CEDR Working Paper | No. 03212025-1

CEDR
Center for Education
Data & Research

Predicting Who Attends (2-Year and 4-Year) College Based on Early (3rd Grade) and Later Test Scores

Dan Goldhaber
University of Washington
American Institutes for Research

Stephanie Liddle
University of Washington

The Center for Education Data & Research
University of Washington Bothell
3610 Albion Pl. N. Ste. A Seattle, WA 98103
(206) 547-5585 cedr@u.washington.edu
www.cedr.us

Acknowledgements: This work is supported by the Bill & Melinda Gates Foundation. All opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the institutions to which the authors are affiliated or the study's funder. We are grateful to the Education Data Research Center and Washington's Office of Superintendent of Public Instruction for providing access to the data that made the analyses in this paper possible. We also wish to acknowledge useful comments from Michael DeArmond. Correspondence regarding this article should be addressed to Dan Goldhaber, Center for Education Data and Research, University of Washington, 3610 Albion Pl. N, Suite A, Seattle, WA 98103. Email: dgoldhab@uw.edu • www.cedr.us

Suggested citation:
Goldhaber, D.; Liddle, S. (2025). Predicting Who Attends (2-Year and 4-Year) College Based on Early (3rd Grade) and Later Test Scores. CEDR Working Paper 03252025-1. University of Washington, Seattle, WA.

You can access other CEDR publications at
http://www.CEDR.us/publications.html

Abstract: The current literature on early predictors of postsecondary outcomes focuses almost exclusively on the relationship between postsecondary outcomes and students' academic achievements and experiences in middle and high school. These relationships are important, but policymakers and educators might get a more complete picture of how K12 relates to college success if we also consider the role played by earlier experiences in schooling. That is what we do in this paper. We use statewide censuses of multiple student cohorts in Washington state to descriptively examine the relationship between students' 3rd grade test performance and college enrollment outcomes. We find that students' 3rd grade test scores strongly predicted the likelihood of enrolling in a 4-year college—roughly as well as high school tests—with significant variation across performance deciles and by students' race/ethnicity and gender.

1 INTRODUCTION

Decades of research show that having a college education yields significant benefits for individuals and society. College graduates enjoy better job opportunities, higher earnings, and better health outcomes than non-graduates (McMahon, 2009; Hout, 2012). College educated citizens also promote democracy and economic growth, and reduce crime and welfare costs (McMahon, 2009; Heckman et al., 2017; Lochner 2011). Even people who attend college without graduating have higher employment rates and earnings than people who never enroll (Giani et al., 2020). Although educational attainment has recently become a fault line in American politics (Dougherty et al., 2024), the benefits of going to college are clear.

Given these benefits, many states and cities have adopted policies that encourage postsecondary enrollment. For example, nearly 100 communities across the country have launched "promise programs" that provide place-based scholarships for students to attend community and technical colleges (e.g., Hershbein et al., 2021). Some states have postsecondary scholarships for students from low-income households, such as Indiana's 21st Century Scholars and Washington's College Bound program.[1] Besides scholarships, states and localities have also implemented a range of initiatives and programs designed to help students think about and plan for life after high school. For example, Washington state—the focus of the current study—requires every student to complete a High School and Beyond Plan (HSBP) before graduating from high school.

Washington's students start exploring their college and career interests via the HSBP in middle school. By beginning with middle schoolers, the HSBP reflects well-established research showing that secondary school experiences matter to postsecondary success. There is much less evidence, however, on whether earlier interventions could also have implications for postsecondary success. To gain a more comprehensive understanding of how students' *entire* K12 journey shapes postsecondary access and success—and to identify other possible points for intervention—researchers and policymakers need to consider students' earlier experiences in school. Luckily, although once untenable, today's state data systems enable linkages across K12 and postsecondary institutions in ways that make this type of investigation possible.

With all that in mind, this paper leverages longitudinal administrative data from Washington state to examine how students' early academic performance—specifically third-grade test scores—predict postsecondary enrollment. We also analyze patterns across different student groups and time periods, motivated by both historical disparities in college enrollments and shifts in state policy over time (e.g., HSBP and College Bound). We address four research questions (RQs):

> RQ 1. To what extent is the probability of enrolling in college predicted by early (3rd grade) test performance and student background characteristics?

[1] Although college completion rates improved between 2011 and 2021, gaps remain: rates increased from 34% to 42% among white Americans, 20% to 28% among Black Americans, 50% to 61% among Asian Americans, and 14% to 21% among Hispanic Americans (census.gov).

RQ 2. Does the predictive power of early tests vary across different racial/ethnic groups and by gender?

RQ 3. To what extent are students anchored to their 3rd grade test scores? And to what extent are tests at different points in a student's academic career predictive of college enrollment?

RQ 4. Does the predictive power of early tests vary across cohorts?

Overall, our results suggest that students' educational trajectories can crystalize early, with elementary outcomes serving as remarkably persistent predictors of future academic success. For example, we find that students' postsecondary enrollment outcomes are strongly predicted by their performance on 3rd grade tests. Students' chances of enrolling in a 4-year college differs substantially based on their academic performance in 3rd grade. The lowest-performing 3rd graders (in the first decile) have about a 6% chance of enrolling in a 4-year college while the highest-performing 3rd graders (in the tenth decile) have about a 48% chance of enrolling in a 4-year college. The likelihood of college enrollment also varies by students' race/ethnicity and gender, with Asian and female students being significantly more likely to enroll in a 4-year college than other students at every point along the 3rd grade test score distribution. However, we find much less variability in the likelihood of attending a 2-year college across the 3rd grade test score distribution and by race/ethnicity.

The disparities we identified highlight other nuances. For example, the gap in college-going rates between high- and low-performing students within each racial/ethnic group steadily increased as students progressed through grades. Within performance deciles, gaps in the likelihood of 4-year college enrollment vary by race/ethnicity as students progressed through grades. High-achieving Asian students, for example, were significantly more likely to enroll in a 4-year college than non-Asians students in every grade, while high-achieving White students were less likely to enroll compared to non-White peers.

Despite statewide initiatives designed to increase college enrollment, we found little evidence that students' likelihood of attending a 4-year college—as predicted by their 3rd-grade test scores—improved across cohorts in ways that reflect the state's goal of supporting historically underserved populations. For example, while we find a slight increase in 4-year college attendance among low-achieving students, we also find that economically disadvantaged students experienced more modest enrollment gains than other groups.

The data we use in this study allow us to make two contributions to the literature on postsecondary outcomes. First, while it is well-known that students' academic experiences in secondary school influence their chances of enrolling in college, there is almost no research connecting students' elementary school experiences to their college enrollment outcomes. We address this gap by following students from 3rd grade all the way through to their first year after high school and into college. Second, what we know about the relationship between college enrollment and earlier schooling experiences is almost entirely based on longitudinal *samples*.

To our knowledge, this is the only study that examines these issues using a statewide *census of students* across multiple cohorts.

2 BACKGROUND ON THE FACTORS PREDICTING COLLEGE ENROLLMENT

A range of factors influence whether high school graduates will enroll in college. Academic preparedness, including high school grades, standardized test scores, and rigorous coursework—such as advanced placement and/or dual enrollment courses—significantly impacts college-going decisions (Schaller et al., 2023; Kremer, 2020; Bowers, 2019). Research also suggests that family income and background play an important role in pursuing higher education (Lovenheim & Reynolds, 2012). Students from families with a history of college attendance, for example, are more likely to pursue higher education themselves (Goodman et al., 2015). There is also clear evidence that college enrollment reflects students' personal ambitions and career goals, although these too may be rooted in family and community contexts (Glick & White, 2004; Mello, 2008).

As noted earlier, much of the existing research leverages students' academic achievements and experiences in high school to predict postsecondary schooling outcomes (for recent examples see: Morris & McKenzie, 2024; Reber et al, 2024; Center for Education Policy Research, 2024). Given high school's connection to college enrollment (e.g., the use of high school transcripts for admissions decisions), this focus on high school makes sense.

There are, however, compelling reasons to look earlier in students' educational experiences to further examine the possible origins of postsecondary attainment. Some studies, for example, have looked slightly farther back and found positive effects of middle school performance on college enrollment and readiness (Holzmann & Duffy, 2024; Bobek & San Pedro, 2021; see Hein et al. (2013) for a review of studies before the last decade). But other research suggests students' level of academic performance may get established even earlier, when they start formal schooling or even before, and that this positioning is remarkably persistent throughout childhood and adolescence (see Alexander, Entwisle, and Dauber 2003; Dauber, Alexander, and Entwisle 1996; Entwisle et al., 2005). This observation about the importance of early schooling is not new. Over 60 years ago Bloom (1964, p.110) wrote: "All subsequent learning in the school is affected and in large part determined by what the child has learned . . . by the end of grade 3."

Despite the clear importance of early educational experiences, quantitative research on the link between elementary school and postsecondary outcomes remains surprisingly limited. This gap in the research has long been noted. Two decades ago, Entwisle and colleagues (2005) highlighted the absence of any national study that could directly connect primary grade school experiences to educational outcomes in young adulthood. Almost a decade later, the same gap was noted in a review by Hein et al. (2013) of the literature on predicting postsecondary outcomes from pre-school to high school. To our knowledge, the gap remains largely unaddressed (for an exception, see Chetty et al., 2011).

3 DATA, MEASURES, AND ANALYTICAL SAMPLE

The data for this study come from several datasets maintained by Washington state's Office of the Superintendent of Public Instruction (OSPI) and the Education Research and Data Center (ERDC) which is the state agency tasked with maintaining the state's P20W data warehouse. The sources of data from OSPI are the Core Student Record System (CSRS) and the Comprehensive Education Data and Research System (CEDARS). From ERDC we use data from the State Board for Community and Technical Colleges Data Warehouse (SBCTC), the Public Centralized Higher Education Enrollment System (PCHEES), and the National Student Clearinghouse (NSC). Individual student records from each data source are longitudinally linkable across years based on a unique student record identifier (ID) maintained by the state.

OSPI Data: CSRS and CEDARS

The CSRS and CEDARS data include annual information for all students enrolled in public schools in Washington state and include demographics (e.g., race and gender), various classifications (e.g., free or reduced-price lunch eligibility, special education, and limited English proficiency status), and educational outcomes. CSRS includes this information in earlier school years from 2004-05 to 2008-09 and CEDARS for later school years from 2009-10 to 2021-22. These systems also include test outcomes for students in 3rd - 8th and 10 or 11th grade in math and reading/English language arts (ELA).[2,3] See **Table 1** for a summary of data holdings across years. As tests have changed over time, we standardize test scores within subject, year, and grade to be mean zero and standard deviation of one. We use students' *averaged* standardized scores on math and ELA tests in different grades to predict college enrollment. From this point on when we refer to test scores, we are referring to these averaged scores.[4] If a student is missing a score in either subject we use the score from whatever subject is observed.[5] For high school tests, we only include scores on the Smarter Balanced Assessments (SBA), which have been required of 11th graders since the 2014-15 school year and 10th graders since the 2017-18 school year.

ERDC Data: SBCTC, PCHEES, and NSC

We merge student-level data from OSPI to various postsecondary data holdings from ERDC that exist from 2013-14 to 2021-22 school years. From the SBCTC data we obtain enrollment data for students enrolled in Washington state's public 2-year community and technical colleges. From PCHEES we obtain that same information from the state's public 4-year colleges. NSC

[2] The tests in Washington changed several times over the period for which we have data. ELA tests changed in the 2014-15 school year while math tests changed twice over the period, first in the 2009-10 and then again in the 2014-15 school year.

[3] Not all students were tested annually. For example, 7.1% of all 3rd graders between 2006 and 2012 had no math or ELA scores. Our samples include only tested students.

[4] When we examine relationships using subject-specific test scores and predictors (i.e., only math or reading), the results are similar and available from the authors upon request.

[5] Of the 507,101 students in our baseline analytic sample, 599 (0.1%) were missing math scores and 1,567 (0.3%) were missing ELA scores.

provides information to ERDC for students who enrolled in private colleges within Washington state as well as students enrolled in public or private colleges outside of Washington. Though as has been documented (see Dynarski et al., 2015), the NSC data are likely to provide a somewhat less complete picture of college-going behavior amongst those students who enrolled in private colleges or outside of Washington.[6] Note, that our data linkages exclude students who attended private school throughout their K-12 education.

Using these data, we construct the main outcome of interest in this paper: an indicator for whether a student was enrolled in a 2-year college, a 4-year college, or not enrolled in a college within one year after their expected year of high school graduation. Students who enrolled in both a 2-year and a 4-year college within the first year after high school are categorized as being enrolled in a 4-year college.[7] We do not distinguish between public/private or in-state/out-of-state enrollment.

3.1 Analytical Samples and Sample Statistics

Our baseline analytic sample is anchored to seven 3rd grade cohorts, from 2006 to 2012, for whom we observe demographic characteristics and 3rd grade test scores (see **Table 2** for cohort setup). This includes 507,101 individual students. For some analyses we restrict the sample to a subset of 3rd graders for whom we also observe test scores in high school. In this subsample we retain 326,943 3rd graders (or 64.5 percent of the baseline sample). We link students in both samples to their college enrollment outcomes within one year after their expected high school graduation.[8]

We report summary statistics for each sample by students' demographics, classifications and 3rd grade test performance in **Table 3.**[9] In comparing the 3rd grade sample (Column 1) and the high school sample (Column 2), we see very little difference in demographic composition or student classifications. Students retained in the high school sample had slightly higher test

[6] For instance, while NSC reports 97% coverage of all students in public and private U.S. colleges, Dynarski et al. (2015) point out that coverage rates may vary for different types of colleges and across student subgroups—e.g., they find lower coverage for colleges with high proportions of first-time or full-time-equivalent students, private for-profit colleges, less selective colleges; and for Black and Hispanic students. Though coverage may have improved in recent years.

[7] There were 10,481 students (or 2.1%) who enrolled in both a 2-year and a 4-year college in the year after high school. They are categorized here as having enrolled in a 4-year college.

[8] Students are in our samples regardless of whether we observe them graduating from a public high school in Washington state. Of the 3rd graders in our baseline sample 143,147 (28.2%) have no observed graduation event from a public high school in Washington state. This doesn't necessarily mean that these students didn't graduate from high school as they could have graduated some time the following year, completed a General Educational Development (GED) credential, received a diploma from a private school in the state, or graduated from an out-of-state high school. Some of these individuals are observed in the postsecondary enrollment data (15,526 (10.9%) in 2-year colleges and 7,119 (5.0%) in 4-year colleges).

[9] 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. Collectively these students comprise 7.9% of 3rd graders across all cohorts. If in the text we refer to students of these races/ethnicities simply as "Other" it is purely for efficiency in writing.

scores, a result consistent with prior research showing that more mobile students tend to have lower test scores (Goldhaber et al., 2022).

In **Table 4** we report college enrollment outcomes by racial/ethnic groups and sample. Compared to 2-year college enrollment, racial/ethnic differences in 4-year college enrollment are much larger, regardless of sample. And while the magnitude of ethno-racial enrollment gaps stayed about the same for 2-year college enrollment across samples (5.5 percentage points), for 4-year colleges the difference in enrollment percentages between the least (Hispanic) and most enrolled (Asian) racial/ethnic group increased between the baseline (32.5) and restricted sample (37.2).

In **Table 5** we report student characteristics by college-going status and sample. Panel A shows this information for 3rd graders in our baseline sample. Panel B shows the same information for our restricted sample of 3rd graders whom we observe in high school. The t-tests in columns 1-2 and 4-5 test the difference in each variable between students in these sample-specific college-going categories and students in the same sample who were not enrolled in college. So, Columns 1 and 2 are compared to Column 3, and Columns 4 and 5 are compared to Column 6.

Overall, 47.2% of 3rd graders attended some college after high school. Of these students, 43.8% attended a 2-year college and 56.2% attend a 4-year college. Again, we see only small differences in student characteristics between the baseline 3rd grade sample and the restricted 3rd grade-to-high school subset within a given college outcome (i.e., compare Column 1 to Column 4, Column 2 to Column 5, and Column 3 to Column 6). Still, there is substantial variation in the types of students who attended a 2- or 4-year college and those who did not. In terms of our baseline sample (Panel A), and when comparing students who did not enroll in college (Column 3) to students who enrolled 4-year colleges (Column 2), enrollees were more likely to be Asian, White, and female. Enrollees were also less likely to have qualified for free or reduced-price lunch (in 3rd grade), to have received English proficiency services, to have been identified for special education services, or to have a learning disability. Many, though not all, of these differences hold when comparing 2-year college-goers (Column 1) to non-college-goers (Column 3), although the differences are much smaller. And 2-year college-goers were slightly more likely to be Hispanic and have received English proficiency services than students who did not enroll in college.

Turning to test scores, we see that both 2- and 4-year college-goers performed better on their 3rd grade tests than students who did not enroll in college. The test scores of students who enrolled in 2-year college were about one-third of a standard deviation higher than students who did not enrolled in college. Four-year college enrollees' scores were over 80% of a standard deviation higher.

4 METHODOLOGY

To assess the extent to which students' 3rd test scores and background characteristics predict their likelihood of college enrollment (RQ 1), we estimate multinomial logistic regressions. The dependent variable indicates each student's college enrollment outcome—i.e.,

2-year college, 4-year college, not enrolled—within the first year after high school. The multinomial logit model is specified as follows:

$$p_{ij} = \frac{e^{(D_i'\beta D_j + R_i'\beta R_j + G_i'\beta G_j + C_i'\beta C_j)}}{\Sigma_{l=1}^{m} e^{(D_i'\beta D_j + R_i'\beta R_j + G_i'\beta G_j + C_i'\beta C_j)}}, \text{j=1, . . ., m,} \quad (1)$$

where p_{ij} is the ratio of the probability of enrolling in college type *j* relative to the probability of not enrolling in college, D_i is a vector of students' 3rd grade test score deciles, R_i is students' race/ethnicity, G_i is students' gender, C_i is a vector of student classifications that includes students' gender, free- or reduced-price lunch eligibility, disability status, and whether students receive special education and/or English-learner services, and the three different enrollment outcomes are indexed by *j*. We also estimate sparse versions of this baseline model that include only test score deciles.

For RQ 2 we build upon the baseline model in equation (1) by including interactions between test decile and students' race/ethnicity and/or gender (akin to Austin et al., 2023 who measure differential "academic mobility" of racial/ethnic subgroups by assessing the extent to which students' *ranks* in the distributions of various academic measures change during their secondary school careers).

For RQ 3 we use our restricted sample and estimate logistic regressions as in equation (2):

$$log\left(\frac{Enroll4YR_i}{1-Enroll4YR_i}\right) = \alpha_0 + D_i'\beta D_j + R_i'\beta R_j + G_i'\beta G_j + C_i'\beta C_j + \varepsilon_i \quad (2)$$

to predict 4-year college enrollment where D_i represents deciles of achievement for either 3rd grade or high school tests, in separate models. Then, to look grade-over-grade, we estimate equation (2) with a three-way interaction between test deciles, race/ethnicity and grade wherein the relationship between test deciles and college enrollment by race/ethnicity is identified by students who move deciles as they progress from grade to grade.

Finally, for RQ 4, we estimate two additional versions of equation (2). The first version tests the extent to which the predictive power of 3rd grade scores, in terms of 4-year college enrollment, have shifted over time for each race/ethnicity by including a three-way interaction between students' 3rd grade test score deciles, 3rd grade cohort, and race/ethnicity. The second version tests this change over time for students' with and without economic disadvantage by including a three-way interaction between students' 3rd grade test score deciles, 3rd grade cohort and free/reduced-lunch eligibility status.

5 FINDINGS

The coefficient estimates from equation (1) predicting college enrollment are reported in Panel B of **Table 1A** in the appendix as relative log odds (Panel A shows a sparse model with only test deciles as covariates). Because these coefficients are difficult to interpret, we illustrate the primary findings in **Figure 1** as predicted probabilities. Prior to discussing those results, and in the interest answering RQ 1, it is worth noting the extent to which other student background

characteristics are related to college enrollment. Females were more likely than males to enroll in either a 2-year college (β=0.273) or a 4-year college (β=0.506) than not to enroll at all, controlling for other covariates. Students who qualified for free or reduced-price lunch (in 3rd grade) were less likely to attend 2-year college (β=-0.557) or a 4-year college (β=-1.107) than not to enroll at all. This pattern also holds for students with a learning disability, though the effects are smaller in magnitude. Students who received English proficiency services were about as likely to attend either a 2-year (β=0.391) or 4-year college (β=0.357) versus not enrolling. And while students identified for special education services were just as likely to attend a 2-year college as to not enroll in college (an effect with no statistical significance), they were less likely to attend a 4-year college (β=-0.056).

5.1 Predicted Probabilities of College Enrollment By Race/Ethnicity and Gender

Figure 1 illustrates our main findings with respect to (RQ 2) by showing the predicted probabilities of students enrolling in 2- or 4-year college one year after high school across deciles of their 3rd grade test scores, by race/ethnicity. Panel A reports the probabilities for 2-year college enrollment and Panel B for 4-year college enrollment.[10]

Focusing first on Panel A, we see that there is little differentiation across deciles of test scores in the likelihood of students enrolling in 2-year colleges. For instance, when considering all students (the black trend line), the total variation in enrollment probability by 3rd grade test score decile is no more than nine percentage points—ranging from a low of 0.15 at the tenth decile to a high of 0.24 at fourth decile.

Differentials by Race/Ethnicity. When considering variation in enrollment probabilities by race/ethnicity, we see the largest differentials at the ends of the 3rd grade test score distribution. For example, among the lowest-performing students (decile 1), Asians were significantly more likely than students of other races/ethnicities, except Black students (see overlapping confidence intervals), to enroll at a 2-year college—by up to 11.9 percentage points. But among the highest performing students (decile 10), Asians were the least likely to attend a 2-year college (by as much as 10.8 percentage points). Hispanics were most likely to attend a 2-year college but overlapping confidence intervals suggest that their likelihood was not significantly different from that of Black students.

Turning to Panel B, the results for predicted 4-year college enrollment stand in stark contrast to those for 2-year enrollment. We find striking differences across both 3rd grade test deciles and racial/ethnic categories. For instance, when considering all students (the black trend line), students in the lowest decile of achievement had a 6% chance of attending a 4-year college while those in the highest decile had a 48% chance—a difference in variation (42 percentage

[10] Predicted probabilities in **Figure 1** are obtained by estimating equation (1) with a two-way interaction between test score deciles and race/ethnicity. We report partial regression output (main effects only) from this model in Panel C of **Table 1A**. Of the 72 interaction terms jointly estimated, 44 were statistically significant: 23 when predicting 2-year college enrollment and 21 when predicting 4-year college enrollment. Full regression results are available from the authors upon request.

points) more than 4 times larger than the range in probabilities of 2-year college enrollment. For the remainder of this paper, we focus on 4-year college enrollment.

In terms of differences by race/ethnicity, Asians were significantly more likely to attend 4-year college at every decile than students of any other race/ethnicity, with average gaps ranging from 4 to 16 percentage points (in deciles 1 and 9, respectively). Black students had the next highest probabilities of attendance—significantly so, for at least the lower half of the test score distribution, with average gaps between non-Asian students of about 5 percentage points. Differences in enrollment probabilities between non-Asian student groups were larger at the lower end of 3rd grade achievement and narrowed toward the top. Compared to all students (i.e., the black trend line) White students were significantly less likely to attend a 4-year college from the bottom of the test score distribution until the top two deciles, where their likelihood of attendance among other high-achieving non-Asian students is, statistically speaking, indistinguishable.

Notably, regardless of college type, the way enrollment probabilities vary across racial/ethnic groups and test score deciles means that students in a particular racial/ethnic subgroup with lower 3rd grade test scores had a higher chance of enrolling in college than students in a different racial/ethnic group with higher levels of achievement. For example, estimates in Panel B show that Black students who scored in the lower half of the 3rd grade test score distribution were, on average, up to 6 percentage points more likely to attend a 4-year college than White students who scored one decile higher.[11]

Differentials by Gender. To see the extent to which there are gender differences in college enrollment, we estimate a version of equation (1) that includes a three-way interaction between students' test score decile, race/ethnicity and gender.[12] Once again, because these coefficients can be difficult to interpret, we plot predicted probabilities. **Figure 2** shows differentials in predicted probabilities of 4-year college enrollment by gender for all students (Panel A) and for each racial/ethnic group (Panels B-F).[13] The solid trend lines in each panel are the same group-specific lines from Panel B of **Figure 1**. The short-dashed trendlines are the predicted probabilities for females; the long-dashed trendlines are the predicted probabilities for males.

We see that females were significantly more likely than males to attend a 4-year college at every decile of 3rd grade achievement and for each race/ethnicity, with few exceptions.[14] The magnitude of this gender gap varies across the 3rd grade achievement distribution with some

[11] Differences are calculated by comparing upper and lower bounds of confidence intervals between racial/ethnic groups from one test decile to the next.
[12] Partial regression results (main effects only) are reported in Panel D of **Table 1A** in the appendix. Of the 170 jointly estimated interaction effects in this model, 46 were statistically significant. Full regression results are available from the authors upon request.
[13] College enrollment probabilities look similar when subject-specific (i.e., math or ELA) 3rd grade test deciles are used (and these results are available upon request). On average, gender differentials were larger for math, which ranged from 6.9 (Other) to 9.9 (Hispanic) percentage points, compared to ELA, which ranged from 4.9 (Asian) to 8.0 (Hispanic) percentage points.
[14] Overlapping confidence intervals suggest no gender differential in the likelihood of 4-year college enrollment for top-achieving Asian or Black students nor for low-achieving students from races/ethnicities categorized as Other.

narrowing at both ends. For all students across all deciles the average gender gap is 7 percentage points (refer to Panel A). The gender gap varies by race/ethnicity—from 5.7 (for Others) up to 8.2 and 8.7 percentage points for Black and Hispanic students, respectively.[15] And, notably, for Black, Hispanic and White students it is often the case that females who scored one decile *lower* on their 3rd grade tests were *more* likely to attend a 4-year college than their higher-achieving (by one decile) male counterparts. Estimates in Panel C of **Figure 2** show, for example, that Black females at the 3rd decile of 3rd grade achievement were up to 10 percentage points more likely to attend a 4-year college than Black males at the 4th decile of 3rd grade achievement.

5.2 CHANGES IN PROPENSITY FOR COLLEGE ENROLLMENT ACROSS GRADES

When considering students' performance and progression through school, we should note that the proportion of students in each achievement decile varies across race/ethnicity and grade. **Figure 3** shows the proportion of students across deciles of the 3rd grade (left side) and high school (right side) test score distributions for each race/ethnicity. Proportions of students in the same achievement deciles vary widely across race/ethnicity. For example, 5.1% of Asian students versus nearly 19% of both Black and Hispanic students scored in the first decile on their 3rd grade test. Indeed, compared to Asians, the distributions of Black and Hispanic students are roughly inverted. For any racial/ethnic group, there is little variability within deciles from 3rd grade to high school except for the notable shift of Asian students into the top decile—from 17.2% in 3rd grade to 24.5% in high school.

Having established in section 5.1 that 3rd grade tests strongly predict 4-year college enrollment, we turn to assessing whether there are points (i.e., grades) in a student's academic career in which students' performance on tests appear to be particularly important for those predictions (RQ 3).

We begin by simply comparing the predictive power of the earliest and latest years of test scores available (i.e., 3rd grade and 10 or 11th grade) on 4-year college enrollment. For comparability across the two grades, we restrict the analytic sample to 3rd graders for whom we also observe high school tests scores (n=326,943) as described in Column 2 of **Table 2**. Using this restricted sample, we first estimate equation (2) in which we predict 4-year college enrollment as a function of 3rd grade test score deciles and student characteristics. We then re-estimate the model using deciles of *high school* test scores instead of 3rd grade scores keeping all other student covariates. Appendix **Table 2A** shows these results. We find that 10th grade scores were more predictive (pseudo R^2= 0.22) than 3rd grade scores (pseudo R^2=0.16) as would be expected given their proximity to college enrollment. When we estimate models for each of the racial/ethnic subgroups separately (see **Appendix Table 3A**), *differences* in the values of pseudo R^2 are not large across subgroups—ranging from 0.049 (Other) to 0.064 (Hispanic). Despite

[15] At places along the 3rd grade achievement distribution where the group-specific gender gap within a race/ethnicity is notably wider than the average gender gap for all students (as it is along much of the 3rd grade achievement distribution for Black and Hispanic students), we see that the wideness is because the group-specific gender differentials are driven by one or both genders' higher or lower likelihoods of attending a 4-year college compared to the average female or male (see **Figure 1A** in the appendix).

these differences, we see that students are, at least to some degree, anchored to their educational achievement levels in elementary school.

Next, we look to see whether there are points in a student's academic career between elementary and high school where their academic performance on grade-level tests is particularly important for predicting college enrollment. To do this we estimate a version of the logistic regression model shown in equation (2) that includes a three-way interaction between test score deciles, race/ethnicity and grade. This model is estimated separately for each racial/ethnic subgroup such that students of a particular racial/ethnic subgroup are compared to all other students (e.g., Asians versus non-Asian).

Figure 4 shows the predicted probabilities from these models by racial/ethnic group across grades for three selected achievement deciles: bottom (1st) decile (short-long-dashed trendlines and circle markers), middle (5th) decile (long-dashed trendlines and diamond markers), and top (10th) decile (solid trendlines and square markers). In each panel, the colored trend lines are the predicted probabilities for students in that particular racial/ethnic group; gray trend lines are the predicted probabilities for students of all other races/ethnicities.[16] So, the solid purple line in Panel A reflects enrollment probabilities for top-performing Asian students and the solid gray line reflects enrollment probabilities for top-performing *non*-Asian students. While we show results for students at various performance deciles across grades, for the remainder of this section we are mainly interested in the top-performing students as they are the most likely to attend a 4-year college.

Given the higher predictive power of high school tests relative to third grade tests—and the increasing proximity to the decision to actually enroll in college—it comes as no surprise that the probabilities of college enrollment steadily increased for top-performing students as they progressed from 3rd grade to high school (note the upward trajectories of the top two trendlines across panels). Except, surprisingly, for Hispanic students whose probabilities of 4-year college enrollment were flat from 7th grade on. In contrast, low-performing students, were consistently very unlikely to attend a 4-year college across all races/ethnicities. Indeed, the growing gap between high- and low-performing students of the same race/ethnicity (compare the colored solid lines to the colored short-long-dashed lines within each panel) is due almost entirely to the increased likelihood of high-achieving students enrolling in college as they progressed through grades since the enrollment probabilities for low-achieving students were relatively flat across grades.[17]

Also note that differentials in the likelihood of college enrollment across grades varies by race even within the same test score decile. Consider high-achievers, shown in the top two trend lines in each subplot. Asian students (solid purple line) were more likely to enroll in college than non-Asian students (solid gray line) in every grade—a difference that is fairly consistent across grades. In contrast, White students (solid red line) were less likely to enroll in college than non-

[16] For the sake parsimony, we do not show students categorized as an "Other" race/ethnicity as their trends are similar to those of Hispanic students.

[17] Asian students were the exception to this pattern in that the likelihood of enrolling in college decreased significantly for low-achieving Asian students in later grades.

White students (solid gray line) with larger differences in later grades. For the most part, high-achieving Black (solid orange line) and Hispanic students (solid blue line) were about as likely to enroll in college as were students of any other race/ethnicity across grades.

Aside from such general trends, this figure can show in what grades and at which points in the test score distribution one racial/ethnic group excels or struggles relative to another (at least for the three achievement deciles displayed). But looking across grades, we find there are not many sudden, statistically significant divergences (or convergences) between racial/ethnic groups where enrollment gaps seem to be widening or narrowing. There are a few exceptions. For example, the top two (solid) trend lines in Panel B shows that among top-performers, Black students (solid orange line) were significantly more likely than students from all other races/ethnicities (solid gray line) to attend a 4-year college only by the time they were in high school, when confidence intervals from the two groups no longer overlap. Likewise, the top two (solid) trendlines in Panel C show that top-performing Hispanic students (solid blue line) in elementary school were less likely to attend a 4-year college than non-Hispanic students (solid gray trendline)—a gap that closed in middle school but reappeared in high school as the likelihood for non-Hispanics to enroll in a 4-year college increased and the likelihood of Hispanic enrollment remained unchanged from 8th grade.

Notably, of all the 216 three-way interactions estimated across these models only 39 (18%) were statistically significant, suggesting that regardless of race/ethnicity or test achievement, the racial/ethnic differences in 4-year college enrollment probabilities are relatively stable across grades—with the exception of top-performing students.

5.3 Changes in Propensity for College Enrollment Across cohorts

In this section we consider whether 3rd grade test performance differentially influenced college enrollment across cohorts (RQ 4). This analysis examines whether state efforts to aid historically disadvantaged students are reflected in changing college trajectories across the cohorts in our sample. Accordingly, we analyze college-going patterns across cohorts by race/ethnicity and economic disadvantage (i.e., eligibility for free or reduced-price lunch).

To test whether 3rd grade achievement differentially predicts enrollment probabilities across cohorts and by race/ethnicity, we estimate a version of equation (2) that includes three-way interactions between students' 3rd grade test score decile, 3rd grade cohort and race/ethnicity.[18] From this model we derive the predicted probabilities (shown in **Figure 5)** of 4-year college enrollment across cohorts and by race/ethnicity for students in the bottom (circle markers) , middle (diamond markers) and top (square markers) deciles of 3rd grade achievement. In general, the overlapping confidence intervals within deciles across cohorts for each racial/ethnic group show little evidence of change over time in enrollment probabilities regardless of students' race/ethnicity or achievement level. Indeed, there are only a few instances where we can detect a

[18] Of the 216 three-way interaction terms estimated in the model, 43 (20%) were statistically significant. Regression results are available from the authors upon request.

statistically significant increase across cohorts—all of which occurred among the lowest-performing students (i.e., in decile 1). For example, compare the predicted enrollment probabilities of low-achieving Black students (Panel B) from the 2006 cohort (0.06) to those in the 2012 cohort (0.14). Because the value of the upper bound of the confidence interval for the 2006 cohort (0.08) is less than the value of lower bound of the confidence interval for the 2012 cohort (0.10), we can infer that 3rd grade achievement increased in predictive power for the lowest-achieving Black students over those years. Taking into account the variability in predicted probabilities captured by the confidence intervals, this increase could have ranged between 2 and 12.5 percentage points. This pattern holds when comparing the earliest cohort to later cohorts (2009 and on) of Hispanic and White students, but with smaller predicted increases—ranging from less than 1 up to 5 percentage points.

Finally, to test the extent to which enrollment probabilities change over time for students with different levels of economic disadvantage, we estimate a version of equation (2) that includes three-way interactions between students' 3rd grade test scores, 3rd grade cohort and free/reduced-price lunch eligibility (in 3rd grade).[19] **Figure 6** shows the model results in the form of predicted probabilities for 4-year college enrollment across cohorts and selected deciles (i.e., 1, 5 and 10) for students who were eligible for free/reduced-price lunch (Panel A) and those who were not (Panel B). Again, in general, overlapping confidence intervals within each performance decile suggests little increase in the probability of 4-year college enrollment across cohorts. The only sustained increases were for the lowest performers (decile 1) in either group, yet the gains are arguably small. For example, estimates in Panel A show that compared to the 2006 cohort, low-performing economically-disadvantaged students in the four most recent cohorts (i.e., 2009-2012) were at most 3 percentage points more likely to attend a 4-year college. Interestingly, enrollment probabilities increased slightly more (though still less than 6 percentage points) for low-achieving students who were *not* economically disadvantaged (see bottom trendline in Panel B).

Despite recent state policies aimed at boosting enrollment of disadvantaged groups, our analysis provides little evidence that the likelihood of 4-year college attendance increased during the studied period—at least when measured by the predictive power of 3rd grade test scores.

6 DISCUSSION AND CONCLUSION

Research on school attainment has long established that long-term trajectories are crucial for understanding educational outcomes (Oakes 1985; Stevenson, Schiller, and Schneider 1994). Policies aimed at improving post-secondary outcomes and attainment typically target secondary school. While secondary-focused policies show some promising results (Holzman et al., 2020; Klasik, 2012), students have already experienced countless developmental milestones and experiences before reaching middle and high school. Few studies have considered whether students' earlier academic achievements *prior* to secondary school influence postsecondary

[19] Of the 54 three-way interaction terms estimated, 6 (11%) were statistically significant. Regression results are available from the authors upon request.

outcomes. Our research addresses this gap by analyzing the relationship between students' 3rd grade test performance and their college enrollment patterns.

Using statewide censuses of students over multiple cohorts, we find that students' 3rd grade test scores strongly predicted their likelihood of enrolling in a 4-year college, with significant variation across performance deciles and notable differences by race/ethnicity and gender. In examining the large disparities we observed between racial/ethnic groups (see Table 4), we find that the lower likelihood of Black and Hispanic students for attending college is not because students with similar 3rd grade test scores had lower outcomes based on their race/ethnicity; indeed, at many deciles of test performance Black and Hispanic students were *more* likely to attend a 4-year college than their comparably-scoring White peers (see Figure 1). Rather, the disparities exist because there are large, systematic racial/ethnic differences in how students scored on the tests. These test score gaps existed in grade 3 and persisted into high school (see Figure 3).

Unsurprisingly, our analysis shows that high school tests, which occur closer to college enrollment, are better predictors of whether a student will attend a 4-year college. But our results also show that you can draw a very similar conclusion based on 3rd grade achievement. Whether the difference in predictive power between 3rd grade and high school tests is large or small is in the eye of the beholder. But it seems reasonable to say that 3rd grade tests do a fairly good job of capturing the relationship between student test performance and college-going. Lastly, although programs designed to help students who typically struggle in school were implemented and expanded during the years we examine, we find little evidence that later cohorts' likelihoods of 4-year college enrollment increased relative to earlier cohorts. This study reinforces two related ideas: (1) a student's academic performance at any given point strongly predicts their college enrollment, and (2) students largely maintain the educational trajectories predicted by their 3rd grade tests. This finding about persistence reflects similar findings that examine academic trajectories and high school outcomes (Goldhaber et al., 2024). On balance, the persistence of trajectories established by 3rd grade suggests that interventions aimed at middle or high school may, for many students, be too late.

Our findings have implications for educational policy at multiple levels. For schools, they underscore the importance of early assessments that can identify struggling students in the primary grades, coupled with immediate, targeted interventions. For school districts, they underscore the importance of elementary education, particularly in high-need areas, and developing continuous support systems that follow students as they progress through grade levels. At the state level, the results support calls to direct additional resources toward early childhood and elementary education in communities where achievement gaps are most pronounced. For research, the results highlight the importance of identifying interventions in elementary education that most effectively alter these predictable trajectories, particularly for historically disadvantaged students (e.g., Xu et al. 2025). Additionally, longitudinal studies that track the impact of sustained early interventions through college enrollment would provide leaders with useful information about which policies might yield more substantial and lasting improvement in educational access and equity.

REFERENCES

Alexander, K. Entwisle, D., & Dauber, S. 2003. *On The Success of Failure: A Reassessment of the Effects of Retention in the Primary Grades*. Cambridge: Cambridge University Press. 2nd Ed. NewYork:Cambridge University Press.

Austin, W., Figlio, D., Goldhaber, D., Hanushek, E., Kilbride, T., Koedel, C., Lee, J. S., Luo, J., Özek, U., Parsons, E., Rivkin S., Sass, T., & Strunk, K. O. (2023). "Academic Mobility in U.S. Public Schools: Evidence from Nearly 3 Million Students." CALDER Working Paper No. 227-0323-3. *National Center for Analysis of Longitudinal Data in Education Research (CALDER).*

Bloom, B.B. (1964.) *Stability and Change in Human Characteristics*. New York: John Wiley.

Bobek, B. & San Pedro, S. (2021). "Planning for the Future: Early Academic Achievement, Educational Aspirations, and Interest-Career Similarity Important for College Enrollment." ACT Research. Issue Brief. https://eric.ed.gov/?id=ED616524

Bowers, A. (2019). "Towards Measures of Different and Useful Aspects of Schooling: Why Schools Need Both Teacher-Assigned Grades and Standardized Assessments." In Brookhart, S. & McMillin, J. (Eds.), *Classroom Assessment and Educational Measurement* (pp. 209-223). Routledge.

Center for Education Policy Research. (2024). "Examining Student Transitions Into College and Career Pathways in Washington: Strategic Data Project Education to Workforce Pathways Diagnostic Research Report." https://sdp.cepr.harvard.edu/sites/hwpi.harvard.edu/files/sdp/files/p20w_washington_state_report_final_for_web_01.pdf

Chetty, R., Friedman, J., Hilger, N., Saez, E., Whitmore-Schanzenbach, D., and Yagan, D. (2011). "How Does Your Kindergarten Class- room Affect Your Earnings? Evidence from Project STAR," *Quarterly Journal of Economics*. 126, 1593–1660.

Dauber, S., Alexander, K., & Entwisle, D. (1996). "Tracking and Transitions through the Middle Grades: Channeling Educational Trajectories." Sociology of Education, 69(4): 290-307.

Dougherty, C., Kiley, J., & Asheer, N. (2024). "Changing Coalitions in a Politically Divided Nation." Pew Research Center. https://www.pewresearch.org/politics/2024/04/09/partisanship-by-race-ethnicity-and-education/

Dynarski, S., Hemelt, S., and Hyman, J. (2015). The Missing Manual: Using National Student Clearinghouse Data to Track Postsecondary Outcomes. *Educational Evaluation and Policy Analysis*, 27(1S):54S-79S.

Entwisle, D., Alexander, K. & Olson, L. (2005). First Grade and Educational Attainment by Age 22: A New Story. *American Journal of Sociology*, 110(5): 1458-1502.

Giani, M., Attewell, P., & Walling, D. (2020). "The Value of an Incomplete Degree: Hetergeneity in the Labor Market Benefits of College Non-Completion." *The Journal of Higher Education*, 91(4): 514-539.

Glick, J. E. & White, M. J. (2004). "Post-Scondary School Participation of Immigrant and Native Youth: The Role of Familial Resources and Educational Experiences." *Social Science Research*, 33:272-299.

Goldhaber, D., Koedel, C., Özek, U., and Parsons, E. (2022). Using Longitudinal Mobility to Identify At-Risk Students. *AERA Open*, 8(1):1-13.

Goldhaber, D., Baird, K., & Gautam, S. (2024). Progression and Predictors of Public-School Student Outcomes in Washington State. CEDR Working Paper No. 05222024-1. University of Washington, Seattle, WA.

Goodman, J., Hurwitz, M., Smith, J., & Fox, J. (2015). "The Relationship Between Siblings' College Choices: Evidence From One Million SAT-Taking Families." *Economics of Education Review*, 48:75-85.

Heckman, J., Humphries, J., & Veramendi, G. (2017). The Non-Market Benefits of Education and Ability." NBER Working Paper Series. http://www.nber.org/papers/w23896

Hein, V., Smerdon, B., & Sambolt, M. (2013). "Predictors of Postsecondary Success." College and Career Readiness & Success Center at American Institutes for Research. https://eric.ed.gov/?id=ED555671

Hershbein, B., McMullen, I., Pittelko, B., & Timmeney, B. (2021). Beyond Degrees: Longer Term Outcomes of the Kalamazoo Promise. Upjohn Institute Working Paper 21-350.

Holzman, B. & Duffy, H. (2024). "Empowering Educational Leaders: On-Track Indicators for College Enrollment." (EdWorkingPaper: 24-960). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/styt-2294.

Holzman, B., Klasik, D. & Baker, R. (2020). "Gaps in the College Application Gauntlet." *Research in Higher Education*, 61:795-822.

Hout, M. 2012. "Social and Economic Returns to College Education in the United States." *Annual Review of Sociology*, 28:279-400.

Kremer, K. (2022). "Predictors of College Success Outcomes in Emerging Adults: The Role of High School Dual Enrollment Courses." *Emerging Adulthood,* 10(1): 188-196.

Klasik, D. (2012). The College Application Gauntlet: A Systematic Analysis of the Steps to Four-Year College Enrollment." *Research in Higher Education*, 53:506-549.

Lochner, L. (2011). Non-production Benefits of Education: Crime, Health, and Good Citizenship." NBER Working Paper Series. http://www.nber.org/papers/w16722

Lovenheim, M. F., & Reynolds, C. L. (2011). "Changes in Postsecondary Choices by Ability and Income: Evidence from the National Longitudinal Surveys of Youth." *Journal of Human Capital*, 5(1): 70-109.

McMahon, W. W. (2009). *Higher learning, greater good: The private and social benefits of higher education*. Johns Hopkins University Press.

Mello, Z. (2008). “Gender Variation in Developmental Trajectories of Educational and Occupational Expectations and Attainment from Adolescence to Adulthood.” *Developmental Psychology*, 44(4) 1069-1080.

Morris, S. & McKenzie, S. (2024). “The Role of High School Freshman Grades, Socioeconomic Status, and School Location on College Enrollment.” *Journal of School Administration Research and Development*, 9(1): 1-11.

Oakes, J. 1985. *Keeping Track: How Schools Structure Inequality*. New Haven, Conn.: Yale University Press.

Reber, S., Rünger, D., & Wong, M. (2024). “The Effects of Charter High Schools on Academic Acheivement and College Enrollment: Evidence from Los Angeles.” *Education Finance and Policy*, 19(4): 567-585.

Schaller, T. K., Routon, P. W., Partridge, M. A., & Berry, R. (2023). “A Systematic Review and Meta-Analysis of Dual Enrollment Research.” *Journal of College Student Retention: Research, Theory & Practice*.

Stevenson, D., Schiller, K., & Schneider, B. (1996). “Sequences of Opportunities for Learning.” *Sociology of Education*, 67(3): 184-98.

United States Census Bureau. (2022). “Educational Attainment in the United States: 2021” https://www.census.gov/newsroom/press-releases/2022/educational-attainment.html

Xu, Z., Özek, U., Levin, J., & Hoon Lee, D. (2025). Effects of Large-Scale Early Math Interventions on Student Outcomes: Evidence From Kentucky’s Math Achievement Fund. Educational Evaluation and Policy Analysis, 0(0). https://doi-org.air.idm.oclc.org/10.3102/01623737241303142

TABLE AND FIGURES

Table 1. Summary of Data Holdings

School Year	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21	2021-22
	CSRS					**CEDARS**												
Demographics and Student Classifications	Race/ethnicity, gender, free- or reduced-lunch eligibility, special education status, limited-English proficiency																	
Test Scores		Math and ELA test scores, standardized by year and grade														No test scores due to COVID-19		
	PCHEES, SBCTC, NSC																	
Postsecondary Outcomes									College enrollment and completion Four-year or two-year; public or private; in-state or out-of-state									

Table 2. Third Grade Cohorts Followed Until One Year After High School

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

Table 3. Summary Statistics of Student Characteristics for Baseline and Restricted Samples

	(1) Baseline 3rd Grade Sample	(2) Restricted High School Sample
Demographics		
Asian (%)	7.5	7.5
Black (%)	4.9	4.2
Hispanic (%)	18.6	19.9
Other (%)	7.9	8.1
White (%)	61.2	60.4
Female (%)	48.9	49.0
Classifications (in 3rd Grade)		
Free/Reduced-Price Lunch (%)	45.3	45.3
English Proficiency Services (%)	9.6	9.9
Special Education Services (%)	11.0	10.7
Learning Disability (%)	3.9	3.6
3rd Grade Test Performance		
Math Score (mean, sd)	0 (1.0)	0.02 (.99)
ELA Score (mean, sd)	0 (1.0)	0.02 (.99)
Average Score (mean, sd)	0 (1.0)	0.02 (.99)
N	507,101	326,943

Table 4. College Enrollment Outcomes, By Race/Ethnicity and Sample

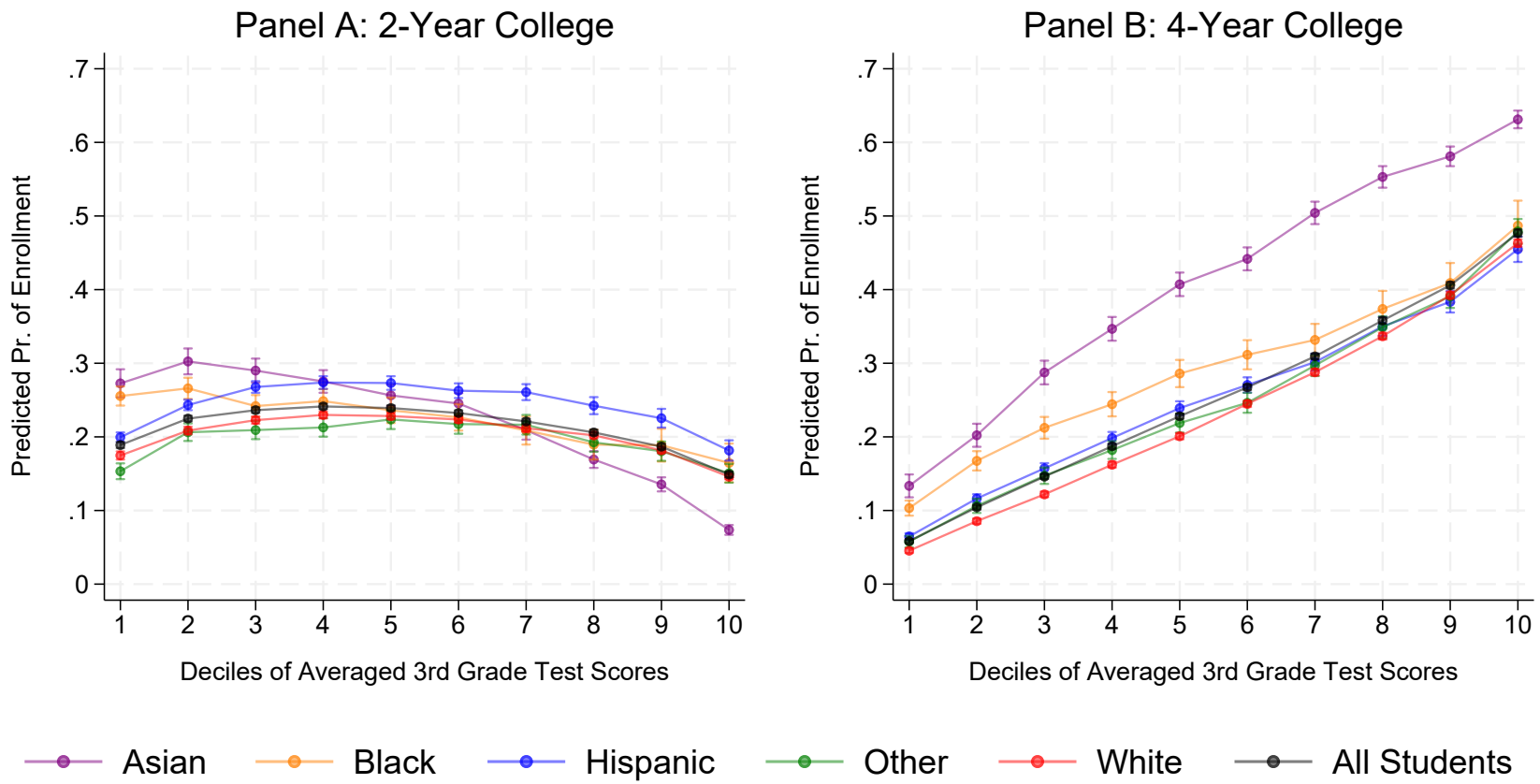
Panel A: Baseline 3rd Grade Sample					
	Asian	Black	Hispanic	Other	White
2-year College	20.1	22.2	24.0	18.6	20.0
4-year College	48.6	20.7	16.1	22.7	28.0
No College	31.3	57.1	60.0	58.7	52.1
N	37,793	24,593	94,364	39,903	310,448
Panel B: Restricted High School Sample					
	Asian	Black	Hispanic	Other	White
2-year College	21.9	26.7	26.7	21.1	21.9
4-year College	56.0	27.6	18.8	27.7	32.6
No College	22.1	45.7	54.5	51.2	45.5
N	24,662	13,681	64,942	26,322	197,336

Table 5. Student Characteristics by College Outcome and Sample

	Panel A: Baseline 3rd Grade Sample			**Panel B: Restricted High School Sample**		
	(1)	(2)	(3)	(4)	(5)	(6)
Enrollment Outcome	2-year	4-year	Not Enrolled	2-year	4-year	Not Enrolled
Demographics						
Asian (%)	7.2***	13.6***	4.4	7.2***	13.6***	3.6
Black (%)	5.2	3.8***	5.3	4.9***	3.7***	4.1
Hispanic (%)	21.6**	11.3***	21.1	23.1*	12.0***	23.5
Other (%)	7.1***	6.7***	8.8	7.4***	7.2***	9.0
White (%)	58.9***	64.6***	60.4	57.5***	63.4***	59.7
Female (%)	50.7***	57.2***	44.1	50.8***	57.3***	42.4
Classifications (in 3rd Grade)						
Free/Reduced-Price Lunch (%)	44.3***	23.7***	56.7	45.1***	24.9***	58.8
English Proficiency Services (%)	12.1***	4.8***	11.0	12.5***	4.9***	11.9
Special Education Services (%)	10.5***	6.1***	13.7	10.4***	6.1***	13.9
Learning Disability (%)	3.4***	1.1***	5.4	3.3***	1.1***	5.5
3rd Grade Test Performance						
Math Score (mean, sd)	-0.07 (.90)***	0.54 (.90)***	-0.25 (.98)	-0.08 (.90)***	0.52 (.92)***	-0.27 (.95)
ELA Score (mean, sd)	-0.06 (.91)***	0.51 (.87)***	-0.24 (.99)	-0.06 (.91)***	0.51 (.87)***	-0.28 (.97)
Average Score (mean, sd)	-0.07 (.89)***	0.57 (.85)***	-0.26 (.99)	-0.07 (.89)***	0.56 (.86)***	-0.30 (.96)
N	104,882	134,623	267,596	75,206	101,391	150,346

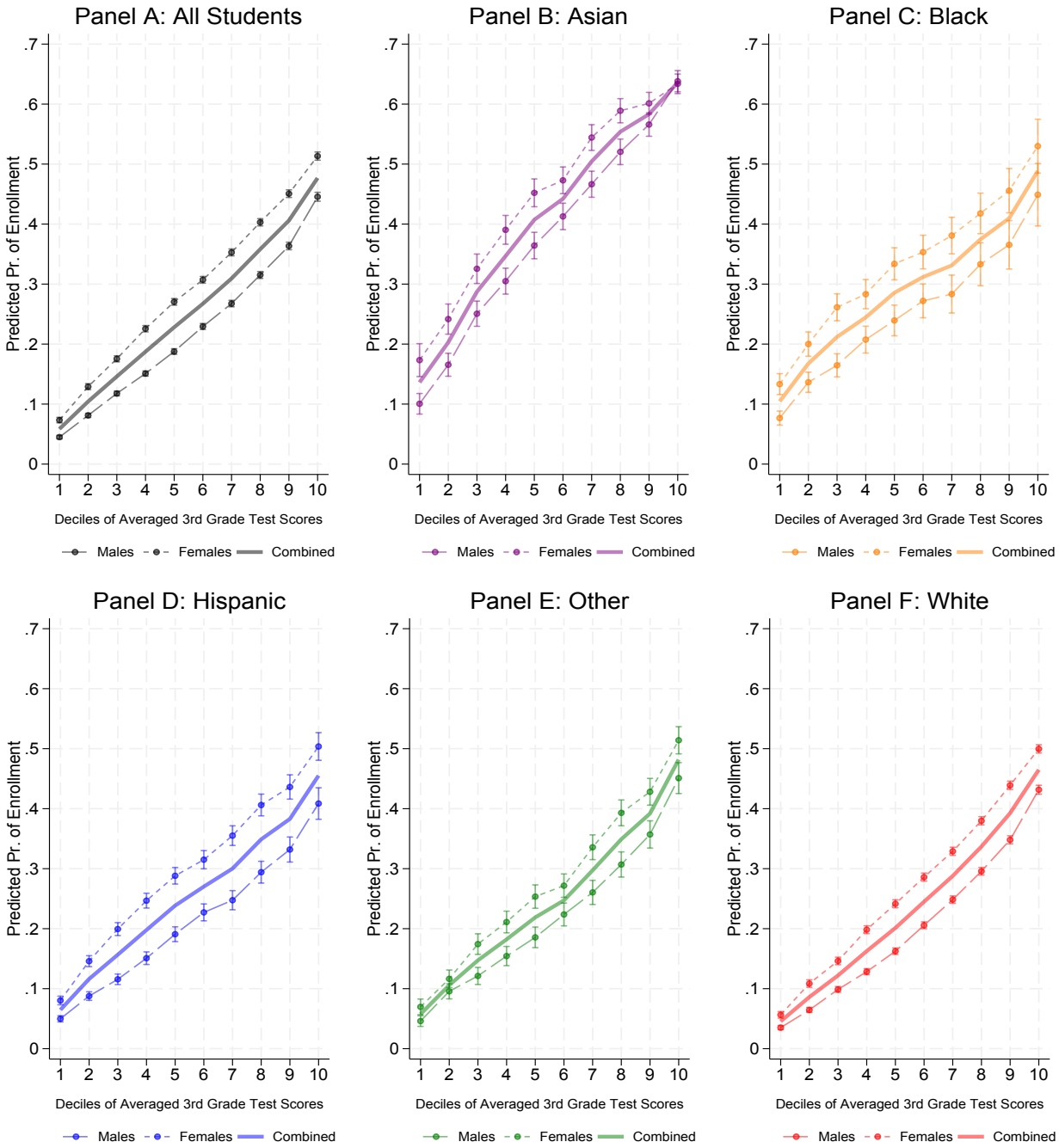
Note. The t-tests in columns 1-2 & 5-6 test the difference in each variable between students in these sample-specific college-going categories and students in the same sample who were not enrolled in college, i.e., Columns 1 and 2 are compared to Column 3 and Columns 4 and 5 are compared to Column 6.

Figure 1. Predicted Probabilities of College Enrollment One Year After High School Across Third Grade Achievement Deciles, By Race/Ethnicity



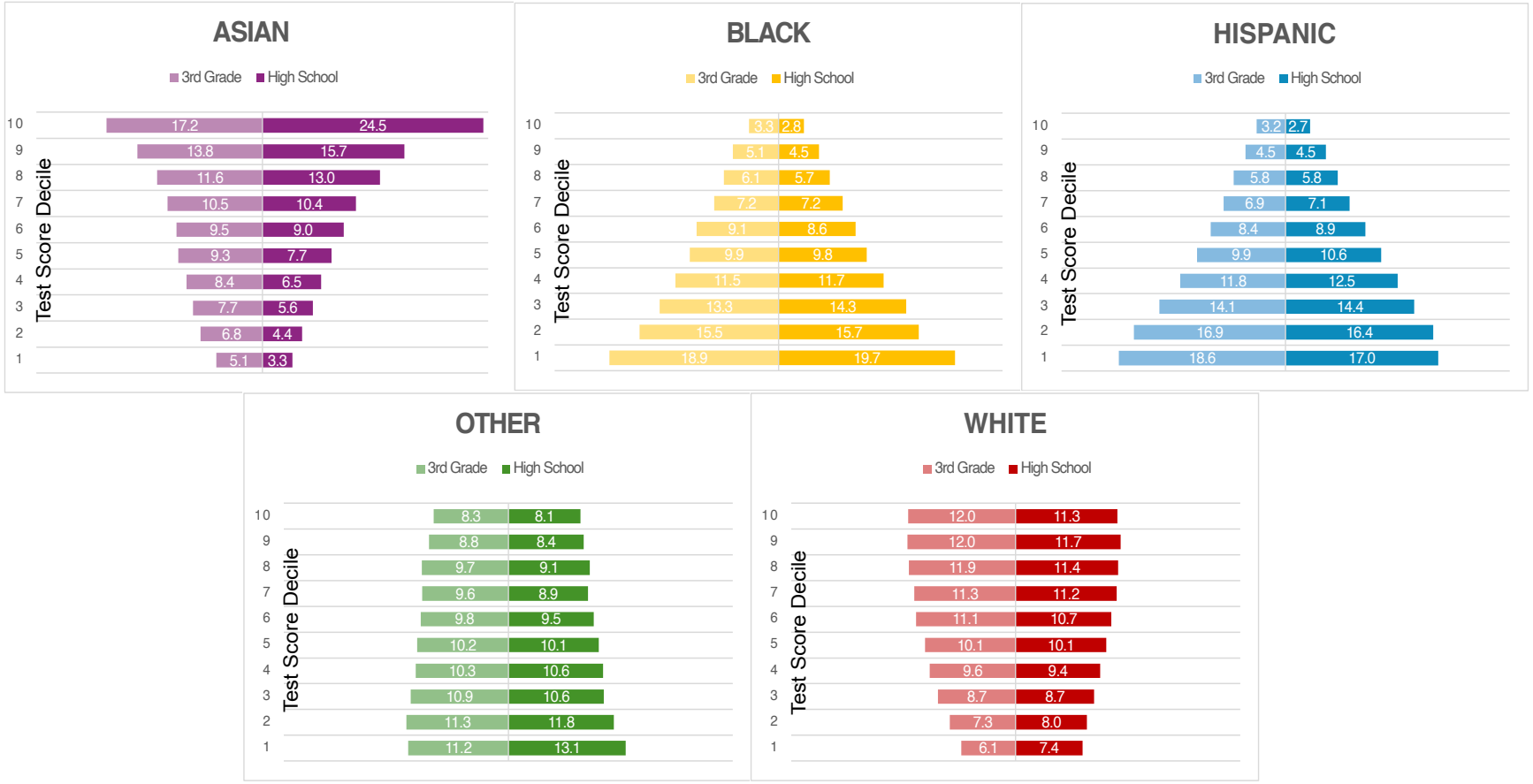
Notes. The predicted probabilities and 95% confidence intervals in this figure are generated based on estimates from a multinomial logistic regression as in equation (1) with two-way interactions between 3rd grade test deciles and race/ethnicity. Predicted probabilities of 2-year (Panel A) and 4-year (Panel B) college enrollment are shown for five racial/ethnic subgroups (Asian=purple, Black=orange, Hispanic=blue, Other=green, White=red) and for all students (black). Relative log odds for the main effects are reported in Column 5 (2-year colleges) and Column 6 (4-year colleges) of Panel C in Table 1A of the appendix.

Figure 2. Predicted Probabilities of 4-Year College Enrollment One Year After High School Across Third Grade Achievement Deciles, By Race/Ethnicity and Gender



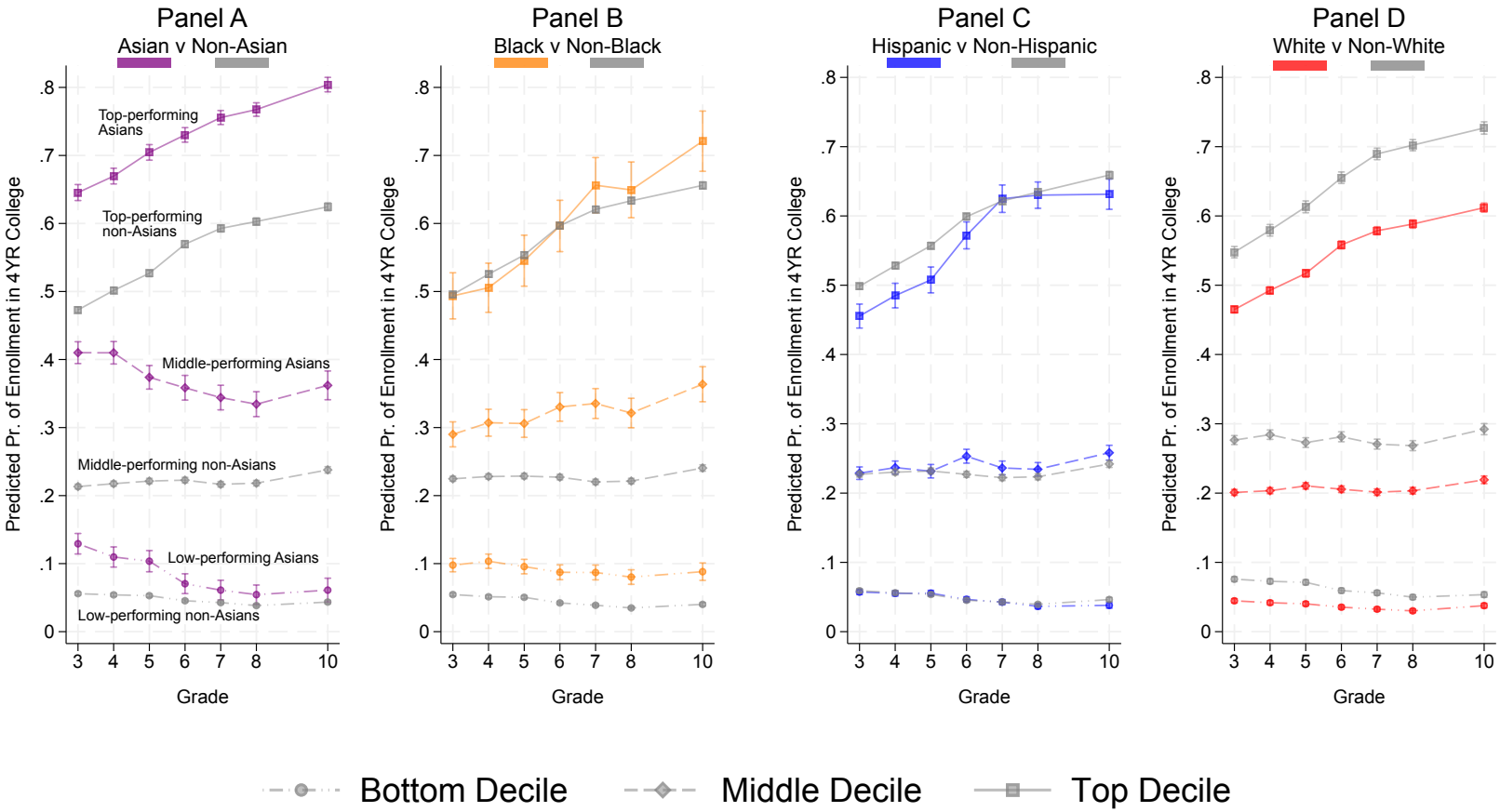
Notes. The predicted probabilities and 95% confidence intervals in this figure are generated based on estimates from a multinomial logistic regression as in equation (1) with three-way interactions between 3rd grade test deciles, race/ethnicity and gender. Within each panel, the solid line is the average likelihood for each racial-ethnic group and is the same line shown in Panel B of Figure 1. The short-dashed lines show predicted probabilities for females; the long-dashed lines show predicted probabilities for males. The average gender gap for all students is 7 percentage points and varies by racial/ethnic groups as follows: Asian (6.3), Black (8.2), Hispanic (8.7), Other (5.7), White (6.7). The relative log odds for the main effects are reported in Column 7 (2-year colleges) and Column 8 (4-year colleges) of Panel D in Table 1A of the appendix.

Figure 3. Proportion of Students Across 3rd Grade and 10th Grade Test Scores Deciles, By Race/Ethnicity



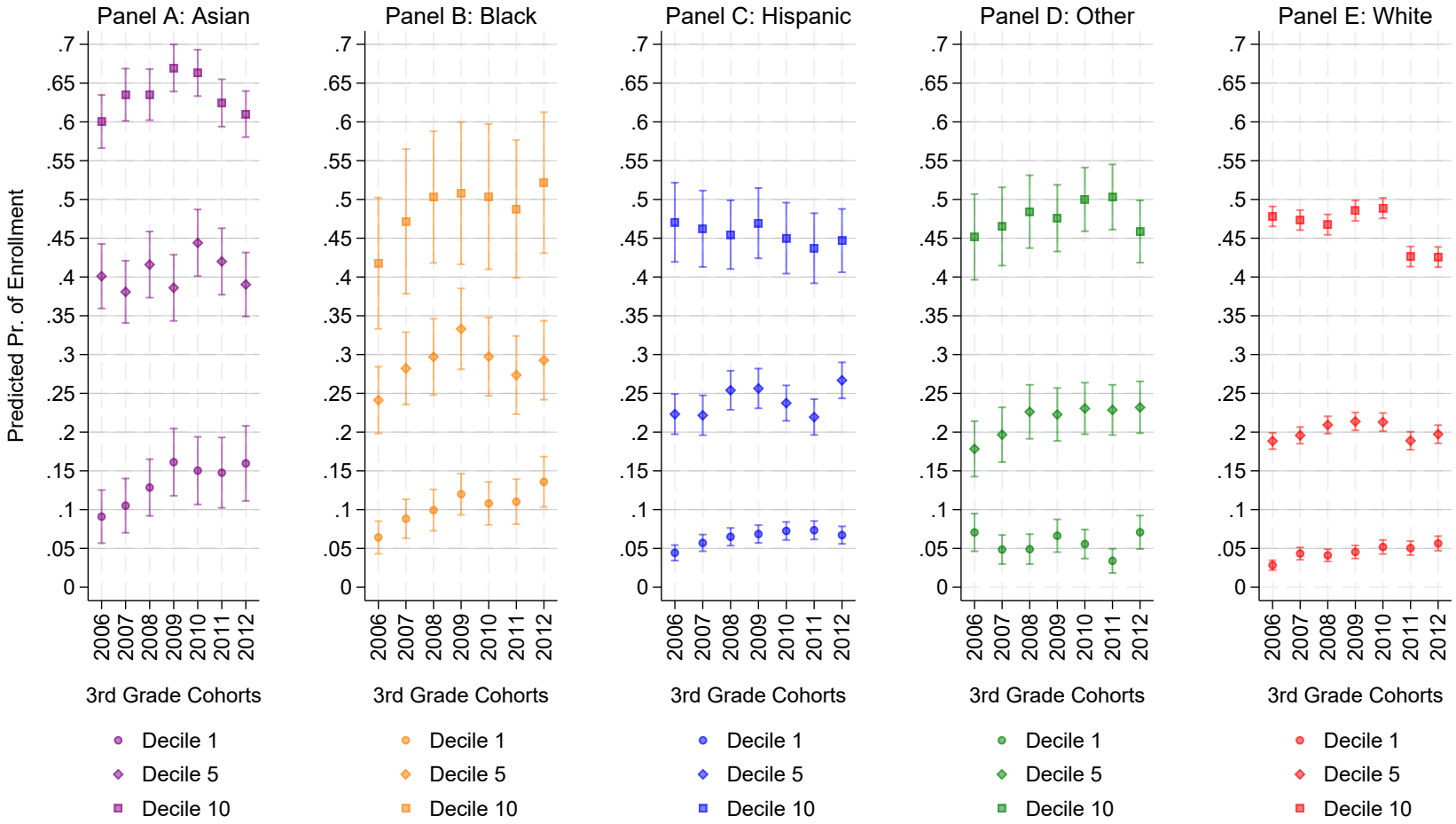
Notes. This figure shows the proportion of students within each race/ethnicity who scored at each the deciles of the test score distribution in 3rd grade (lighter, left-hand-side bars) and 10th grade (darker, right-hand-side bars).

Figure 4. Predicted Probabilities of Enrolling in 4-Year College Across Grades and by Deciles of Achievement



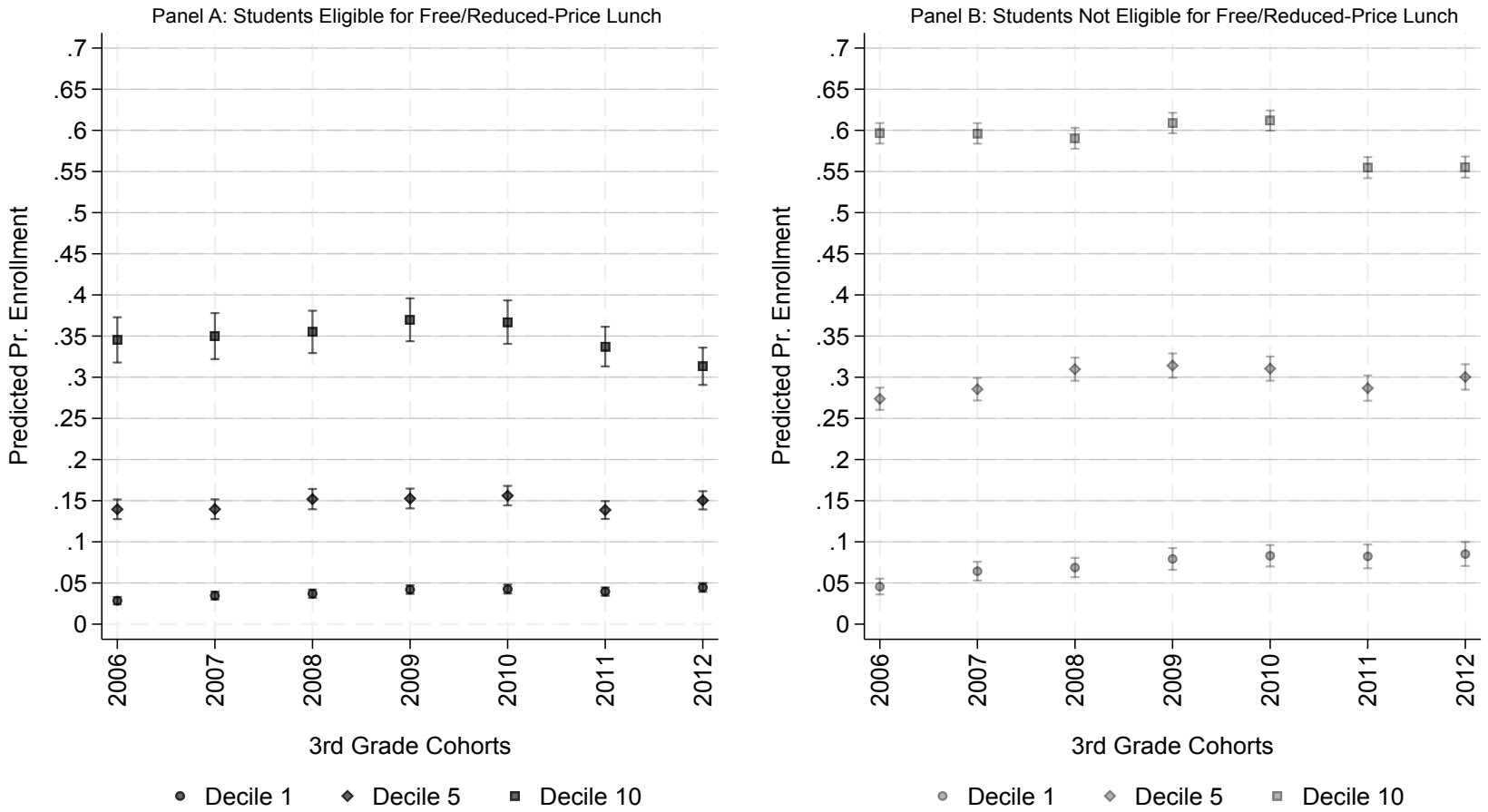
Notes. The predicted probabilities and 95% confidence intervals in this figure are generated based on estimates from separate logistic regressions for each racial/ethnic group as shown in equation (2) with three-way interactions between test score deciles, race/ethnicity and grade. We show results for three selected achievement deciles: bottom (decile 1) with short-long-dashed trendlines and circle markers, middle (decile 5) with long-dashed trendlines and diamond markers, and top (decile 10) with solid trendlines and square markers. In each panel, the colored trend lines are the predicted probabilities for students in that racial/ethnic group; gray trend lines are the predicted probabilities for students of all other races/ethnicities. So, the solid purple line in Panel A shows enrollment probabilities for top-performing Asian students and the solid gray line shows enrollment probabilities for top-performing *non*-Asian students.

Figure 5. Predicted Probabilities of 4-Year College Enrollment for Low-Achieving Students Across Cohorts, By Race/Ethnicity



Notes. The predicted probabilities and 95% confidence intervals in this figure are generated based on estimates from a logistic regression as in equation (2) with three-way interactions between students' 3rd grade test score decile, 3rd grade cohort and race/ethnicity. Within each panel, we show results for three selected achievement deciles: decile 1 (circle markers), decile 5 (diamond markers), and decile 10 (square markers).

Figure 6. Predicted Probabilities of 4-Year College Enrollment for Low-Achieving Students Across Cohorts, By Free/Reduced Price Lunch Status



Notes. The predicted probabilities in this figure are generated based on estimates from a logistic regression as in equation (2) with three-way interactions between students' 3rd grade test score decile, 3rd grade cohort and free/reduced-price lunch eligibility. Results for students who were eligible for free/reduced-price lunch are shown in Panel A; results for students not eligible are shown in Panel B. Within each panel, we show results for three selected achievement deciles: decile 1 (circle markers), decile 5 (diamond markers), and decile 10 (square markers).

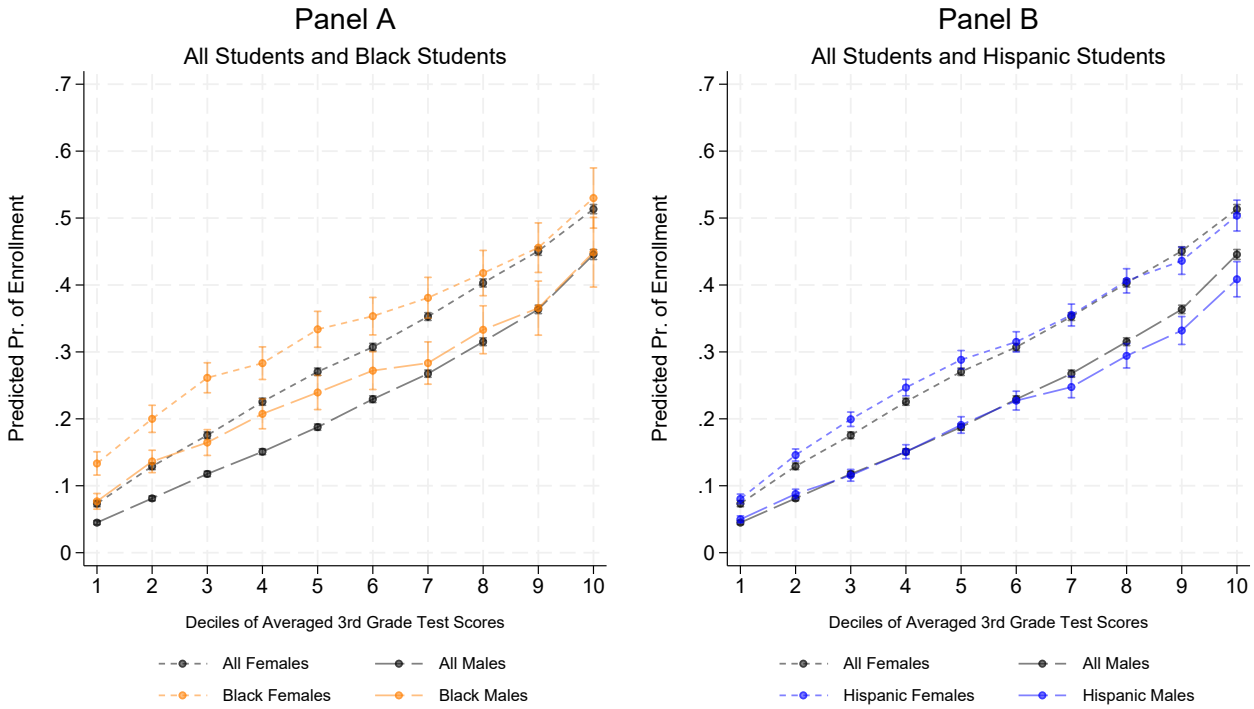
APPENDIX

Table 1A. Relative Log Odds of Enrolling in a 2- or 4-year College One Year After High School

	Panel A		Panel B		Panel C		Panel D	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2-Year	4-Year	2-Year	4-Year	2-Year	4-Year	2-Year	4-Year
3rd Grade Avg. Test Score Deciles (ref. Decile 1)								
Decile 2	0.318***	0.799***	0.313***	0.707***	0.286***	0.750***	0.321***	0.706***
	(0.016)	(0.027)	(0.016)	(0.027)	(0.026)	(0.045)	(0.035)	(0.065)
Decile 3	0.457***	1.272***	0.459***	1.127***	0.433***	1.194***	0.442***	1.203***
	(0.016)	(0.025)	(0.017)	(0.026)	(0.025)	(0.043)	(0.034)	(0.061)
Decile 4	0.563***	1.642***	0.567***	1.463***	0.547***	1.569***	0.599***	1.543***
	(0.016)	(0.025)	(0.017)	(0.026)	(0.025)	(0.042)	(0.034)	(0.059)
Decile 5	0.633***	1.955***	0.635***	1.744***	0.610***	1.858***	0.652***	1.840***
	(0.016)	(0.024)	(0.017)	(0.026)	(0.025)	(0.041)	(0.033)	(0.058)
Decile 6	0.688***	2.237***	0.678***	1.983***	0.668***	2.141***	0.762***	2.168***
	(0.016)	(0.024)	(0.017)	(0.025)	(0.024)	(0.041)	(0.033)	(0.058)
Decile 7	0.720***	2.509***	0.696***	2.214***	0.679***	2.375***	0.786***	2.428***
	(0.016)	(0.024)	(0.018)	(0.025)	(0.025)	(0.041)	(0.034)	(0.057)
Decile 8	0.749***	2.789***	0.713***	2.459***	0.723***	2.630***	0.851***	2.696***
	(0.017)	(0.024)	(0.018)	(0.025)	(0.025)	(0.040)	(0.034)	(0.057)
Decile 9	0.743***	3.067***	0.682***	2.679***	0.705***	2.878***	0.838***	2.943***
	(0.017)	(0.024)	(0.019)	(0.025)	(0.025)	(0.040)	(0.034)	(0.057)
Decile 10	0.621***	3.414***	0.520***	2.938***	0.579***	3.148***	0.774***	3.294***
	(0.019)	(0.024)	(0.020)	(0.025)	(0.026)	(0.040)	(0.036)	(0.057)
Student Demographics								
Race/Ethnicity (ref. White)								
Asian			0.512***	1.106***	0.741***	1.394***	0.879***	1.343***
			(0.016)	(0.014)	(0.055)	(0.082)	(0.069)	(0.114)
Black			0.322***	0.520***	0.592***	1.048***	0.706***	0.990***
			(0.017)	(0.019)	(0.040)	(0.070)	(0.053)	(0.101)
Hispanic			0.281***	0.229***	0.197***	0.423***	0.216***	0.408***
			(0.011)	(0.012)	(0.028)	(0.054)	(0.038)	(0.078)
Other Races/Ethnicities			-0.026	0.067***	-0.142**	0.229**	-0.047	0.277*
			(0.014)	(0.014)	(0.047)	(0.083)	(0.063)	(0.118)
Gender (ref. Male)								
Female			0.273***	0.506***	0.274***	0.507***	0.457***	0.592***
			(0.007)	(0.007)	(0.007)	(0.007)	(0.039)	(0.076)
Student Classifications (in 3rd Grade)								
Free-Reduced Price Lunch			-0.557***	-1.107***	-0.559***	-1.108***	-0.560***	-1.109***
			(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)
English Proficiency Services			0.391***	0.357***	0.385***	0.311***	0.386***	0.313***
			(0.014)	(0.017)	(0.014)	(0.018)	(0.014)	(0.018)
Special Education Services			0.021	-0.056***	0.020	-0.051**	0.025	-0.046**
			(0.014)	(0.016)	(0.014)	(0.016)	(0.014)	(0.016)
Learning Disability			-0.112***	-0.204***	-0.109***	-0.192***	-0.108***	-0.189***
			(0.023)	(0.033)	(0.023)	(0.033)	(0.023)	(0.033)
Interaction Effects:								
*Race/Ethnicity*Test Score Decile*	No	No	No	No	Yes	Yes	Yes	Yes
*Race/Ethnicity*Test Score Decile*Gender*	No	No	No	No	No	No	Yes	Yes
Constant	-1.462***	-2.843***	-1.439***	-2.577***	-1.433***	-2.730***	-1.518***	-2.771***
	(0.012)	(0.021)	(0.015)	(0.024)	(0.021)	(0.039)	(0.028)	(0.055)
Observations	507,101	507,101	506,661	506,661	506,661	506,661	506,661	506,661
Pseudo R-squared	0.0688	0.0688	0.0989	0.0989	0.0999	0.0999	0.100	0.100

Standard errors in parentheses *** p<0.001, ** p<0.01, * p<0.05

Figure 1A. Predicted Probabilities of 4-Year College Enrollment One Year After High School Across Third Grade Achievement Deciles, By and Gender, For Black, Hispanic and All Students



Notes. This figure shows the same enrollment probabilities from Figure 2 rearranged to make for easier comparisons, by race/ethnicity and gender, between all students and either Black (Panel A) or Hispanic students (Panel B). The short-dashed lines are the predicted probabilities for females and the long-dashed lines are the predicted probabilities for males. This focus of this plot is on points along the 3rd grade achievement distribution where the gender gaps for Black or Hispanic students are especially large relative to the gaps for all students. For Black students, both females and males drive the larger-than-average gender gap since each gender is significantly more likely to attend a 4-year college than the average female or male (compare, in particular, the gaps at deciles 1- 6 in Panel A). This is especially true for Black females whose difference from the average female (0.064) is larger than the difference between Black males and the average male (0.047). For Hispanics, it is females who have higher-than-average likelihoods of college enrollment among lower-achieving 3rd graders (see, for example, deciles 3 and 4 in Panel B). But at among high-achieving 3rd graders, it is Hispanic males who have significantly lower-than-average enrollment probabilities (see deciles 9 and 10 in Panel B).

Table 2A. Log Odds of Enrolling in a 4-year College Based on 3rd Grade and High School Tests

	Panel A	Panel B
	(1)	(2)
	3rd Grade Tests	High School Tests
Avg. Test Score Deciles (ref.Decile 1)		
Decile 2	0.542***	0.658***
	(0.030)	(0.039)
Decile 3	0.925***	1.233***
	(0.029)	(0.036)
Decile 4	1.206***	1.617***
	(0.029)	(0.035)
Decile 5	1.469***	2.024***
	(0.029)	(0.034)
Decile 6	1.714***	2.335***
	(0.028)	(0.034)
Decile 7	1.944***	2.635***
	(0.028)	(0.034)
Decile 8	2.187***	2.965***
	(0.028)	(0.034)
Decile 9	2.449***	3.350***
	(0.028)	(0.034)
Decile 10	2.816***	3.846***
	(0.029)	(0.034)
Student Demographics/Classifications		
Race/Ethnicity (ref. White)		
Asian	1.059***	0.803***
	(0.016)	(0.016)
Black	0.628***	0.784***
	(0.022)	(0.023)
Hispanic	0.135***	0.219***
	(0.014)	(0.014)
Other	0.099***	0.137***
	(0.016)	(0.017)
Female	0.490***	0.474***
	(0.008)	(0.009)
Free/Reduced-Price Lunch	-0.938***	-0.798***
	(0.010)	(0.010)
English Proficiency Services	0.207***	0.010
	(0.019)	(0.020)
Special Education Services	-0.073***	-0.142***
	(0.018)	(0.019)
Learning Disability	-0.179***	-0.189***
	(0.037)	(0.039)
Constant	-2.522***	-3.194***
	(0.027)	(0.033)
Observations	326,943	326,943
Pseudo R-squared	0.164	0.217

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 3A. Log Odds of Enrolling in a 4-year College, by Race/Ethnicity

	Panel A: 3rd Grade Test Deciles					Panel B: High School Test Deciles				
	(1) Asian	(2) Black	(3) Hispanic	(4) Other	(5) White	(6) Asian	(7) Black	(8) Hispanic	(9) Other	(10) White
Avg. Test Score Deciles (ref.Decile 1)										
Decile 2	0.394***	0.575***	0.555***	0.658***	0.578***	0.827***	0.754***	0.766***	0.657***	0.569***
	(0.099)	(0.089)	(0.052)	(0.103)	(0.052)	(0.190)	(0.104)	(0.071)	(0.112)	(0.061)
Decile 3	0.890***	0.905***	0.915***	1.032***	0.996***	1.443***	1.282***	1.398***	1.122***	1.169***
	(0.095)	(0.090)	(0.051)	(0.100)	(0.049)	(0.177)	(0.100)	(0.067)	(0.106)	(0.056)
Decile 4	1.233***	1.162***	1.172***	1.284***	1.293***	1.814***	1.564***	1.769***	1.500***	1.587***
	(0.093)	(0.091)	(0.052)	(0.099)	(0.048)	(0.173)	(0.101)	(0.067)	(0.103)	(0.054)
Decile 5	1.482***	1.355***	1.438***	1.513***	1.573***	2.290***	1.898***	2.166***	1.828***	2.016***
	(0.093)	(0.093)	(0.052)	(0.098)	(0.047)	(0.170)	(0.102)	(0.066)	(0.101)	(0.053)
Decile 6	1.709***	1.562***	1.619***	1.707***	1.846***	2.575***	2.249***	2.475***	2.122***	2.330***
	(0.093)	(0.094)	(0.053)	(0.097)	(0.047)	(0.169)	(0.103)	(0.066)	(0.101)	(0.052)
Decile 7	2.020***	1.638***	1.792***	1.972***	2.078***	2.949***	2.569***	2.773***	2.352***	2.629***
	(0.093)	(0.098)	(0.055)	(0.097)	(0.046)	(0.168)	(0.105)	(0.067)	(0.100)	(0.052)
Decile 8	2.263***	1.941***	2.032***	2.190***	2.320***	3.312***	2.671***	3.008***	2.677***	2.978***
	(0.093)	(0.101)	(0.056)	(0.096)	(0.046)	(0.168)	(0.111)	(0.068)	(0.100)	(0.052)
Decile 9	2.562***	2.104***	2.244***	2.399***	2.588***	3.735***	3.200***	3.435***	3.136***	3.336***
	(0.093)	(0.106)	(0.059)	(0.097)	(0.046)	(0.168)	(0.120)	(0.071)	(0.101)	(0.052)
Decile 10	3.009***	2.641***	2.552***	2.852***	2.933***	4.395***	3.496***	3.839***	3.595***	3.812***
	(0.094)	(0.124)	(0.063)	(0.098)	(0.046)	(0.168)	(0.143)	(0.078)	(0.103)	(0.052)
Student Demographics/Classifications										
Female	0.337***	0.522***	0.574***	0.414***	0.501***	0.373***	0.403***	0.555***	0.400***	0.485***
	(0.028)	(0.041)	(0.022)	(0.030)	(0.011)	(0.030)	(0.043)	(0.023)	(0.032)	(0.011)
Free/Reduced-Price Lunch	-0.514***	-0.457***	-0.506***	-0.874***	-1.190***	-0.418***	-0.313***	-0.377***	-0.652***	-1.047***
	(0.031)	(0.044)	(0.025)	(0.032)	(0.013)	(0.032)	(0.046)	(0.026)	(0.033)	(0.014)
English Proficiency Services	0.313***	0.653***	-0.033	0.060	0.048	0.057	0.316***	-0.112***	-0.034	-0.302***
	(0.044)	(0.076)	(0.027)	(0.113)	(0.059)	(0.043)	(0.076)	(0.027)	(0.115)	(0.060)
Special Education Services	-0.072	-0.175*	-0.024	-0.042	-0.059***	-0.144*	-0.159	-0.042	-0.121*	-0.149***
	(0.070)	(0.094)	(0.052)	(0.064)	(0.022)	(0.074)	(0.097)	(0.054)	(0.065)	(0.023)
Learning Disability	-0.622***	-0.100	-0.263***	-0.218*	-0.096**	-0.446***	-0.087	-0.134	-0.204	-0.192***
	(0.164)	(0.153)	(0.093)	(0.131)	(0.047)	(0.172)	(0.157)	(0.096)	(0.133)	(0.048)
Constant	-1.603***	-2.129***	-2.599***	-2.434***	-2.592***	-2.822***	-2.636***	-3.413***	-2.870***	-3.128***
	(0.087)	(0.084)	(0.050)	(0.091)	(0.045)	(0.166)	(0.094)	(0.065)	(0.094)	(0.050)
Observations	24,662	13,681	64,942	26,322	197,336	24,662	13,681	64,942	26,322	197,336
Pseudo R-squared	0.138	0.101	0.107	0.147	0.153	0.201	0.160	0.171	0.196	0.205

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1