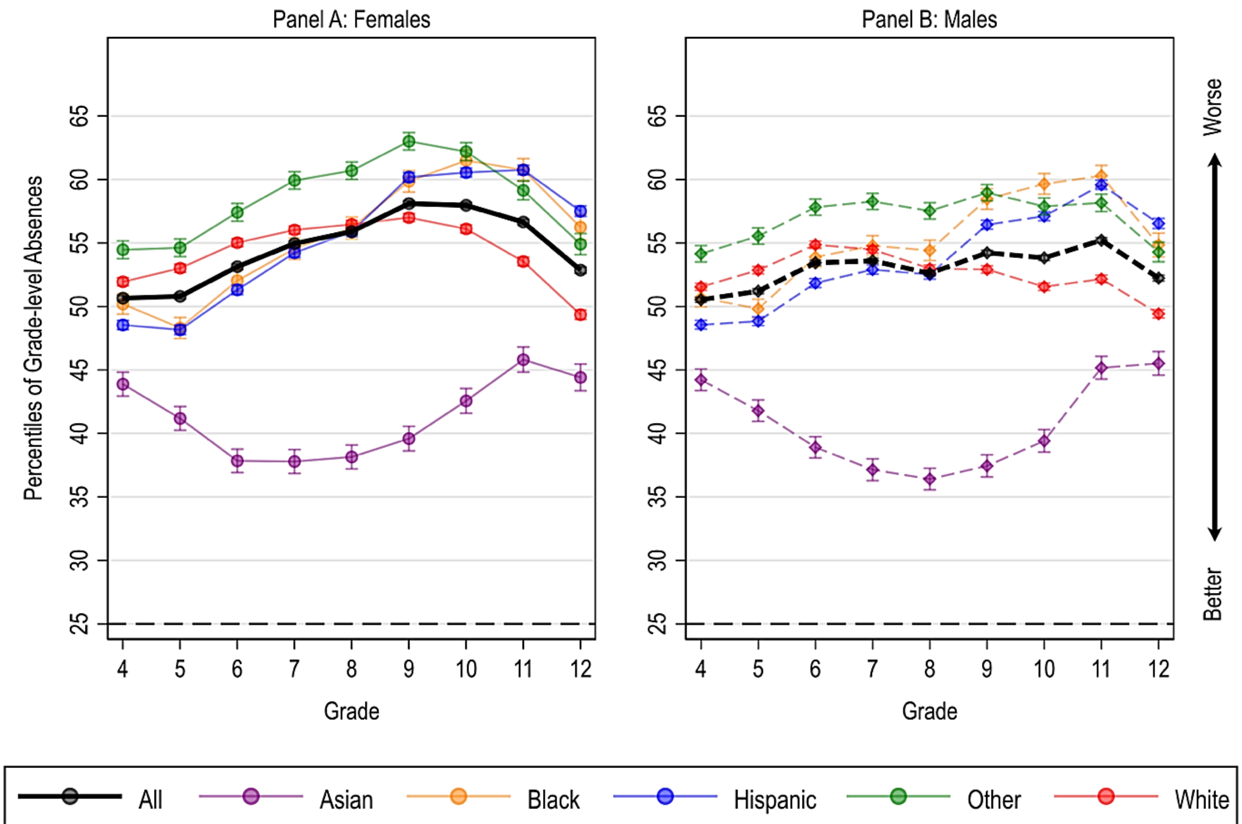
Note. This figure shows estimates of academic mobility in high school disciplinary incidents for low-achieving 3rd graders (i.e., 3rd graders whose averaged math and ELA scores were at the 25th percentile). Results are shown for each student subgroup. Panel A shows results for females (circles) and Panel B shows results for males (diamonds). Racial/ethnic groups are color-coded: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). In each panel, solid horizontal black lines show the average academic mobility for that gender. The dashed horizontal black line is plotted at 25th percentile of the outcome variable. Because higher values of this outcome are undesirable, estimates above the dashed line reflect downward mobility and estimates below it reflect upward mobility. Estimates that cross it reflect no statistically significant academic mobility.

Figure 6. Academic Mobility of Math Test Scores for Low-Performing 3rd Graders from 4th Grade to High School



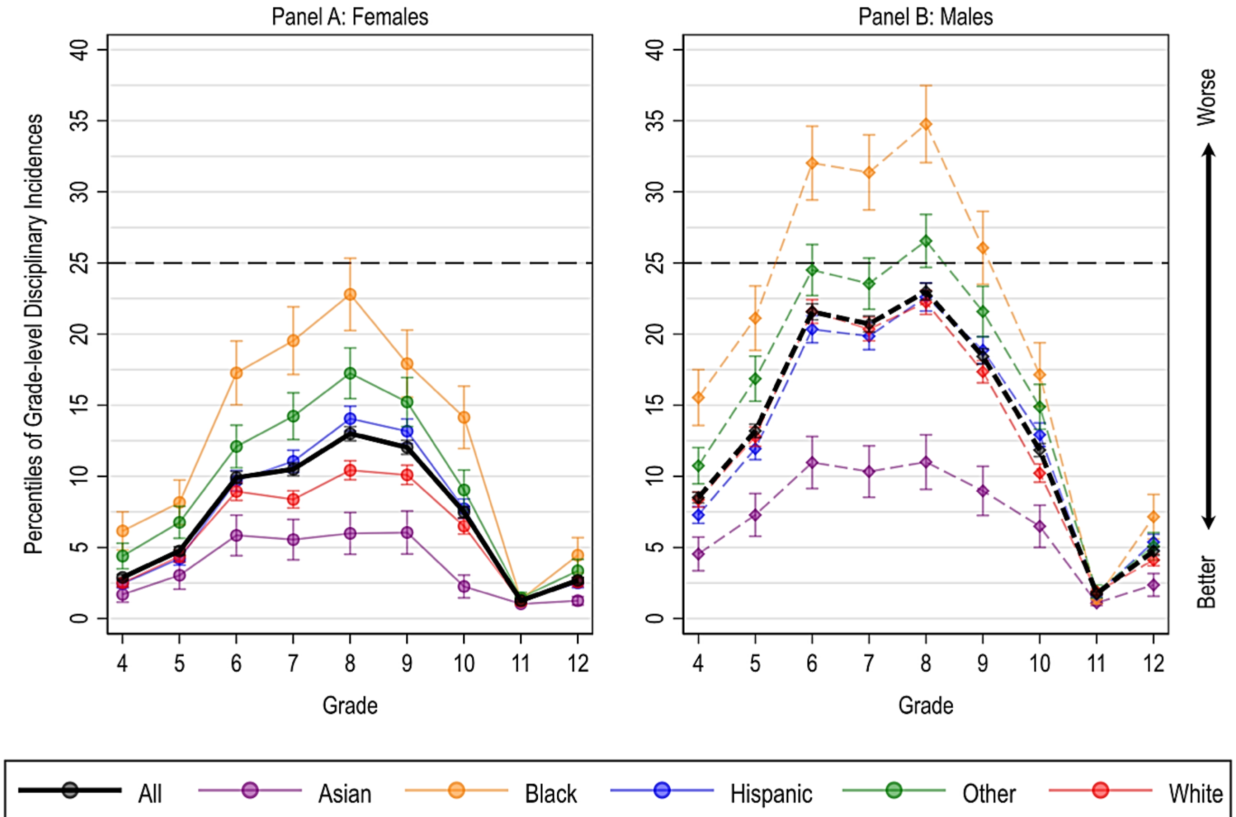
Note. This figure shows the *progression* in academic mobility (with 95% confidence intervals) for low-achieving 3rd graders (i.e., 3rd graders who scored at the 25th percentile) in math test scores for each observable grade from 4th grade to high school. Females are shown in Panel A (circle symbols) and males in Panel B (diamond symbols). Racial/ethnic groups are color-coded as follows: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). Estimates for all students, by gender, are shown via the black trend lines of each panel. The dashed horizontal line is plotted at the 25th percentile of the grade-level outcomes. Estimates above this line reflect upward academic mobility from 3rd grade to each subsequent grade; estimates below this line reflect downward academic mobility.

Figure 7. Academic Mobility in Absenteeism for Low-Performing 3rd Graders from Grades 4-12



Note. This figure shows the *progression* in academic mobility (with 95% confidence intervals) for low-achieving 3rd graders (i.e., 3rd graders whose averaged math and ELA scores were at the 25th percentile) in absenteeism from grades 4-12. Females are shown in Panel A (circle symbols) and males in Panel B (diamond symbols). Racial/ethnic groups are color-coded as follows: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). Estimates for all students, by gender, are shown via the black trend lines of each panel. The dashed horizontal line is plotted at the 25th percentile of the grade-level outcomes. Because higher values of this outcome are undesirable, estimates above the dashed line reflect downward mobility and estimates below it reflect upward mobility. Estimates that cross it reflect no statistically significant academic mobility.

Figure 8. Academic Mobility in Disciplinary Incidences for Low-Performing 3rd Graders from Grades 4-12



Note. This figure shows the *progression* in academic mobility (with 95% confidence intervals) for low-achieving 3rd graders (i.e., 3rd graders whose averaged math and ELA scores were at the 25th percentile) in disciplinary incidences from grades 4-12. Females are shown in Panel A (circle symbols) and males in Panel B (diamond symbols). Racial/ethnic groups are color-coded as follows: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). Estimates for all students, by gender, are shown via the black trend lines of each panel. The dashed horizontal line is plotted at the 25th percentile of the grade-level outcomes. Because higher values of this outcome are undesirable, estimates above the dashed line reflect downward mobility and estimates below it reflect upward mobility. Estimates that cross it reflect no statistically significant academic mobility.

Appendix

Table A.1. Cohorts of 3rd Graders with Observable High School Outcomes

School Year	Cohorts 1	2	3	4	5	6	7	8
2006	**3**	2	1	K	age 4-5	age 3-4	age 2-3	age 1-2
2007	**4**	**3**	2	1	K	age 4-5	age 3-4	age 2-3
2008	**5**	**4**	**3**	2	1	K	age 4-5	age 3-4
2009	**6**	**5**	**4**	**3**	2	1	K	age 4-5
2010	**7**	**6**	**5**	**4**	**3**	2	1	K
2011	**8**	**7**	**6**	**5**	**4**	**3**	2	1
2012	**9**	**8**	**7**	**6**	**5**	**4**	**3**	2
2013	**10**	**9**	**8**	**7**	**6**	**5**	**4**	**3**
2014	**11**	**10**	**9**	**8**	**7**	**6**	**5**	**4**
2015	**12**	**11**	**10**	**9**	**8**	**7**	**6**	**5**
2016	PS1	**12**	**11**	**10**	**9**	**8**	**7**	**6**
2017	PS2	PS1	**12**	**11**	**10**	**9**	**8**	**7**
2018	PS3	PS2	PS1	**12**	**11**	**10**	**9**	**8**
2019	PS4	PS3	PS2	PS1	**12**	**11**	**10**	**9**
2020	PS5	PS4	PS3	PS2	PS1	12	**11**	**10**
2021	PS6	PS5	PS4	PS3	PS2	PS1	**12**	**11**
2022	PS7	PS6	PS5	PS4	PS3	PS2	PS1	**12**

Note. We follow 8 cohorts of 3rd graders through high school (shown in bold). Because of changes to statewide testing policy, the COVID-19 pandemic and different starting dates for the collection of specific data elements, we do not observe all high school outcomes in each year/grade or for all cohorts. We observe high school math and ELA scores for Cohorts 2-7. For all other outcomes, we observe students in at least some grades from every cohort. See Table 1 for details of what grades/cohorts are excluded for each outcome.

Table A.2. OLS Regressions of Subject-Specific 3rd Grade Tests Scores on Student Demographics

	(1)	(2)
	Standardized 3rd Grade Math Score	Standardized 3rd Grade ELA Score
Student Race/Ethnicity (referent: White)		
Asian	0.241***	0.090***
	(0.005)	(0.004)
Black	-0.408***	-0.333***
	(0.006)	(0.005)
Hispanic	-0.195***	-0.170***
	(0.003)	(0.003)
Other race	-0.179***	-0.164***
	(0.004)	(0.004)
Student Gender (referent: Male)		
Female	-0.017***	0.174***
	(0.002)	(0.002)
Student Classifications		
Eligible for Free/Reduced Lunch	-0.393***	-0.400***
	(0.003)	(0.003)
Identified for Special Education Services	-0.504***	-0.545***
	(0.005)	(0.005)
Has a Learning Disability	-0.521***	-0.715***
	(0.008)	(0.008)
Has Limited English Proficiency	-0.557***	-0.682***
	(0.004)	(0.004)
Identified as Highly Capable	0.977***	0.801***
	(0.007)	(0.006)
Constant	0.344***	0.283***
	(0.002)	(0.002)
Observations	581,094	581,094
R^2	0.240	0.271

Note. Results are from OLS regressions and can be interpreted in standard deviation units. Standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.3. Average Marginal Effects from Regressions Predicting Non-Test High School Outcomes with 3rd Grade Math Tests

	(1)	(2)	(3)	(4)	(5)
	No. Adv. Math Courses	No. of Absences	No. of Disciplinary Incidents	12th Grade GPA	Graduated On-time
Student Test Scores (3rd Grade)					
Math	0.269***	-9.788***	-0.250***	0.243***	0.060***
	(0.002)	(0.100)	(0.003)	(0.001)	(0.001)
Student Demographics (3rd Grade)					
Asian	0.355***	-23.266***	-0.490***	0.289***	0.087***
	(0.004)	(0.342)	(0.012)	(0.005)	(0.003)
Black	0.147***	12.516***	0.257***	-0.156***	0.002
	(0.007)	(0.428)	(0.011)	(0.006)	(0.002)
Hispanic	0.076***	7.591***	0.056***	-0.088***	0.020***
	(0.004)	(0.262)	(0.007)	(0.004)	(0.001)
Other Race	0.059***	8.193***	0.121***	-0.105***	-0.018***
	(0.005)	(0.322)	(0.008)	(0.005)	(0.002)
Female	0.011***	2.646***	-0.494***	0.252***	0.045***
	(0.002)	(0.172)	(0.006)	(0.002)	(0.001)
Student Classifications (3rd Grade)					
Free/Reduced-Lunch	-0.263***	25.433***	0.414***	-0.347***	-0.103***
	(0.003)	(0.199)	(0.006)	(0.003)	(0.001)
Special Education	-0.051***	-4.717***	0.004	0.053***	-0.018***
	(0.005)	(0.334)	(0.008)	(0.005)	(0.002)
Learning Disability	-0.216***	8.144***	-0.036**	-0.028***	0.021***
	(0.013)	(0.539)	(0.013)	(0.008)	(0.003)
Limited English	0.060***	-4.555***	-0.181***	0.079***	0.030***
	(0.006)	(0.329)	(0.009)	(0.005)	(0.002)
Highly Capable	0.220***	-8.099***	-0.316***	0.292***	0.039***
	(0.005)	(0.493)	(0.017)	(0.007)	(0.004)
N	521,391	521,391	517,319	467,162	464,855
(Pseudo) R^2	0.099	0.011	0.045	0.201	0.084

Note. All results are reported as average marginal effects. Model specifications vary depending on the outcome. We estimate non-linear models with negative binomial (Columns 1-3) or logistic (Column 5) regressions. Standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.4. Average Marginal Effects from Regressions Predicting Non-Test High School Outcomes with 3rd Grade ELA Tests

	(1)	(2)	(3)	(4)	(5)
	No. Adv. ELA Courses	No. of Absences	No. of Disciplinary Incidents	12th Grade GPA	Graduated On-time
Student Test Scores (3rd Grade)					
ELA	0.204***	-8.734***	-0.240***	0.266***	0.054***
	(0.002)	(0.102)	(0.003)	(0.001)	(0.001)
Student Demographics (3rd Grade)					
Asian	0.341***	-24.362***	-0.509***	0.250***	0.094***
	(0.005)	(0.342)	(0.012)	(0.005)	(0.003)
Black	0.299***	13.535***	0.278***	-0.130***	-0.005*
	(0.007)	(0.428)	(0.011)	(0.006)	(0.002)
Hispanic	0.141***	8.021***	0.066***	-0.077***	0.017***
	(0.004)	(0.262)	(0.007)	(0.004)	(0.001)
Other Race	0.096***	8.663***	0.129***	-0.098***	-0.021***
	(0.005)	(0.323)	(0.008)	(0.004)	(0.002)
Female	0.140***	4.265***	-0.447***	0.299***	0.034***
	(0.003)	(0.173)	(0.006)	(0.002)	(0.001)
Student Classifications (3rd Grade)					
Free/Reduced-Lunch	-0.199***	25.737***	0.413***	-0.340***	-0.104***
	(0.003)	(0.200)	(0.006)	(0.003)	(0.001)
Special Education	-0.086***	-4.635***	-0.003	0.055***	-0.020***
	(0.006)	(0.336)	(0.009)	(0.005)	(0.002)
Learning Disability	-0.176***	6.902***	-0.080***	-0.062***	0.029***
	(0.012)	(0.542)	(0.013)	(0.008)	(0.003)
Limited English	0.083***	-4.946***	-0.195***	0.064***	0.033***
	(0.005)	(0.332)	(0.009)	(0.005)	(0.002)
Highly Capable	0.217***	-10.570***	-0.358***	0.228***	0.052***
	(0.006)	(0.491)	(0.017)	(0.007)	(0.004)
N	521,391	521,391	517,319	467,162	464,855
(Pseudo) R^2	0.049	0.011	0.045	0.214	0.089

Note. All results are reported as average marginal effects. Model specifications vary depending on the outcome. We estimate non-linear models with negative binomial (Columns 1-3) or logistic (Column 5) regressions. Standard errors in parentheses * p < 0.05, ** p < 0.01, *** p < 0.001.

Table A.5. Average Marginal Effects from Regressions Predicting High School Outcomes with 8th Grade Predictors

	(1)	(2)	(3)	(4)	(5)	(6)
	No. Adv. Math Courses	No. Adv. ELA Courses	No. of Absences	No. of Disciplinary Incidents	12th Grade GPA	Graduated On-time
Student Test Scores (8th Grade)						
8th Grade Math	0.324***	0.154***	-14.781***	-0.291***	0.319***	0.070***
	(0.003)	(0.002)	(0.135)	(0.004)	(0.002)	(0.001)
8th Grade ELA	0.119***	0.189***	-6.038***	-0.207***	0.191***	0.034***
	(0.002)	(0.002)	(0.136)	(0.004)	(0.002)	(0.001)
Student Demographics (8th Grade)						
Asian	0.249***	0.247***	-18.684***	-0.367***	0.127***	0.058***
	(0.004)	(0.005)	(0.333)	(0.012)	(0.004)	(0.003)
Black	0.206***	0.369***	8.352***	0.168***	-0.046***	0.018***
	(0.007)	(0.007)	(0.433)	(0.010)	(0.006)	(0.002)
Hispanic	0.099***	0.162***	3.775***	-0.024***	-0.020***	0.029***
	(0.004)	(0.005)	(0.248)	(0.006)	(0.003)	(0.001)
Other Race	0.074***	0.116***	5.714***	0.083***	-0.062***	-0.008***
	(0.005)	(0.005)	(0.325)	(0.008)	(0.004)	(0.002)
Female	-0.022***	0.136***	5.412***	-0.392***	0.234***	0.027***
	(0.003)	(0.003)	(0.175)	(0.006)	(0.002)	(0.001)
Student Classifications (8th Grade)						
Free/Reduced-Lunch	-0.175***	-0.134***	21.456***	0.301***	-0.238***	-0.075***
	(0.003)	(0.003)	(0.198)	(0.006)	(0.003)	(0.001)
Special Education	-0.033***	-0.065***	-6.904***	0.043***	0.151***	-0.004*
	(0.006)	(0.006)	(0.434)	(0.010)	(0.006)	(0.002)
Learning Disability	-0.128***	-0.153***	1.163*	-0.240***	0.079***	0.062***
	(0.014)	(0.013)	(0.566)	(0.014)	(0.008)	(0.003)
Limited English	0.060***	0.069***	-3.283***	-0.172***	0.064***	0.027***
	(0.006)	(0.006)	(0.461)	(0.011)	(0.006)	(0.002)
Highly Capable	0.114***	0.109***	2.040***	-0.035**	-0.105***	0.011***
	(0.005)	(0.006)	(0.350)	(0.012)	(0.005)	(0.003)
N	459,309	459,309	459,309	456,034	416,762	420,372
(Pseudo) R^2	0.142	0.067	0.018	0.073	0.325	0.137

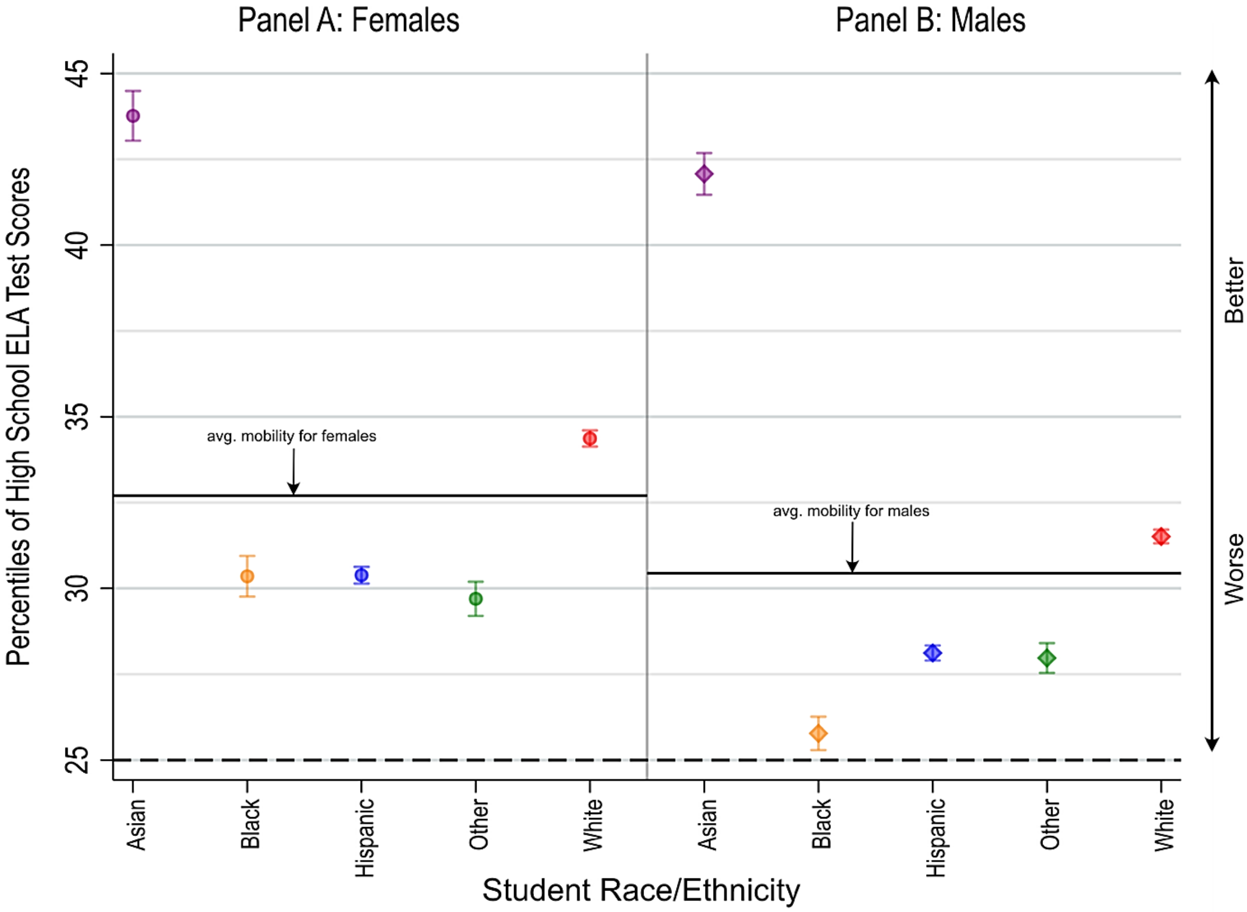
Note. All results are reported as average marginal effects. Model specifications vary depending on the outcome. We estimate non-linear models with negative binomial (Columns 1-4) or logistic (Column 6) regressions.
Standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.6. Beta Coefficients and Intercepts from EIV Regressions for Each High School Outcome and Student Subgroup

High School Outcome	Statistic	Student Subgroups									
		Female					Male				
		Asian	Black	Hispanic	Other	White	Asian	Black	Hispanic	Other	White
Test Scores											
Standardized High School Math Score	β	0.72	0.67	0.66	0.76	0.72	0.78	0.69	0.69	0.79	0.77
		(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
	α	24.19	12.37	13.21	10.58	14.81	21.09	8.76	10.95	8.62	12.13
		(0.58)	(0.44)	(0.20)	(0.36)	(0.17)	(0.56)	(0.43)	(0.20)	(0.37)	(0.18)
Standardized High School ELA Score	β	0.68	0.67	0.68	0.78	0.71	0.71	0.66	0.69	0.76	0.71
		(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
	α	26.81	13.69	13.28	10.31	16.68	24.30	9.22	10.96	8.90	13.86
		(0.55)	(0.49)	(0.21)	(0.40)	(0.18)	(0.47)	(0.41)	(0.18)	(0.35)	(0.16)
Non-test Outcomes											
Number Adv. Math Courses (9-12 grade)	β	0.64	0.33	0.32	0.42	0.45	0.71	0.35	0.34	0.48	0.49
		(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
	α	2.50	3.19	1.59	-3.10	-5.87	0.86	1.12	0.20	-4.55	-6.78
		(0.70)	(0.48)	(0.20)	(0.36)	(0.16)	(0.58)	(0.39)	(0.18)	(0.31)	(0.14)
Number Adv. ELA Courses (9-12 grade)	β	0.41	0.29	0.32	0.38	0.43	0.46	0.27	0.29	0.35	0.38
		(0.01)	(0.02)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
	α	18.19	18.54	10.73	6.24	1.50	11.06	12.62	6.00	2.84	-1.75
		(0.76)	(0.65)	(0.27)	(0.46)	(0.19)	(0.63)	(0.52)	(0.21)	(0.37)	(0.15)
Number of Absences (9-12 grade)	β	-0.22	-0.24	-0.25	-0.31	-0.20	-0.26	-0.24	-0.24	-0.32	-0.21
		(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
	α	49.47	68.93	68.79	70.71	60.85	49.02	67.79	66.80	68.31	58.48
		(0.57)	(0.51)	(0.22)	(0.41)	(0.18)	(0.52)	(0.47)	(0.22)	(0.38)	(0.16)
Number of Disciplinary Incidences (9-12 grade)	β	-0.09	-0.31	-0.18	-0.28	-0.18	-0.17	-0.36	-0.28	-0.34	-0.29
		(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.02)	(0.01)	(0.01)	(0.00)
	α	10.77	34.28	22.44	30.41	20.93	20.07	48.94	36.88	42.18	35.48
		(0.45)	(0.71)	(0.29)	(0.55)	(0.20)	(0.55)	(0.71)	(0.32)	(0.57)	(0.23)
High School GPA in 12th Grade	β	0.45	0.39	0.42	0.51	0.48	0.49	0.35	0.39	0.47	0.46
		(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
	α	41.47	27.88	28.85	24.71	30.95	32.89	20.86	21.11	18.98	23.05
		(0.50)	(0.44)	(0.21)	(0.37)	(0.17)	(0.45)	(0.36)	(0.18)	(0.33)	(0.15)
Student Graduated High School On Time	β	0.14	0.27	0.24	0.32	0.25	0.20	0.39	0.33	0.39	0.31
		(0.00)	(0.00)	(0.00)	(0.00)	0.00	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	α	85.96	71.90	75.64	67.32	74.50	81.43	60.84	66.55	60.51	68.01
		(0.08)	(0.07)	(0.03)	(0.06)	(0.03)	(0.09)	(0.09)	(0.03)	(0.05)	(0.03)

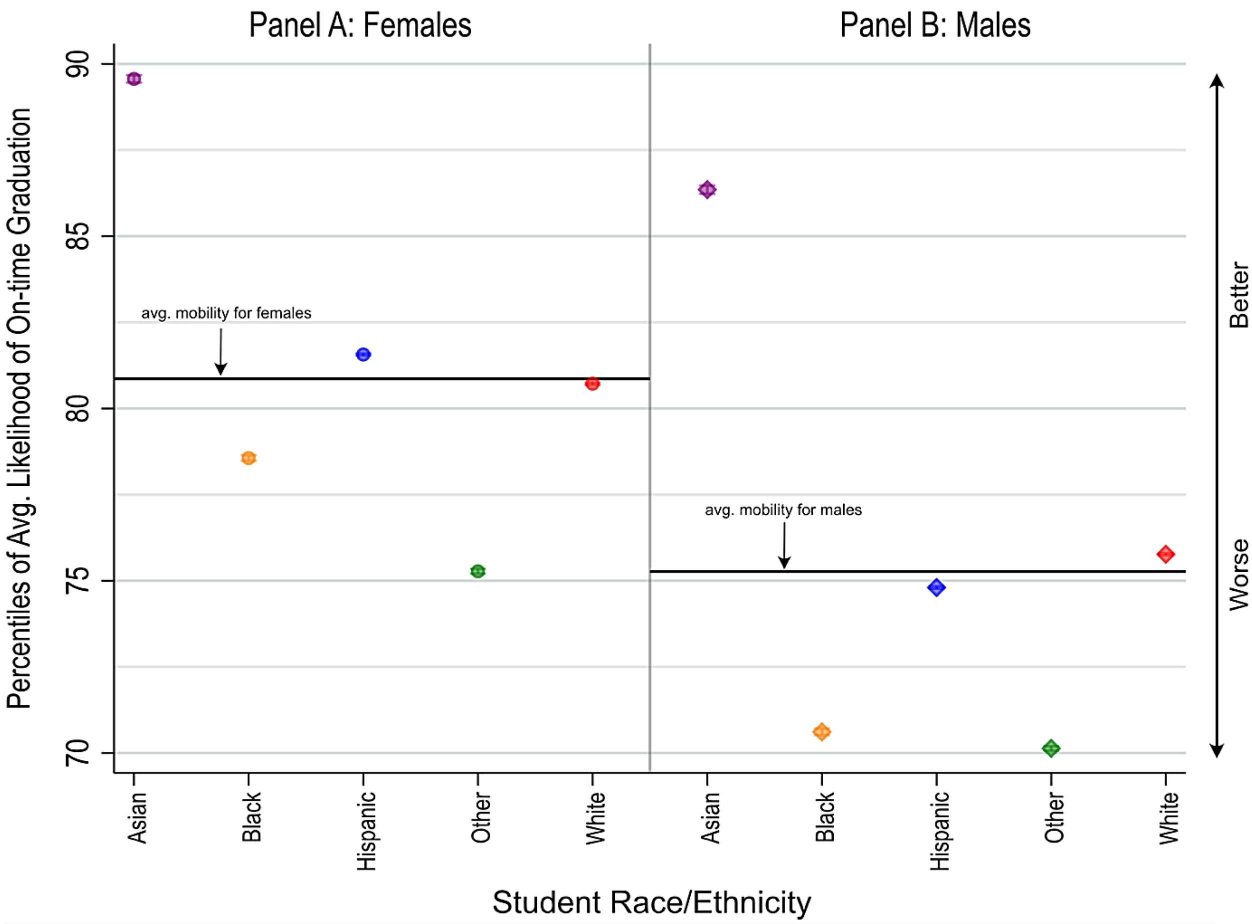
Note. This table reports beta coefficients and y-intercepts from rank-to-rank errors-in-variables (EIV) regressions derived from equation (2) for each high school outcome and each student subgroup. (Standard errors in parentheses.)

Figure A.1. Academic Mobility in High School ELA Test Scores for Low-Performing 3rd Graders, by Race/Ethnicity and Gender



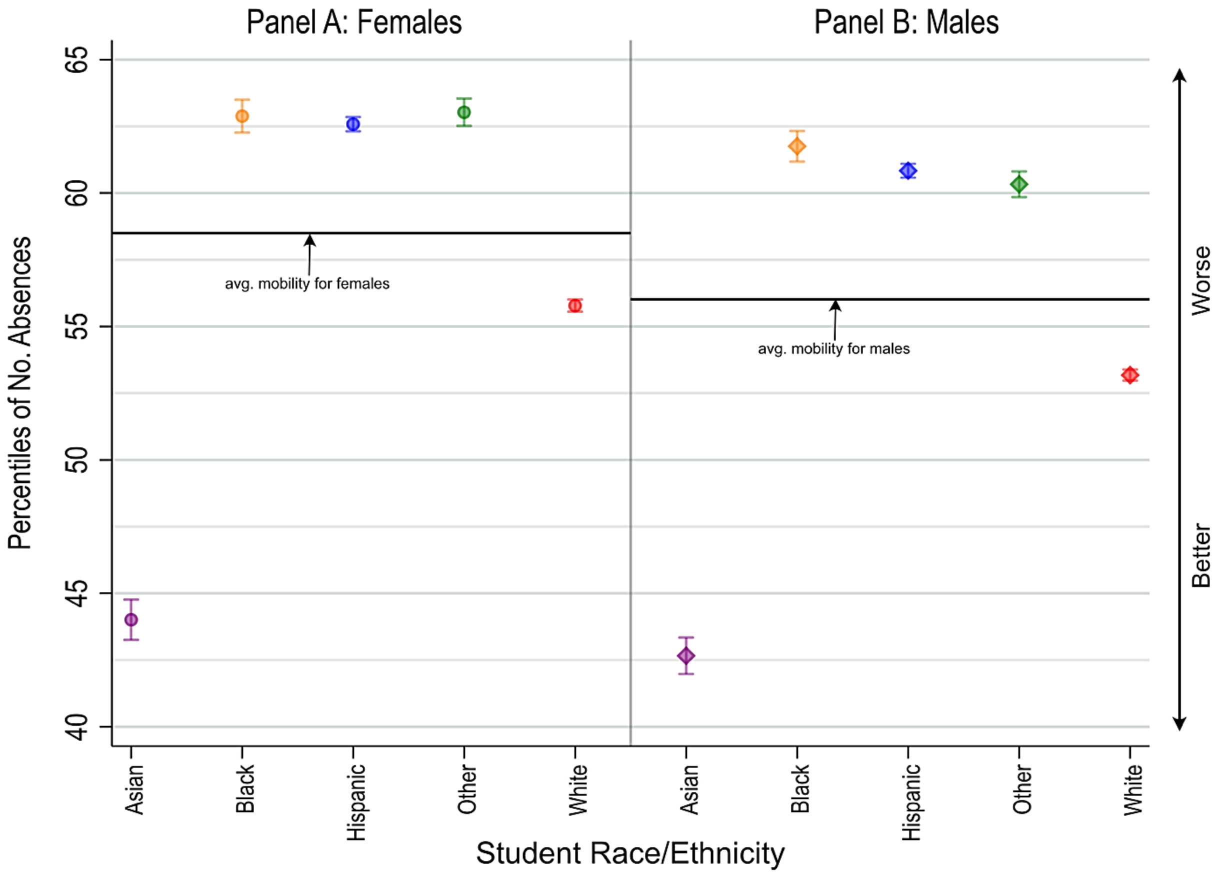
Note. This figure shows estimates of academic mobility in high school ELA tests for low-achieving 3rd graders (i.e., 3rd graders who scored in the 25th percentile on ELA tests). Results are shown for each student subgroup. Panel A shows results for females (circle symbols) and Panel B shows results for males (diamond symbols). Racial/ethnic groups are color-coded: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). In each panel, solid horizontal black lines show the average academic mobility for that gender. The dashed horizontal black line is plotted at 25th percentile of the outcome variable. Estimates above the dashed line reflect upward mobility, estimates below it reflect downward mobility and estimates that cross it reflect no statistically significant academic mobility.

Figure A.2. Academic Mobility in On-time Graduation for Low-Performing 3rd Graders, by Race/Ethnicity and Gender



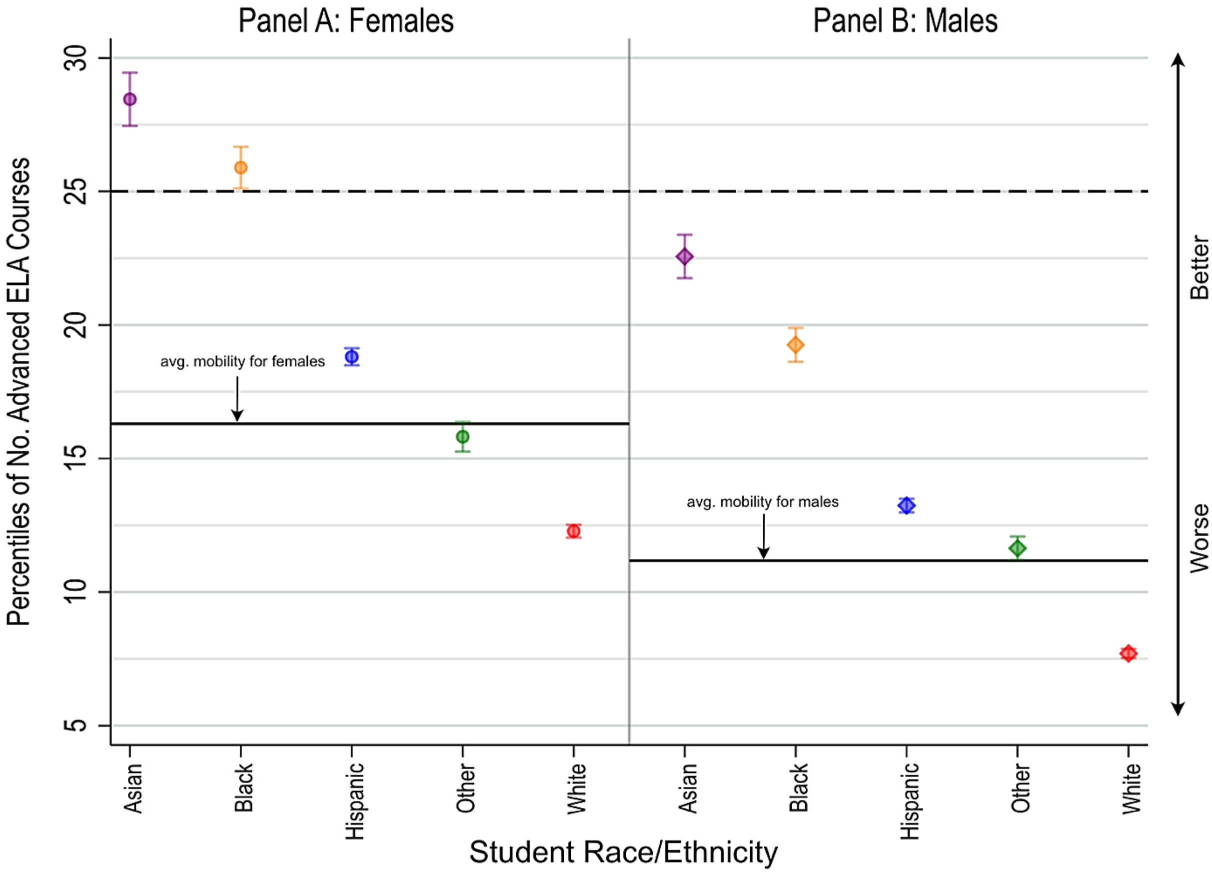
Note. This figure shows estimates of academic mobility in on-time high school graduation for low-achieving 3rd graders (i.e., 3rd graders whose averaged math and ELA scores were at the 25th percentile). Results are shown for each student subgroup. Panel A shows results for females (circles) and Panel B shows results for males (diamonds). Racial/ethnic groups are color-coded: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). In each panel, solid horizontal black lines show the average academic mobility for that gender. The dashed horizontal black line is plotted at 25th percentile of the outcome variable. Estimates above the dashed line reflect upward mobility, estimates below it reflect downward mobility and estimates that cross it reflect no statistically significant academic mobility.

Figure A.3. Academic Mobility in Absenteeism for Low-Performing 3rd Graders, by Race/Ethnicity and Gender



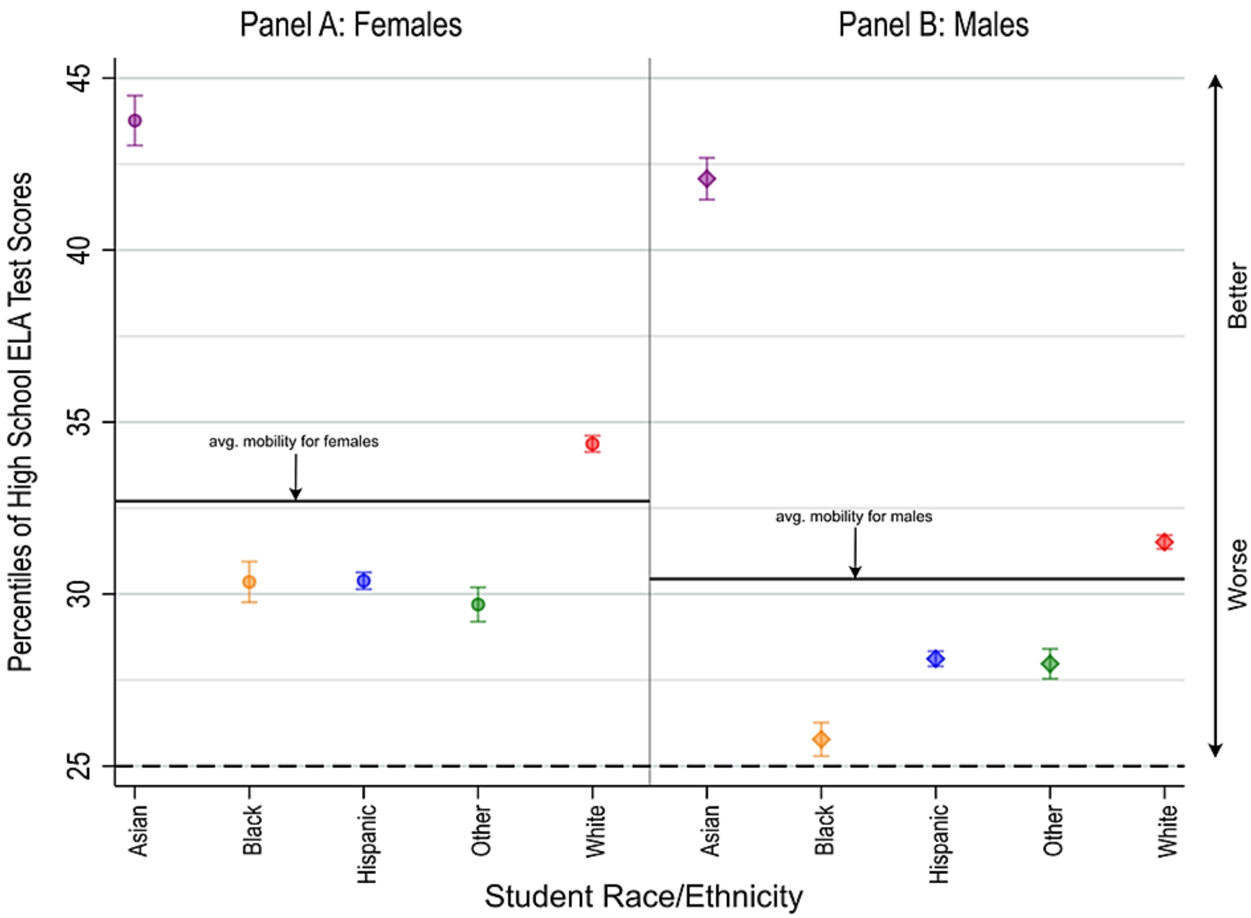
Note. This figure shows estimates of academic mobility in high school absenteeism for low-achieving 3rd graders (i.e., 3rd graders whose averaged math and ELA scores were at the 25th percentile). Results are shown for each student subgroup. Panel A shows results for females (circles) and Panel B shows results for males (diamonds). Racial/ethnic groups are color-coded: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). In each panel, solid horizontal black lines show the average academic mobility for that gender. The dashed horizontal black line is plotted at 25th percentile of the outcome variable. Because higher values of this outcome are undesirable, estimates above the dashed line reflect downward mobility and estimates below it reflect upward mobility. Estimates that cross it reflect no statistically significant academic mobility.

Figure A.4. Academic Mobility in College-level ELA Course-taking for Low-Performing 3rd Graders, by Race/Ethnicity and Gender



Note. This figure shows estimates of academic mobility in high school ELA course-taking for low-achieving 3rd graders (i.e., 3rd graders whose averaged math and ELA scores were at the 25th percentile). Results are shown for each student subgroup. Panel A shows results for females (circles) and Panel B shows results for males (diamonds). Racial/ethnic groups are color-coded: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). In each panel, solid horizontal black lines show the average academic mobility for that gender. The dashed horizontal black line is plotted at 25th percentile of the outcome variable. Estimates above the dashed line reflect upward mobility, estimates below it reflect downward mobility and estimates that cross it reflect no statistically significant academic mobility.

Figure A.5. Academic Mobility of ELA Test Scores for Low-Performing 3rd Graders from 4th Grade to High School



Note. This figure shows estimates of academic mobility in high school ELA tests for low-achieving 3rd graders (i.e., 3rd graders who scored in the 25th percentile on ELA tests). Results are shown for each student subgroup. Panel A shows results for females (circle symbols) and Panel B shows results for males (diamond symbols). Racial/ethnic groups are color-coded: Asian (purple), Black (orange), Hispanic (blue), Other (green) and White (red). In each panel, solid horizontal black lines show the average academic mobility for that gender. The dashed horizontal black line is plotted at 25th percentile of the outcome variable. Estimates above the dashed line reflect upward mobility, estimates below it reflect downward mobility and estimates that cross it reflect no statistically significant academic mobility.