Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 25-1241

Identifying Indicators to Support Educational Attainment for Different Groups of English Learners in High School

Elaine Allensworth
University of Chicago

Marisa de la Torre
University of Chicago

Kaitlyn Franklin
University of Chicago

Jinqing (Jean) Xu
University of Chicago

This study examines a broad array of potential indicators for early warning and college readiness indicator systems for different subgroups of English Learners in high school. Using data from 2008 through 2021 from the Chicago Public Schools, the study follows cohorts of students from eighth grade through college graduation, using logistic regression models to run exploratory analyses, and confirmatory analyses with omitted cohorts and machine learning methods. The study shows that current reporting obscures high achievement among some subgroups of ELs and identifies strategies with the most leverage for improving attainment for different EL subgroups.

VERSION: July 2025

Suggested citation: Allensworth, Elaine, Marisa de la Torre, Kaitlyn Franklin, and Jinqing Xu. (2025). Identifying Indicators to Support Educational Attainment for Different Groups of English Learners in High School. (EdWorkingPaper: 25-1241). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/yf35-e829

Identifying Indicators to Support Educational Attainment for Different Groups of English Learners in High School

Abstract

This study examines a broad array of potential indicators for early warning and college readiness indicator systems for different subgroups of English Learners in high school. Using data from 2008 through 2021 from the Chicago Public Schools, the study follows cohorts of students from eighth grade through college graduation, using logistic regression models to run exploratory analyses, and confirmatory analyses with omitted cohorts and machine learning methods. The study shows that current reporting obscures high achievement among some subgroups of ELs and identifies strategies with the most leverage for improving attainment for different EL subgroups.

Key words (3-5 allowed): English learners, multilingual learners, high school achievement, graduation, college readiness

Elaine Allensworth, 773-834-3061, elainea@uchicago.edu

Marisa de la Torre, 773-834-2698, mdelatorre@uchciago.edu

Kaitlyn Franklin, 470-388-2727, kaitlynfranklin@uchicago.edu

Jinqing (Jean) Xu, 646-464-1942, jinqingxu@uchicago.edu

All authors can be contacted at:

University of Chicago Consortium on School Research

Edith Abbott Hall, 969 E. 60th St., Chicago, IL 60637

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A220430 to the University of Chicago. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Identifying Indicators to Support Educational Attainment for Different Groups of English Learners in High School

English learners[1] (ELs) have the potential to bring much-needed multilingual skills to the professional workforce, especially if they have opportunities to graduate from high school and college. Most ELs aspire to graduate high school and earn a post-secondary credential (Gwynne et al., 2012; Shi & Watkinson, 2019), but high school ELs graduate high school, enroll in college, and obtain college degrees at far lower rates than their English-proficient peers (U.S. Department of Education n.d.a; Kanno & Cromley, 2013). There is a need to know how school staff could have the greatest impact for better supporting their educational attainment.

Most reporting on ELs focuses on standardized test performance, even though the tests are typically administered in English, a language in which ELs are, by definition, not yet fully proficient. Indicators of performance in high school other than test scores, including students' course grades, and behaviors (attendance, discipline records, study habits), are highly predictive of high school graduation and college outcomes in the general population of students (Bowen et al., 2009; Jackson et al., 2024), and used in Early Warning Indicator Systems (EWIS) to identify students who are at risk of not graduating from high school, and College Readiness Indicator Systems (CRIS), to identify whether students are likely to eventually obtain a college degree (see review in Balfanz & Byrnes, 2019). Discerning ELs' needs based on these types of indicators could be key to providing better support and increasing their educational attainment.

Case studies have noted that different subgroups of ELs have distinct academic needs (e.g., Clark-Gareca et al., 2019; Freeman et al., 2003; Le, et al., 2024), so improving educational

[1] There is debate about whether the term "English learners" is the right term to use to describe students whose home language is not English who have not passed an English proficiency exam and are enrolled in schools in the U.S. We use the term in this manuscript for consistency with usage by the Illinois State Board of Education and Chicago Public Schools.

attainment outcomes for subgroups of ELs may require different strategies. Many students transition out of EL status before high school, and these former ELs are rarely included on statistics about English learners. Other students who enroll in school in kindergarten or the early elementary grades never pass the English proficiency exam and enter high school as long-term ELs (LTELs). A disproportionate percentage of LTELs are identified as having a learning disability and qualify for both special education and English language services. Another group of ELs enroll in school in the U.S. for the first time in the middle grades or high school–these students are called "late-arriving" or "newcomer" ELs.

This study examines a wide array of potential indicators of progress towards high school graduation, college enrollment, and college degree attainment for different groups of ELs in high school in the Chicago Public Schools (CPS). ELs in CPS high schools are demographically similar to ELs in public schools across the nation: about 80 percent of Chicago ELs in high school come from Spanish-language backgrounds, which is similar to the nation where 75 percent of ELs have a Spanish-language background (U.S. Department of Education, 2020), and other ELs in CPS speak 70 different languages (de la Torre et al., 2019). Over 90 percent of ELs in Chicago are eligible for free or reduced-price lunch.

Once we understand how indicators of academic performance are related to educational attainment for different subgroups of ELs, we examine the performance of each of the subgroups of ELs on those indicators. This helps identify where strategies are most needed for improving educational attainment for that group. For each EL subgroup we ask:

RQ1: How are academic indicators in grades eight to twelve related to educational attainment?

a. Which single indicators are most predictive?
b. Which combinations of indicators substantially improve the prediction?
c. Do the indicators remain predictive when taking into account school effects?

d. What is the shape of the relationships of the strongest indicators?

RQ2: How do students in each subgroup perform on the indicators that are most predictive?

Prior Literature

In the fall of 2017, five million public school students across the country were classified as ELs, representing 10 percent of the school population, up from 3.5 million in the fall of 2000. Seven percent of ninth graders nationally are classified as ELs, and many other high school students were classified as English learners in earlier grades (U.S. Department of Education, 2020). Most ELs, like other students, aspire to more than a high school diploma (Gwynne et al., 2012; Shi & Watkinson, 2019). Yet, in 2015–16, only 67 percent of ELs nationwide graduated from high school in four years, compared to 84 percent for non-ELs (USDOE, n.d.a). Analyses using the National Education Longitudinal Study of 1988 showed that 18 percent of ELs enrolled in four-year colleges upon high school graduation, compared with 43 percent of native English-speaking students (Kanno & Cromley, 2015). These statistics indicate that schools in this country are not sufficiently supporting ELs to graduate high school and attend college.

Schools Often Support Educational Attainment Through Indicator Systems

A number of studies have shown positive impacts on students' outcomes associated with the use of EWIS and CRIS (Balfanz & Byrnes, 2019; Davis et al., 2019; Faria et al., 2017; Roderick et al., 2014). By organizing pieces of data on student performance into coherent data systems, school practitioners develop and test school strategies to improve students' educational attainment. Effective indicator systems require knowing which indicators (e.g., attendance, grades, test scores, suspension rates, English proficiency) are strongly related to educational attainment, and the degree to which students are struggling/succeeding on those indicators.

A wide array of indicators are used for early intervention around high school graduation and college readiness. For example, Bowers, Sprott, and Taff (2013) identified 110 predictors of

high school dropout/graduation that had been used in the literature, including course failures, low grades, disciplinary problems, grade retention, and low standardized test scores. Based on an extensive review of literature, Gurantz and Borsato (2012) suggested a menu of college-readiness indicators that included an array of academic indicators, indicators of tenacity, and indicators of college knowledge. Having an overwhelming number of data points can make it difficult to set priorities for improvement efforts. If some indicators are not strongly predictive for a particular group, or do not add value to the prediction beyond other potential indicators, data systems can be simplified to focus on those indicators that provide unique information.

Most high school graduation indicator systems focus on students' grades, attendance, and behavior. These are characterized in some places as the "ABCs" (attendance, behavior, course performance), and others as "BAG" (behavior, attendance, and grades). The focus on these indicators is supported by a number of studies showing that they are predictive of high school graduation in the middle grades and the first year of high school, with other indicators of achievement providing limited, if any, additional information about whether students will graduate (Allensworth & Easton, 2007, Balfanz & Byrnes, 2019; Bowers, 2010; Hartman et al., 2011; Kieffer & Marinell, 2012; Norbury et al., 2012; Stuit et al., 2016; Zau & Betts, 2008).

College readiness indicator systems also usually include students' course grades as a primary indicator, supported by research showing they tend to be the strongest predictors of college grades and of college graduation, compared to other potential academic indicators (Bowen et al., 2009; Camara & Echternacht, 2000; Geiser & Santelices, 2007). Standardized test scores, particularly the ACT and SAT, also feature prominently in college readiness indicator systems since they factor into college admissions and can determine whether students are eligible for financial aid (Tierney et al., 2009), as well as coursework and application steps.

When considering English learners, there may be additional indicators of academic performance that matter for educational attainment. English proficiency levels might matter– at the beginning of high school when students begin their high school program of study, or in eleventh grade when students take college entrance exams, or in twelfth grade immediately before entering college. We might also wonder whether the predictiveness of indicators changes during the high school years for English learners who are newcomers in the middle and high school years and might be adjusting to a new cultural context at the start of high school. Therefore, we examine a broad array of potential indicators of educational attainment, including those used in indicator systems for the general population, as well as indicators that are specific to EL-students around English proficiency, in the ninth-grade year and in later years.

High School ELs Are a Diverse Group of Students

Researchers generally classify high school ELs into two major groups: 1) long-term ELs; and 2) late-arriving ELs (Menken, 2013; Umansky et al., 2018). LTELs are students who, despite having been enrolled in U.S. schools for six or more years, have not passed the proficiency exam (Menken, 2013; Sugarman & Geary, 2018). Compared to the population of ELs in the early grades, LTELs are more likely to be male, eligible for special education services, and have attended multiple schools (Menken et al., 2012; Umansky et al., 2017). In Chicago, about one-fifth of students who began kindergarten as ELs and remained stably enrolled through eighth grade did not reach English proficiency before high school. Among LTELs in Chicago in eighth grade, 55 percent had an identified disability (de la Torre et al., 2019). Thus, over half of LTELs are doubly-identified as needing academic supports in high school, and LTELs with IEPs likely have different experiences and outcomes in high school than LTELs without IEPs–requiring further differentiation when comparing factors influencing students' educational attainment.

Late-arriving ELs are another group of active ELs in high school. According to estimates from the 2012-16 American Community Survey, approximately 35 percent of Illinois ELs in grades 6-12 were born outside of the U.S., compared to 13 percent of Illinois ELs in earlier grades (Sugarman & Geary, 2018). Schools serving late-arriving ELs must not only facilitate their English language development but also provide support as they adapt to a new cultural and schooling context, deal with the social challenges of dislocation, and potentially contend with academic needs related to interrupted formal education (Menken, 2013; Short & Boyson, 2012; Umansky et al., 2018).

Still another group to consider is former ELs. Most students who begin kindergarten as ELs pass the English proficiency exam in grades 3-5 and then are no longer considered ELs. Traditionally, states and districts have not reported on former ELs or included them in their statistics on ELs. With the signing of the federal Every Student Succeeds Act (ESSA) in 2015, states can now choose to include former ELs for up to four years after achieving proficiency in their metrics on ELs. Most of the research that exists on former ELs has been limited to high school graduation rates (e.g., Huang et al., 2016; Kieffer & Parker, 2017). Little is known about college attainment for this group, and the supports that matter most for them.

This study identifies the best early indicators for diverse groups of English learners

The primary goal of this exploratory study is to identify which indicators are most strongly associated with educational attainment outcomes for each of several subgroups of English learners: former ELs, LTELs without IEPs, LTELs with IEPs, and late-arriving ELs. This can point school practitioners to the indicators that are most critical in their Early Warning and College Readiness Indicator Systems. We then further identify how students in the subgroups perform on those indicators to identify where school staff would likely have the most impact for improving students' educational attainment. For example, we might find that a

particular indicator is highly predictive, but that almost all students in a given subgroup perform at levels that make them very likely to obtain an education milestone, so focusing on that indicator it would make little difference.

Sample and Methodology

We draw on an extensive longitudinal archive of administrative data, using information from the middle grades through post-secondary years of students in CPS. Through a long-standing data sharing agreement with CPS, the research team has access to data on the population of students in the district in each semester, with a student identifier allowing us to follow students' academic performance over time, including annual attendance, standardized assessments, and demographic data. The administrative data files also include information on the annual ACCESS assessment that determines whether ELs demonstrate English proficiency each year. Data from the National Student Clearinghouse (NSC) provide information on students' post-secondary enrollment and graduation.

The analytic sample used for this study consists of cohorts of students who started ninth grade for the first time in the fall of 2008 through 2017 in district-run (non-charter) schools in CPS. These students would have graduated high school in the spring of 2012 through 2021, if they graduated in four years, and potentially enrolled in college in fall 2012 through 2021. To study college degree completion, we limit the analyses to cohorts who started ninth grade in the fall of 2008 through 2011, allowing ten years to obtain a college degree (four years of high school and six years of college). Students in charter schools are not included because the district does not have complete records of their course performance. If students transferred to a charter school after the fall of their ninth-grade year, they are still part of the analytic sample but might have missing data on course grades, leading to some missing data in the analysis of indicators for a small number of students. We required students to be enrolled for at least 110 days during their first time in ninth grade to ensure that students were enrolled in CPS schools long enough to take

the ACCESS test–which measures English proficiency in CPS–and is typically taken in January. That allows us to identify active ELs. We also removed students who had a validated transfer out of CPS within the first four years to another school district since their outcomes are unknown.

We held back two cohorts from our exploratory analyses so that we could conduct confirmatory analyses, described at the end of the manuscript. The main analyses are based on all cohorts except students who started ninth grade in 2010 and 2016; these cohorts were neither the earliest nor the latest cohorts being studied and include a cohort with data through college graduation as well as a more recent cohort.

Students are classified in a group based on their status at the beginning of ninth grade:

- Never ELs: Students who have never been classified as ELs.
- Former ELs: Students who were formerly classified as English Learners during their time in CPS but demonstrated English proficiency through the ACCESS test and exited EL status by the start of ninth grade.
- Active ELs: Students who were classified as an English Learner at the start of ninth grade. Among active ELs we define the following groups:
    - *Late-arriving ELs: Active ELs* in the ninth grade who had been in CPS for *fewer than six years* (i.e. after third grade).
    - *LTELs without an Individualized Education Plan (IEP): Active ELs* in the ninth grade who had been in CPS for *six or more years* (i.e. third grade or earlier) without demonstrating English proficiency through the ACCESS test and *did not* have an IEP for an *identified disability* at the beginning of ninth grade.
    - *LTELs with an IEP: Active ELs* in the ninth grade who had been in CPS for *six or more years* (i.e. third grade or earlier) without demonstrating English proficiency

through the ACCESS test and had an IEP for an *identified disability* at the beginning of ninth grade.

Former and active ELs in CPS high schools were more likely to be eligible for free or reduced-price lunch (FRPL) and to be Latinx than students who were never ELs (see Table 1). Over 90 percent of LTELs were both FRPL eligible and Latinx. LTELs with IEPs were more likely to be male (63 percent) compared to around 50 percent of students in other groups. Late-Arriving ELs were less likely to be Latinx and more likely to be Asian than students in other EL groups. Both former ELs and late-arriving ELs were less likely to be identified with a disability than never ELs or LTELs. Late-arriving ELs were more likely to be new to CPS in ninth grade than other students; 22 percent first enrolled in the district in ninth grade.

[Table 1]

Educational attainment is measured with three dichotomous variables (0 or 1): high school graduation, immediate college enrollment, and college degree completion:

- High school graduation coded as "1" if students graduate high school within four years of starting ninth grade, with GED and alternative high school diploma recipients counted as non-graduates. Students with IEPs are more likely than other students to take more than four years to graduate, so we also ran models that allowed six years to graduate; the conclusions about which variables mattered most were the same. The only difference was that graduation rates were slightly higher for all groups.
- Immediate college enrollment coded as "1" if students enroll in any college (two- or four-year) within four years of starting ninth grade. We use ninth grade as the base for the indicator so that non-graduates are included (most will be non-college enrollees).
- College degree attainment is coded as "1" for students who receive either a 2-year or 4-year degree within 10 years of starting ninth grade. We combine 2- and 4-year colleges

because of the complexity of the analyses that include three outcomes and five subgroups of students. Future studies can separately examine attainment of 2-year versus 4-year degrees. We initially ran models that predicted college degree attainment within eight years of ninth grade, allowing four years for high school and four years for college. However, because many students take five or six years in college to attain a degree, we also decided to look at attainment within 10 years. The indicators that matter were similar in either analysis, but we show the 10-year rates to be more comprehensive.

Figure 1 shows the educational attainment outcomes for all ten cohorts of students by EL subgroup. Former ELs often have better educational attainment outcomes compared to students who were never ELs, while students who were active ELs in high school have lower attainment rates. Attainment rates at each milestone are the lowest for LTELs, especially LTELs with IEPs.

Methodology

This study is done as part of a research-practice partnership that includes external researchers, district representatives, and policy advocates. The questions and methods are designed to provide information to practitioners and nonprofits working with schools and families as they support higher levels of educational attainment for English learners. The primary analyses are exploratory, with additional analyses to confirm the results would be similar with different samples and methods, described in the section on robustness checks.

To determine which variables are the best predictors of educational attainment (RQ1a), we use logistic regression models predicting each educational attainment outcome (e.g., high school graduation, college enrollment, and college degree attainment) separately for each subgroup. We build the models up with an increasing number of variables and compare McFadden pseudo-R^2 statistics (defined as one minus the ratio of the log likelihood value from the estimated model and the log likelihood value of a model with only an intercept and no covariates) from across the models to gauge whether additional variables improve the prediction

beyond the indicators already in the model. First, we enter each variable by itself, then pairs of predictors, then all potential high school indicators, and finally we include background characteristics. All of the high school achievement variables were entered as categorical variables (or a series of dummy variables) which allowed for non-linear relationships, and for inclusion of students with missing data on any one variable. However, we found that the pseudo-R^2 statistics from models of indicators in later grades were overly influenced by students with missing data because non-graduates were included in the category of students without data, which often indicated non-graduates. Therefore, we re-ran the models only including students with complete data, which is what we show here. The same indicator variables emerge as the strongest predictors as in the models that include students with missing data.

The models include background characteristics and potential high school indicators, including attendance, GPAs, days suspended, course pass rates, and standardized test scores. For active ELs, we include ACCESS proficiency levels, whether a student reached English proficiency and was reclassified out of EL status, and their annual growth on the ACCESS. Each variable is measured in eighth grade and at each high school grade level, see Table 2 for details.

[Table 2]

Student background characteristics used in the models included: gender, race/ethnicity, census block social status, eligibility for FRPL, special education status, census block concentration of poverty, age at entry to high school, and school mobility in middle grades. Social status and concentration of poverty provide additional information on economic status beyond free- or reduced-price lunch status, for which the vast majority of students in CPS qualify. These are created from U.S. Census data linked to students' residential block groups; there are approximately 2,500 block groups among CPS students and are often equivalent to one city block. Social status is a measure calculated from the percent of employed persons who are

managers or executives and average education level among adults, and the concentration of poverty is calculated from the percentage of families under the poverty line and male unemployment rates. Mobility in the middle grades is calculated as the number of different schools a student attended from 6th through 8th grade. Age at entry to high school has two measures - number of months old for high school, which is a continuous variable for students who are over the age of 14 when they enter 9th grade, and a true/false indicator for whether a student is `young for high school' and enters 9th grade before the age of 14.

To identify the shape of the relationships of the indicators with the outcomes (RQ1b) we use bar graphs showing the percentage of students attaining each outcome at different levels of the most-predictive indicators. These figures allow us to determine at what points the probability of attaining a degree is nearly universal (about 90% or higher), at some risk (around 70%), highly uncertain (around 50%), or highly unlikely (less than 25%). This has implications for the kinds of interventions that are needed, since students at some risk require monitoring but may not need an intervention. Students whose outcomes are highly uncertain are most likely to show a change in outcomes based on whether they receive support. And students whose outcomes are highly unlikely need intensive supports.

To determine whether systems that contain multiple indicators would produce substantively better predictions of attainment outcomes (RQ1b), we build on the models presented for RQ1a by examining the pseudo-R^2 statistics that result from pairs of indicators. We focus on indicators from the ninth-grade year, since this year is the focus of many early warning indicator systems. If the pseudo-R^2 statistics improve by more than 0.02 points above the best single indicator we consider it potentially worthwhile to increase the complexity of the indicator system by including information on multiple indicators. We then combine all of the ninth-grade indicators as predictors in models of each attainment outcome to determine whether the inclusion

of more than two indicators would substantively improve the prediction of the outcome. Finally, we run a model that also includes students' background characteristics (e.g., race, ethnicity, economic status, and eighth grade academic achievement) to determine whether background matters beyond school performance indicators.

It is possible that some of the high school achievement indicators are predictive of educational outcomes not because achievement matters, but because they attend schools where all students are more likely to graduate, attend college, and attain college degrees. If an indicator of educational attainment is predictive only because it is an indicator of which school a student attends, and not because of an individual student's academic readiness, then working to improve that indicator will be unlikely to improve a students' likelihood of attaining the educational milestone without changing their high school. We run a series of models that include the most predictive achievement indicators as predictors of each outcome using standardized, continuous versions of the indicator instead of the non-parametric series of dummy variables along with high school fixed effects. This allows us to discern whether the size of the relationship (the coefficients on the indicators) changes when we add school fixed effects to the models (RQ1d).

Finally, we use descriptive methods to show how each subgroup performs on the high school indicators that are most predictive of educational attainment (RQ2). This shows us where there is the greatest potential for improving the educational attainment of each group.

Results

Which single indicators are most predictive of educational attainment?

We begin by showing the strength of the relationship of each potential indicator with each educational attainment outcome through pseudo-R^2 statistics. Note that these are not the same as R^2 statistics from a linear model, but a higher pseudo-R^2 does indicate a more accurate prediction of the outcome than lower pseudo-R^2 of the same outcome with the same sample. In

general, values ranging from 0.2 to 0.4 indicate an excellent fit (McFadden, 1979). Pseudo-R^2 are shown in Table 3 for models run separately for each of the EL subgroups, as well as for students who were never ELs as a comparison. Coefficients for grades eleven and twelve are not shown for outcomes other than English proficiency because of table length; they are similar in size to those for tenth grade and available from authors.

[Table 3]

Across all three attainment outcomes, students' GPAs have the strongest relationships with later educational attainment compared to the other potential indicators. Attendance rates and course failure rates also tend to have strong relationships with high school graduation. For example, tenth grade GPAs have pseudo-R^2 statistics of between 0.28 - 0.33 for all EL groups except LTELs with IEPs when predicting high school graduation, which are the highest pseudo-R^2s of any tenth-grade indicators. These are followed by pseudo-R^2 statistics for tenth grade course failure rates of 0.23 - 0.29, and attendance of 0.21 - 0.27 when predicting high school graduation among all EL groups except LTELs with IEPs. When predicting college enrollment, the pseudo-R^2 statistics for tenth grade GPA range from 0.13 to 0.16, and those for college degree completion range from 0.16 to 0.22 for all EL groups except LTELs with IEPs. These are the highest pseudo-R^2 statistics at grade 10 for the EL groups.

None of the indicators based on ACCESS scores show strong relationships with high school graduation or college outcomes for any of the EL groups; the pseudo-R^2 statistics are all below 0.10 with the exception of ACCESS proficiency level for LTELs with IEPs. Reaching proficiency on the ACCESS so that a student is no longer considered an active EL in high school is not a strong predictor of any of the attainment outcomes for any of the EL groups.

Ninth and tenth grade exams taken in English show relationships that are just slightly higher than the relationships of the ACCESS proficiency levels for high school graduation and

college enrollment; the pseudo-R^2 statistics of the high school exams (EXPLORE, PLAN, ACT or PSAT, SAT) are all below 0.10 for EL groups. For college degree completion, the pseudo-R^2 statistics for the high school exams are higher for the former EL and late-arriving EL groups, with pseudo-R^2 statistics around 0.09-0.13. The standardized tests, which are administered in English, are stronger predictors of college enrollment and college completion for never EL students than for either former or active EL students in high school.

There are several differences across grade levels and subgroups

In general, ninth grade indicators tend to be considerably more predictive than eighth grade indicators for all groups, suggesting that the transition to high school is a year when many students' academic trajectories change. Tenth grade indicators tend to be even more predictive than ninth grade indicators for all students, but for late-arriving EL students, there is more improvement in the predictions of educational attainment outcomes in tenth and eleventh grade, compared to the ninth grade, than for the other groups. For late-arriving ELs, the fact that ninth grade indicators are somewhat less predictive may indicate that the first year of high school is a greater year of flux for them. They may need more support for navigating high school during the ninth-grade year than other students. This group of students is more likely to be new to the district in ninth grade than students in other EL groups.

For LTELs with IEPs, the attainment outcomes are more poorly predicted by high school achievement indicators than for other groups of ELs. The strongest predictors for students with IEPs have much lower pseudo-R^2 values than the strongest predictors for the other EL groups. Across the outcomes, the strongest predictor for LTELs with IEPs has pseudo-R^2 that is about half the size of the strongest predictor for the other groups, suggesting it is more difficult to predict educational attainment outcomes based on high school indicators for LTELs with IEPs.

We wondered whether the indicators were less predictive for LTELs with IEPs because they were doubly-identified for services, or if the patterns were similar for never EL students with IEPs. Therefore, we ran the same models just for never ELs with IEPs. As observed with ELs, the outcomes of never EL students with IEPs were much less precisely predicted than for never EL students without IEP (see Table A1 in Online Appendix A). At the same time, the indicators were even less predictive of outcomes for LTELs with IEPs than for students with IEPs who were never ELs.

Which combinations of indicators substantially improve the prediction?

The number of potential combinations of indicators is extremely large. Therefore, we focus on combinations of ninth grade indicators for the remaining analyses as their pseudo-R^2 are not dramatically different from the indicators in later grades, and focusing in ninth grade helps practitioners identify students at risk early in high school. Table 4 presents the pseudo-R^2 statistics for pairs of ninth grade high school indicators in the middle of the table, with those from single-indicator models at the top of the table for comparison. At the bottom of the table are the pseudo-R^2 statistics from models that include all ninth-grade indicators together; these can be compared to the two-indicator models to discern whether the prediction of the attainment outcome would improve substantially with more than two indicators. Finally, there are pseudo-R^2 statistics for models that include students' background characteristics. These show whether providing additional background information beyond high school achievement indicators would improve the prediction further.

[Table 4]

High school graduation is best predicted by attendance in combination with either GPAs or course failure rates for most of the EL groups. For former ELs and long term ELs without IEPs, the pseudo-R^2 statistics are around 0.33 when combining GPA and attendance, while the

pseudo-R^2 statistic is smaller for late-arriving ELs at 0.24 because ninth grade indicators are less predictive for late-arriving ELs (see Table 4). The combination of attendance with GPA increases the pseudo-R^2 statistics by 0.04 - 0.06 points over ninth grade GPA alone. Including all of the ninth-grade indicators only increases the pseudo-R^2 statistics by about 0.02 above the most predictive models with two indicators for these groups.

The exception is among LTELs with IEPs. For LTELs with IEPs, the combination of ACCESS scores along with either GPA, failure rates, or attendance provides the strongest prediction, with pseudo-R^2 statistics of about 0.20. This more than doubles the pseudo-R^2 statistics relative to the single-indicator models for this group. However, even with multiple indicators, the pseudo-R^2 statistics are lower than those for the other groups of students, indicating more imprecision in predicting high school graduation for ELs with IEPs.

College enrollment is less strongly predicted by high school achievement for students who were former and active ELs than those who were never ELs. Background factors play a larger role in college enrollment for several EL groups than for never ELs (see Table 4). For late-arriving ELs and LTELs without IEPs, background characteristics increase the pseudo-R^2 statistics by 0.05 to 0.08 beyond models without background characteristics, indicating that enrollment decisions for these groups are more strongly influenced by factors outside of school performance. Combinations of GPA, attendance, failure rates, and standardized test scores improve the prediction of who will enroll in college beyond single indicators.

GPA and attendance together provide the best prediction of college enrollment for all groups except students with IEPs, with pseudo-R^2 statistics of 0.13 - 0.15 for students in the EL groups. Other two-indicator combinations are about as predictive, including combinations with failure rates and standardized test scores. Combining all ninth-grade indicators together increases the pseudo-R^2 statistics just slightly higher than the 2-indicator combination for former ELs and

LTELs without IEPs (by about 0.02 compared to the most predictive combination), but a larger increase (of about 0.04) for LTELs with IEPs and late-arriving ELs. For students with IEPs, even combining all of the indicators together produces a low pseudo-R^2 (0.11). It is also more difficult to predict immediate college enrollment for students in any of the EL groups than for students who were never EL, for whom the pseudo-R^2 statistics are larger than for the EL groups.

The combination of GPA with test scores leads to a better prediction of who will eventually obtain a *college degree* than GPA alone. For all EL groups, combining students' GPA with their standardized test (EXPLORE/PSAT) score produces the best two-indicator prediction (with pseudo-R^2 statistics of 0.15 to 0.22, see Table 4). Adding all the indicators together in the model improves the prediction by less than 0.02 points, except for LTELs with IEPs for whom the prediction becomes almost as strong as the prediction for former ELs when all indicators are considered (0.19 compared to 0.21). As seen with college enrollment, background characteristics improve the prediction of college degree attainment for all groups.

Do the Indicators Remain Predictive Taking into Account School Effects?

To adjust for potential high school effects, we ran models predicting educational attainment with and without school fixed effects (FEs) (see Table 5). With the prior models we were interested in the model statistics (the pseudo-R^2) to discern which indicators explained the most variance in outcomes. Here we are interested in the coefficients representing each indicator to discern changes in their relationships with each outcome when school fixed effects are added. To make it easy to compare coefficients, we used continuous versions of each of the most-predictive indicators from the earlier models: ninth grade attendance, ninth grade GPAs, and ninth grade test scores. The indicators were standardized so they each have the same meaning: the log-odds change in the probability of the outcome for each standard deviation change in the indicator. All three indicators are entered simultaneously in the models shown. The relationships of the indicators with the outcomes are modeled separately for each EL group through a no-

intercept model with interaction terms between each EL group and each indicator variable. The main effect terms for each EL group simply represent the average outcome for each group when all three achievement indicators are at the overall mean. The sample size changes very slightly in the models with school fixed effects because some very small schools have no variation in the outcomes (e.g., all students graduate or no students obtain a college degree).

[Table 5]

For *high school graduation*, students' attendance and GPAs have the same or similar values in models with and without school fixed effects. The coefficient for GPAs is over 1.0 for all groups except students with IEPs, for whom it is 0.82, and remains about as strong or stronger when school FEs are added. The coefficient for attendance ranges from 0.6 to 0.9 across the groups and is about the same size in the model with FEs as the model without FEs. For most groups, the coefficients associated with test scores get smaller when school FEs are added, but test scores show relatively modest relationships with high school graduation regardless of whether school FEs are included.

For *college outcomes*, school effects explain some of the relationships of attendance and test scores with college outcomes. The changes are moderate when predicting college enrollment; for example, for former ELs, the standardized test coefficient declines from 0.535 to 0.359. The changes are larger when predicting college degree attainment; the relationships of standardized test scores with college completion are about half their original size for never ELs, former ELs, and LTELs without IEPs (e.g., shrinking from 0.655 to 0.325 for former ELs), and somewhat smaller for late-arriving ELs and LTELs with IEPs (e.g., shrinking from 1.025 to 0.876 for late-arriving ELs). In general, attendance is less predictive of college degree attainment than GPAs or test scores when the three indicators are included in the models than when entered singly (not shown), and the size of the coefficient is even smaller when school FEs are added.

The coefficients for GPAs get bigger when controlling school FEs, which is a pattern seen in prior studies (e.g., Bowen et al., 2009). There is a suppression effect because it is harder to get high grades in schools with high achieving students and those are schools with higher college degree completion rates. As a result, GPAs emerge as the strongest predictor of college degree completion when taking into account school effects for never ELs, former ELs, and LTELs without IEPs. But both test scores and GPAs remain strong predictors of college degree completion for late-arriving ELs and LTELs with IEPs, even though the tests are taken in English.

What Is the Shape of the Relationships?

Figure 2 shows how ninth grade GPA, semester course failures, and attendance are related to *high school graduation*. In most subgroups, students with a 2.0 GPA (C average) or higher in ninth grade or no semester course failures, have about a 90 percent or higher probability of graduating high school. One notable exception is for students with IEPs, whose probability of graduating is well below 90 percent even if they have GPAs above a 2.0, or no course failures. We wondered if this was because some students' IEPs permitted more than four years to graduate, so we also looked at graduation patterns allowing students six years to graduate. This increased the graduation rates for all students, but the large discrepancies in graduation rates for students with IEPs with high achievement remained.

An additional minor exception is for late-arriving ELs whose probability of graduating high school is 87 percent if they have a GPA of 2.0 - 3.0 (C to B average), but it is 94 percent if they have a 3.0 average or higher. Late-arriving ELs also have a slightly lower probability of graduating than students in other groups if they have a GPA between 1.0 and 2.0 (66 percent versus 69-74 percent for other groups). These slightly lower probabilities correspond with the

relatively greater instability of ninth grade indicators for late-arriving ELs than for other student groups. Tenth grade GPA is notably more predictive for this group than ninth grade GPA.

In all subgroups, students with a ninth grade GPA at or below a 1.0 (D Average) are unlikely to graduate high school; they have about a 1-in-4 chance of graduating (about 25 percent graduate in each group). Likewise, students who have four or more semester Fs in ninth grade have about a 1-in-3 chance of graduating, or less. There is also a considerable difference in graduation rates among students with C-D averages compared to students with B-C averages; students with a 1.0 - 2.0 average have about a 70 percent chance of graduating, while students with above a 2.0 have about a 90 percent chance of graduating, except for students with IEPs. Differences among the EL and never EL groups in graduation rates are small relative to the differences in graduation rates by students' ninth grade GPA.

There is a gradual decline in graduation rates with lower attendance for all of the student groups, but attendance shows a stronger relationship with graduation for students who are active ELs in high school than students who are not active ELs in high school. Graduation rates are above 95 percent for all groups of students whose ninth-grade attendance is 98-100 percent, except for LTELs with IEPs whose graduation rates are 84 percent even with close to perfect attendance in the ninth-grade year. Among students with slightly lower attendance in ninth grade (95-98 percent), students in the three active-EL groups in high school have lower graduation rates than the former or never ELs, and these differences increase further among students with 90-95 percent attendance. Among students who attend 90 percent or less, there are differences in graduation rates observed among all of the subgroups, with former ELs showing lower graduation rates than never ELs, followed by late-arriving ELs, and then LTELs.

Attendance provides much further differentiation of students' risk of not graduating among students with ninth grade GPAs below a "C" (2.0) GPA, as seen in Table 6. Among

active ELs in high school (not including LTELs with IEPs), graduation rates for students with "C-D" GPAs (1.0 - 2.0) are between 46 and 87 percent depending on their ninth-grade attendance. Among students with "D-F" ninth grade GPAs (less than 1.0), graduation rates range from 14 to 49 percent, depending on their attendance. For students with IEPs, ninth grade ACCESS scores provide additional information beyond students' GPAs, more so than attendance (see Table 7). The relationship between ninth grade GPAs and high school graduation rates is similar to that of students without IEPs among students with IEPs and ninth grade ACCESS scores of 3 or higher. But for students with IEPs who have ACCESS scores below 3.0, there is not a clear relationship between GPAs and graduation rates.

[Table 6 and Table 7]

College outcomes are strongly associated with GPAs for active ELs, but their college outcomes are lower than students with similar achievement who were not active ELs. Each grade-point increase in ninth grade GPAs is associated with a 10-25 percentage point increase in the probability of enrolling in college for all groups, except students with IEPs, as seen in Figure 3. At the same time, at all GPA levels, college enrollment rates are lower among students who were active ELs in high school compared to students with similar academic records who were not active ELs. For example, among students with A-B averages in ninth grade, college enrollment rates ranged from 29 to 62 percent across the active EL groups compared to enrollment rates of 73-77 percent among never and former ELs. College enrollment rates were particularly low for LTELs with IEPs compared to all other EL groups among students with similar GPAs. The same pattern can be seen in college degree attainment rates, where active ELs are much less likely to attain a college degree in ten years after starting high school than never or former ELs with the same ninth grade GPAs, while ELs with IEPs are even less likely to attain a college degree than students in the other EL groups.

Overall, the differences in college outcomes based on ninth grade GPAs are much larger than the differences based on EL group status. College degree attainment rates are only above 25 percent for students with an A-B ninth grade GPA in all groups, see the right panel of Figure 3, with the exception of students with IEPs for whom college degree attainment rates are below 15 percent at all GPA levels. As noted earlier, for late-arriving ELs and LTELs with IEPs, ninth grade test scores provide additional information beyond GPAs about college graduation, even after controlling for school effects. As shown in Table 8, among students with "A-B" averages in these groups, college degree attainment rates are at or above 50 percent if their ninth grade test scores were average or above relative to the district average, but if they were more than half a standard deviation below the mean, their attainment rates were less than 33 percent for late-arriving ELs and less than 17 percent for LTELs with IEPs. About a third of late-arriving ELs and less than 10 percent of LTELs with IEPs have average scores or above on the tests that are given in English in ninth grade (see Figure 4).

[Table 8]

How Do Students in Each Subgroup Perform on The High School Indicators?

There are considerable differences across the EL groups in their performance on the indicators that are most predictive of educational attainment, and these differences correspond to differences in educational attainment shown previously in Figure 1.

Ninth grade GPAs and attendance were stronger among former ELs and late-arriving ELs than never ELs and other subgroups. One-third of students in these two groups finished ninth grade with an A-B GPA, and another third finished with a B-C GPA (see Figure 4). About a third of students in these groups finished the ninth grade with less than a "C average" GPA. Former and late-arriving ELs also had very strong ninth grade attendance, on average, with over a

quarter of students in these groups attending 98 percent or more and more than half attending 95 percent or more. Their grades and attendance were stronger than students who were never EL.

LTELs without IEPs had much lower GPAs than students in other subgroups, driven by a high number of course failures. Fourteen percent of long term ELs without IEPs had a A-B GPA in ninth grade, while 34 percent had a C-D average, and 17 percent had less than a D average. Twenty percent of LTELs without IEPs had 4 or more course failures in ninth grade, compared to 10 to 12 percent for their peers. In contrast, LTELs with IEPs were not more likely to have GPAs below a D average, although were much less likely to have A-B averages than students in other groups. Both LTELs with and without IEPs were more likely to have very low attendance than students in other EL groups, although not lower than students who were never EL. Close to a fifth of LTELs (19 percent) had attendance rates of less than 85 percent, while about a third were chronically absent, with attendance rates of less than 90 percent.

Active ELs were more likely to score low or very low on standardized tests given in English than former and never ELs. This is not surprising, given that English language proficiency in the district is determined by a standardized test of English-language arts (ELA) skills. Among late-arriving ELs, 63 percent scored in the low or very low range. Seventy-five percent of 4 LTELs without IEPs and 93 percent of LTELs with IEPs scored in the low or very low range.

Robustness Checks

As an exploratory study, we did not begin with *a priori* hypotheses but allowed the data to show which indicators demonstrated the strongest relationships with educational attainment. One danger in allowing the data to inform our conclusions is that the models will be overly fitted to the specific sample of students used for analysis. We conducted several different analyses to verify that the main findings remain similar with changes to the sample and methods.

Confirmatory Analyses

We reran the analyses with just the excluded 2010 and 2016 cohorts (for college completion we just used the 2010 cohort). In general, the same variables come out as the best predictors for these two cohorts as with the other cohorts; see Table A2 in the Online Appendix.

Addressing Missing Data

The model results shown in Table 3 include all students with non-missing data, which results in different sample sizes in each cell. We also ran models that include dummy variables for students who are missing any given variable, allowing for a consistent sample of 147,780 students for each analysis in each cell. The same indicators emerged as the strongest predictors of the achievement outcomes in the models that included dummy variables for missing data, so using these alternative models would not affect our conclusions (see Table A3 in the Online Appendix).

Using Alternative Methods

We used machine learning techniques to ensure the results were reliable and consistent across different methodologies. The Least Absolute Shrinkage and Selection Operator (LASSO) method produces sparse models with decision rules for selecting only the most important predictors, offering a good balance between prediction accuracy and model complexity. For a detailed explanation of this method and figures showing the results see Online Appendix B.

In general, we find that LASSO models identified the same key indicators for predicting the different educational attainment outcomes as the more traditional methods. GPA, attendance, and course failures were the variables selected through LASSO for predicting high school graduation. For predicting college enrollment, the LASSO results indicate that GPA is the most predictive variable for all students, and standardized test scores improved the prediction for the

subgroups other than LTELs. GPA and test scores are the key indicators for predicting degree attainment for all groups of students except for long-term English learners, for whom the model did not identify any indicators with the one standard error rule. Some demographic variables, including gender, race, and FRPL improve the prediction of the college outcomes for some groups.

Discussion and Areas for Future Research

Many active ELs and former ELs have strong high school achievement, with correspondingly strong educational attainment rates. Former ELs and late-arriving ELs show strong high school performance, on average, which may not be recognized because reporting often groups them together with the general population (in the case of former ELs) or other active English learners (in the case of late-arriving ELs). The strong performance of late-arriving ELs in Chicago suggests that the language supports that are in place in the district allow most students to engage in learning at a similar level as students whose native language is English.

These results may differ from other places, and it could affect our determination of which indicators were most predictive of later outcomes. Active English learners in some other places are less likely to take college-preparatory coursework (Callahan & Shifrer, 2016; Stephens & Francis, 2018), but we did not find this to be the case in Chicago; English learners in all of our subgroups took similar numbers of college-prep courses as former and never ELs (analysis available from authors). That may be why we find such modest relationships between English proficiency levels and educational attainment outcomes. It would be worthwhile to compare policies and practices around enrolling English learners in classes in different places, and the resulting outcomes for students.

At the same time, active English learners in Chicago high schools do not have equivalent attainment outcomes to students who were former or never ELs, suggesting a need for changing

or strengthening practices to better support educational equity. Disproportionate numbers of LTELs need more support to pass their classes and earn good grades in high school. In addition, active ELs in high school face more barriers to college degree attainment than never ELs who have similar achievement, and background factors such as economic status matter more than for students who were never EL.

Supporting attendance and grades would most benefit all groups' educational attainment

English proficiency may seem like the priority for ELs, but these analyses show that improving grades and attendance is critical for supporting ELs' educational attainment. Language supports can be provided in different ways; some may be designed to support students' success in their classes, while others focus on English language development without a clear connection to students' other classes. For example, schools could allocate time for English as a second language (ESL) teachers to collaborate with content class teachers to help embed language supports and activities appropriate for developing language skills in core content classes. ESL and bilingual teachers could also provide direct language and academic support to students as needed in their core content classes, supporting vocabulary development or providing academic assistance with course assignments (e.g., reading passages, writing responses, and papers). These might be more helpful for improving students' grades in core content than stand-alone ELA-development classes unconnected to students' other coursework, especially for LTELs who have achieved oral proficiency in English. Such strategies also have more evidence of effectiveness for middle-grade students than stand-alone classes (Baker et al, 2014).

Schools don't need different systems of early warning and college readiness indicators for EL students and the general population, but thresholds for which students are at risk differ. Course grades are the indicators that are most predictive of educational attainment for never ELs and ELs. At the same time, the threshold at which students become at risk for meeting/not

meeting an educational attainment milestone is different for students in particular EL groups than for students who were not active ELs in high school. High attendance rates are particularly important for active-ELs in high school, relative to former and never ELs, as they start to be at risk of not graduating with attendance less than 95 percent, instead of 90 percent, and their graduation rates decline more sharply as attendance declines.

College attainment outcomes are also lower for active-ELs in high school than for other students who have similar high school achievement, and background characteristics such as economic status matter more. This highlights that supports for college planning, applications, and navigating financial aid are particularly needed for students with high achievement in these groups to improve their college attainment rates.

Beyond general strategies for all students, these results suggest particular implications for each subgroup:

Former ELs: The strong performance of former ELs often goes unrecognized–in Chicago they have higher educational attainment outcomes than never ELs. If properly supported, they have a high probability of leaving school multilingual with high levels of achievement. If more former ELs earned GPAs of 3.0 or higher they would likely further increase their college degree attainment rates. Over a third of former ELs have ninth grade GPAs between 2.0 and 3.0 which means they are engaged in school, but not at the level where they are likely to attain a college degree.

Late-arriving ELs: Late-arriving ELs have strong academic performance in high school in Chicago, which suggests the language supports provided in Chicago are effective for the average late-arriving EL to fully engage in school. Yet, the grades and attendance of late-arriving ELs seem to be more in-flux between ninth and tenth grade than typical for other groups of students, suggesting a need to monitor and check-in with late-arriving ELs through their first two

years of high school, not just their ninth-grade year, as they adjust to their new high school environment. In addition, strong attendance seems to be more critical for active ELs than for former and never ELs. When it comes to college, barriers beyond high school achievement play a larger role for late-arriving ELs than for former and never ELs, which could include knowledge about navigating the post-secondary search process, access to financial aid, and a lack of documented proficiency in English (Rodriguez & Cruz, 2009).

*LTELs without IEPs***:** Many LTELs without IEPs are successful in high school; almost half had at least a "B" average and passed all of their classes in ninth grade. However, they are at elevated risk of struggling; about 40 percent of LTELs without IEPs in our sample needed more support in the transition to high school to pass their classes and eventually graduate. It is not their level of English proficiency that predicts whether or not they graduate high school and attain a college degree, but success in their classes. A number of studies have shown that LTELs are more likely to be chronically absent and disengaged from school (Clark, et al., 2020); supports around their grades and attendance are critically needed. This could include the use of EWI systems to keep students from falling behind, tutoring that is directly connected to their coursework, assistance with any barriers they may face coming to school or doing homework, and making sure the school context is inviting for LTELs. Prior case studies suggest the label of long-term EL itself can have a negative effect on expectations for students and their academic mindsets, leading to a lack of motivation (Menken et al., 2012; Clark-Gerexa et al., 2020).

LTELs with IEPs: Educational outcomes are harder to predict for this group, although improving attendance is particularly important. They were equally at risk of low outcomes as other groups if they had low attendance, and 33 percent were chronically absent. One issue for this group is that high grades do not ensure that they will graduate from high school. The fact that achievement is less predictive of attainment outcomes for this group means that both school

staff and parents may have difficulty knowing how much support students in this group need to be likely to graduate. For students with IEPs, assessing student performance on multiple indicators–including standardized tests, grades, and attendance together, can help school practitioners further narrow down who is at risk of not graduating, and the level of support they may need. This may require more complex indicators for this specific group of students.

In addition, there is a need to better understand the school experiences of LTELs with IEPs to understand why so many students with IEPs have low attendance rates, relative to other ELs, and why they are still at risk of not graduating from high school even if they have high grades. This group of students is also very unlikely to attain a college credential, even if they have strong grades in high school. Again, this suggests more research is needed to understand their experiences preparing for life after high school, and their college and workforce outcomes.

References

Adair, J. K. (2015). *The impact of discrimination on the early schooling experiences of children from immigrant families*. Migration Policy Institute.

Allensworth, E. M., & Easton, J. Q. (2007). What Matters for Staying On-Track and Graduating in Chicago Public High Schools: A Close Look at Course Grades, Failures, and Attendance in the Freshman Year. Research Report. *Consortium on Chicago School Research*.

Baker, S., Lesaux, N., Jayanthi, M., Dimino, J., Proctor, C. P., Morris, J., ... & Newman-Gonchar, R. (2014). Teaching academic content and literacy to English learners in elementary and middle school. IES Practice Guide. NCEE 2014-4012. *What Works Clearinghouse*.

Balfanz, R., & Byrnes, V. (2019). Early warning indicators and intervention systems: State of the field. *Handbook of Student Engagement Interventions*, 45-55. https://pathwaystoadultsuccess.org/wp-content/uploads/2021/03/EWSStateoftheField.pdf

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies, 81*(2), 608-650.

Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line: Completing college at America's public universities*. Princeton University Press.

Bowers, A. J., Sprott, R., & Taff, S. A. (2013). Do we know who will drop out? A review of the predictors of dropping out of high school: precision, sensitivity, and specificity. *The High School Journal, 96*(2), 77-100.

Bowers, A. J. (2010). Grades and graduation: A longitudinal risk perspective to identify student dropouts. *The Journal of Educational Research*, *103*(3), 191-207.Callahan, R. M., & Shifrer, D. (2016). Equitable access for secondary English learner students: Course taking

as evidence of EL program effectiveness. *Educational Administration Quarterly, 52*(3), 463-496. https://doi.org/10.1177/0013161X16648190

Camara, W. J., & Echternacht, G. (2000). The SAT [R] I and high school grades: Utility in predicting success in college. Research Notes.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal, 21*(1), C1–C68.

Clark-Gareca, B., Short, D., Lukes, M., & Sharp-Ross, M. (2020). Long-term English learners: Current research, policy, and practice. *TESOL Journal*, *11*(1), e00452.

Davis, M. H., Mac Iver, M. A., Balfanz, R. W., Stein, M. L., & Fox, J. H. (2019). Implementation of an early warning indicator and intervention system. *Preventing School Failure: Alternative Education for Children and Youth*, *63*(1), 77-88.

de la Torre, M., Blanchard, A., Allensworth, E. M., & Freire, S. (2019). *English learners in CPS: A new perspective*. University of Chicago Consortium on School Research.

Faria, A. M., Sorensen, N., Heppen, J., Bowdon, J., Taylor, S., Eisner, R., & Foster, S. (2017). Getting students on track for graduation: Impacts of the Early Warning Intervention and Monitoring System after one year. REL 2017-272. *Regional Educational Laboratory Midwest*.

Freeman, Y., Freeman, D., & Mercuri, S. (2003). Helping middle and high school age English language learners achieve academic success. *NABE Journal of Research and Practice*, *1*(1), 110-122.

Geiser, S., & Santelices, M. V. (2007). Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators

of four-year college outcomes. University of California Berkeley. https://cshe.berkeley.edu/sites/default/files/publications/rops.geiser._sat_6.13.07.pdf.

Gwynne, J., Pareja, A. S., Ehrlich, S. B., & Allensworth, E. (2012). *What matters for staying on-track and graduating in Chicago Public Schools: A focus on English language learners*. University of Chicago Consortium on Chicago School Research.

Gurantz, O., & Borsato, G. N. (2012). Building and implementing a college readiness indicator system: Lessons from the first two years of the CRIS Initiative. *Voices in Urban Education, 35*, 5-15.

Huang, M., Haas, E., Zhu, N., & Tran, L. (2016). *High school graduation rates across English learner student subgroups in Arizona* (REL 2017–205). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West.

Hartman, J., Wilkins, C., Gregory, L., Gould, L. F., & D'Souza, S. (2011). Applying an on-track indicator for high school graduation: Adapting the Consortium on Chicago School Research indicator for five Texas districts. Issues & Answers. REL 2011-No. 100. *Regional Educational Laboratory Southwest (NJ1)*.

Jackson, C. K., Kiguel, S., Porter, S. C., & Easton, J. Q. (2024). Who benefits from attending effective high schools? *Journal of Labor Economics*, *42*(3), 717-751.

Kanno, Y., & Cromley, J. (2013). English language learners' access to and attainment in postsecondary education. *TESOL Quarterly, 47*(1), 89–121.

Kieffer, M. J., & Marinell, W. H. (2012). *Navigating the middle grades: Evidence from New York City*. Research Alliance for New York City Schools.

Kieffer, M. J., & Parker, C. E. (2017). *Graduation outcomes of students who entered New York City public schools in grade 5 or 6 as English learner students* (REL 2017–237). U.S.

Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands.

Le, B., Black, K. E., Carlson, C., Miciak, J., Romano, L., Francis, D., & Kieffer, M. J. (2024). Ever English learner 4-year graduation: Toward an intersectional approach. *Educational Researcher, 0*(0). https://doi.org/10.3102/0013189X241246747

McFadden, D. (1977). *Quantitative methods for analyzing travel behavior of individuals: Some recent developments*. Cowles Foundation Discussion Papers. 707. https://elischolar.library.yale.edu/cowles-discussion-paper-series/707

Menken, K., Kleyn, T., & Chae, N. (2012). Spotlight on "long-term English language learners": Characteristics and prior schooling experiences of an invisible population. *International Multilingual Research Journal*, *6*(2), 121-142.

Menken, K. (2013). Emergent bilingual students in secondary school: Along the academic language and literacy continuum. *Language Teaching, 46(4)*, 438-476.

Menken, K., Kleyn, T., & Chae, N. (2012). Spotlight on "long-term English language learners": Characteristics and prior schooling experiences of an invisible population. *International Multilingual Research Journal*, *6*(2), 121–142.

National Center for Education Statistics. (2020). *The Condition of Education.* U.S. Department of Education. https://nces.ed.gov/pubs2020/2020144.pdf.

Norbury, H., Wong, M., Wan, Y., Reese, K., Dhillon, S., & Gerdeman, R. D. (2012). Using the freshman on-track indicator to predict graduation in two urban districts in the midwest region. Issues & Answers. REL 2012-No. 134. *Regional Educational Laboratory Midwest*.

Roderick, M., Kelley-Kemple, T., Johnson, D. W., & Beechum, N. O. (2014). *Preventable failure: Improvements in long-term outcomes when high schools focused on the ninth-grade year. Research Summary*. University of Chicago Consortium on Chicago School Research.

Rodriguez, G. M., & Cruz, L. (2009). The transition to college of English learner and undocumented immigrant students: Resource and policy implications. *Teachers College Record*, *111*(10), 2385-2418.

Saunders, W. M., & Marcelletti, D. J. (2013). The gap that can't go away: The catch-22 of reclassification in monitoring the progress of English Learners. *Educational Evaluation and Policy Analysis, 35*(2), 139–156.

Shi, Q., & Watkinson, J. (2019). English Language Learners and school belonging: Implications for school counselor practice. *Professional School Counseling, 22*(1b), 1-9.

Short, D. J., & Boyson, B. A. (2012). *Helping newcomer students succeed in secondary schools and beyond*. Center for Applied Linguistics.

Stephens, A., & Francis, D. (Eds.). (2018). *English learners in STEM subjects: Transforming classrooms, schools, and lives*. National Academies Press.

Stuit, D., Lindsay, J., & Loney, E. (2016). Identifying early warning indicators in three Ohio school districts. Stated Briefly. REL 2016-207. *Regional Educational Laboratory Midwest*.

Sugarman, J., & Geary, C. (2018). *English learners in Illinois: Demographics, outcomes, and state accountability policies.* Migration Policy Institute.

Umansky, I. M., Thompson, K. D., & Díaz, G. (2017). Using an ever-English learner framework to examine disproportionality in special education. *Exceptional Children, 84*(1)*,* 76-96.

Umansky, I. M., Hopkins, M., Dabach, D. B., Porter, L., Thompson, K., & Pompa, D. (2018). *Understanding and supporting the educational needs of recently arrived immigrant English learner students: Lessons for state and local education agencies.* Council of Chief State School Officers.

Tierney, W. G., Bailey, T., Constantine, J., Finkelstein, N., & Hurd, N. F. (2009). Helping Students Navigate the Path to College: What High Schools Can Do. IES Practice Guide. NCEE 2009-4066. *What Works Clearinghouse*.

U.S. Department of Education. (n.d.a) *Academic performance and outcomes for English Learners: Performance on national assessments and on-time graduation rates*. Retrieved from https://www2.ed.gov/datastory/el-outcomes/index.html#two.

Zau, A., & Betts, J. R. (2008). *Predicting success, preventing failure: An investigation of the California high school exit exam*. Public Policy Institute of California.

Tables

Table 1

Demographic Characteristics of Analytic Sample

Characteristic	All	Never ELs	Former ELs	Active ELs		
				LTELs Without IEPs	LTELs With IEPs	Late-arriving ELs
Number	185,441	117,814	53,371	4,658	4,809	4,789
Male	48%	47%	49%	52%	63%	54%
FRPL eligible	82%	77%	91%	96%	96%	87%
Race						
Latino	46%	25%	84%	94%	92%	56%
Black	37%	58%	1%	1%	1%	6%
White	10%	12%	7%	2%	4%	11%
Asian	5%	3%	7%	3%	3%	24%
Other	2%	3%	1%	<0.5%	1%	3%
IEP	14%	14%	7%		100%	6%
Learning disability	9%	9%	5%		75%	4%
Cognitive disability	3%	3%	1%		19%	1%
Physical disability	1%	1%	<0.5%		4%	1%
Behavioral disability	1%	1%	<0.5%		2%	<0.5%
Speech and language disability	<0.5%	<0.5%	<0.5%		<0.5%	<0.5%
New to CPS in ninth grade	4%	5%	0%		0%	22%

Table 2

High School Indicators Used to Predict Attainment Outcomes

Indicator	Data	Categories for Dummy Variable Groups
Standardized test scores	8th grade: ISAT before 2015, PARCC in 2015 and later 9th grade: EXPLORE before 2015, PSAT in 2015 and later 10th grade: PLAN before 2015, PSAT in 2015 and later 11th grade: ACT in 2015 and 2016, SAT in 2017 and later	Five categories created from the mean and standard deviation of all students in sample who took the exams: Very High: 1.5+ standard deviations above the mean; High: 0.5 to 1.5 standard deviations above the mean; mis: 0.5 standard deviations above to 0.5 standard deviations below the mean; Low: 0.5 to 1.5 standard deviations below the mean; Very Low: 1.5+ standard deviations below the mean; students with missing data
Grade point average	Unweighted GPA based on grades in all classes for 8th, 9th, 10th, 11th, and 12th grades	4.00-3.01, 3.00-2.01, 2.00-1.01, 1.00-0.00, missing
Course failures	Number of semester courses in which a student received an F for 9th, 10th, 11th, and 12th grades	0, 1, 2-3, 4+, missing
Attendance	Days present divided by days enrolled for 8th-12th grades	98%-100%, 95%-98%, 90%-95%, 85%-90%, below 85%, missing
Days suspended	Indicators based on the number of out-of-school suspension days in 9th, 10th, 11th, and 12th grades	0,1-2, 3-5, 6+
ACCESS proficiency level	For active ELs for 9th-12th grades only. Proficiency levels go from 1 to 6 on the ACCESS test	1.0-1.9, 2.0-2.9, 3.0-3.9, 4.0-4.9, 5.0-6.0, missing
ACCESS growth	Difference in a student's ACCESS score relative to the prior year for active ELs for 9th, 10th, 11th, and 12th grades.	Five groups created from the mean and standard deviation of all students in sample who took the exam: Very High, High, Mid, Low, Very Low, and missing (see standardized test scores row for rules for each category)
Reached proficiency	Student reached proficiency on the ACCESS test so no longer considered an EL in 9th-12th grades	1=yes, 0=no

Table 3

Pseudo R^2 Statistics From Single-Indicator Models Predicting High School Graduation and College Outcomes

Indicator		High School Graduation					College Enrollment					College Degree Completion				
		Former ELs	Late-Arriving ELs	LTELs Without IEPs	LTELs With IEPs	All Never ELs	Former ELs	Late-Arriving ELs	LTELs Without IEPs	LTELs With IEPs	All Never ELs	Former ELs	Late-Arriving ELs	LTELs Without IEPs	LTELs With IEPs	All Never ELs
Background		0.065	0.064	0.038	0.016	0.094	0.065	0.094	0.051	0.027	0.116	0.090	0.149	0.072	0.025	0.169
Test scores																
	8th grade - English	0.058	0.030	0.021	0.017	0.074	0.062	0.045	0.021	0.019	0.105	0.088	0.076	0.033	0.053	0.149
	8th grade - math	0.074	0.050	0.020	0.010	0.085	0.074	0.059	0.025	0.012	0.109	0.109	0.111	0.050	0.068	0.162
	9th grade - composite	0.086	0.048	0.033	0.021	0.100	0.086	0.065	0.031	0.023	0.128	0.118	0.122	0.071	0.084	0.194
	10th grade - composite	0.085	0.070	0.033	0.015	0.094	0.087	0.066	0.031	0.017	0.122	0.128	0.091	0.043	0.056	0.196
GPA	8th	0.141	0.090	0.114	0.007	0.129	0.090	0.077	0.079	0.017	0.119	0.109	0.114	0.065	0.013	0.167
	9th	0.272	0.194	0.260	0.086	0.236	0.122	0.105	0.107	0.033	0.147	0.171	0.177	0.133	0.059	0.207
	10th	0.320	0.273	0.308	0.107	0.270	0.136	0.127	0.138	0.043	0.161	0.188	0.223	0.162	0.068	0.208
Attendance	8th	0.104	0.062	0.083	0.055	0.105	0.045	0.026	0.040	0.029	0.055	0.049	0.043	0.026	0.031	0.058
	9th	0.235	0.158	0.205	0.077	0.214	0.112	0.087	0.100	0.068	0.143	0.121	0.079	0.067	0.067	0.173
	10th	0.265	0.203	0.238	0.084	0.223	0.126	0.096	0.116	0.068	0.150	0.135	0.110	0.102	0.063	0.175
Course failures	9th	0.252	0.165	0.245	0.072	0.230	0.104	0.072	0.104	0.040	0.121	0.109	0.096	0.083	0.025	0.129
	10th	0.286	0.228	0.283	0.074	0.249	0.113	0.079	0.114	0.043	0.130	0.126	0.130	0.108	0.056	0.153
Days suspended	9th	0.050	0.025	0.031	0.016	0.074	0.019	0.011	0.014	0.009	0.044	0.019	0.017	0.019	0.003	0.056
	10th	0.040	0.016	0.026	0.009	0.045	0.018	0.008	0.019	0.008	0.034	0.018	0.001	0.015	0.000	0.048
ACCESS score	8th		0.025	0.016	0.055			0.033	0.024	0.044			0.053	0.010	0.058	
	9th		0.047	0.032	0.078			0.055	0.035	0.060			0.065	0.040	0.074	
	10th		0.045	0.033	0.052			0.066	0.030	0.043			0.081	0.026	0.057	
	11th		0.042	0.007	0.045			0.087	0.008	0.024			0.089	0.085	0.021	
	12th		0.029	0.008	0.026			0.050	0.033	0.030			0.086	0.056	0.031	
ACCESS growth	8th-9th		0.006	0.003	0.005			0.001	0.001	0.003			0.023	0.009	0.006	
	9th - 10th		0.022	0.016	0.007			0.017	0.014	0.011			0.040	0.020	0.033	
	10th - 11th		0.017	0.003	0.010			0.020	0.006	0.013			0.056	0.019	0.033	
	11th - 12th		0.004	0.004	0.001			0.020	0.002	0.014			0.061	0.014	0.030	
Reached proficiency	9th		0.022	0.016	0.013			0.016	0.014	0.014			0.035	0.022	0.031	
	10th		0.023	0.005	0.008			0.015	0.003	0.006			0.039	0.007	0.021	
	11th		0.003	0.000	0.003			0.009	0.000	0.001			0.037	0.025	0.004	
	12th		0.009	0.000	0.000			0.004	0.002	0.001			0.053	0.023	0.002	

Note. LTEL stands for "long-term English learner." Background characteristics include gender, race/ethnicity, free or reduced-price lunch eligibility, special education status, social status, concentration of poverty, age at entry to high school, and school mobility in middle grades.

Table 4

Models With Combinations of Predictors for High School Graduation and College Outcomes: Pseudo-R2 Statistics

Ninth Grade Indicator(s)	High School Graduation					College Enrollment					College Degree Completion				
	Former ELs	Late-Arriving ELs	LTEL No IEPs	LTELs With IEPs	All Never ELs	Former ELs	Late-Arriving ELs	LTELs No IEPs	LTELs With IEPs	All Never ELs	Former ELs	Late-Arriving ELs	LTELs No IEPs	LTEL With IEPs	All Never ELs
GPA	0.279	0.197	0.264	0.088	0.239	0.114	0.107	0.111	0.029	0.148	0.158	0.177	0.133	0.058	0.207
Course failures	0.258	0.166	0.251	0.065	0.232	0.101	0.075	0.108	0.035	0.121	0.107	0.096	0.083	0.025	0.127
Attendance rate	0.240	0.161	0.213	0.088	0.215	0.109	0.087	0.104	0.077	0.142	0.108	0.079	0.067	0.065	0.173
Stand. test score	0.084	0.052	0.036	0.021	0.100	0.079	0.068	0.036	0.024	0.126	0.102	0.122	0.071	0.084	0.184
Days suspended	0.046	0.024	0.032	0.012	0.074	0.016	0.011	0.015	0.008	0.043	0.017	0.017	0.019	0.002	0.053
ACCESS level		0.047	0.032	0.129			0.055	0.035	0.077			0.065	0.040	0.074	
ACCESS growth		0.006	0.003	0.013			0.001	0.001	0.005			0.023	0.009	0.006	
GPA + failures	0.300	0.207	0.291	0.100	0.259	0.124	0.111	0.127	0.041	0.157	0.162	0.179	0.135	0.050	0.209
GPA + attendance	0.337	0.240	0.327	0.135	0.297	0.148	0.134	0.148	0.079	0.189	0.179	0.185	0.147	0.087	0.247
GPA + test score	0.287	0.213	0.249	0.154	0.270	0.140	0.139	0.119	0.064	0.191	0.189	0.217	0.165	0.152	0.267
GPA + suspensions	0.293	0.194	0.269	0.092	0.256	0.123	0.116	0.116	0.039	0.160	0.159	0.177	0.136	0.059	0.214
GPA + A. proficiency		0.223	0.252	0.200			0.138	0.115	0.096			0.197	0.141	0.137	
GPA + ACCESS growth		0.212	0.263	0.133			0.104	0.106	0.051			0.181	0.135	0.069	
Failures + attendance	0.325	0.220	0.318	0.102	0.293	0.142	0.112	0.147	0.081	0.175	0.148	0.119	0.106	0.060	0.206
Failures + test score	0.265	0.193	0.240	0.163	0.267	0.132	0.115	0.115	0.069	0.182	0.162	0.175	0.129	0.117	0.237
Failures + suspensions	0.267	0.163	0.254	0.064	0.248	0.102	0.073	0.112	0.040	0.132	0.110	0.098	0.088	0.027	0.144
Failures + A. proficiency		0.200	0.253	0.206			0.114	0.121	0.109			0.136	0.099	0.096	
Failures + ACESS growth		0.182	0.250	0.138			0.072	0.104	0.056			0.112	0.073	0.033	
Attendance + test score	0.236	0.173	0.194	0.137	0.227	0.133	0.120	0.114	0.085	0.179	0.157	0.170	0.121	0.129	0.238
Attendance + suspensions	0.244	0.158	0.210	0.078	0.224	0.108	0.087	0.098	0.073	0.146	0.109	0.082	0.072	0.068	0.178
Attendance + proficiency		0.183	0.211	0.190			0.127	0.121	0.126			0.129	0.095	0.119	

Ninth Grade Indicator(s)	High School Graduation					College Enrollment					College Degree Completion				
	Former ELs	Late-Arriving ELs	LTEL No IEPs	LTELs With IEPs	All Never ELs	Former ELs	Late-Arriving ELs	LTELs No IEPs	LTELs With IEPs	All Never ELs	Former ELs	Late-Arriving ELs	LTELs No IEPs	LTEL With IEPs	All Never ELs
Attendance + A. growth		0.162	0.204	0.133			0.086	0.100	0.078			0.094	0.062	0.071	
Test score + suspensions	0.105	0.067	0.053	0.037	0.141	0.082	0.072	0.040	0.030	0.138	0.109	0.128	0.072	0.087	0.197
Test score + proficiency		0.059	0.048	0.029			0.094	0.045	0.040			0.125	0.090	0.098	
Test score + A. growth		0.054	0.038	0.022			0.067	0.030	0.022			0.157	0.053	0.087	
Suspensions + proficiency		0.068	0.058	0.111			0.060	0.047	0.073			0.078	0.043	0.074	
Suspensions + A. growth		0.027	0.033	0.032			0.011	0.015	0.011			0.036	0.020	0.004	
A. proficiency + growth		0.049	0.032	0.066			0.057	0.034	0.057			0.097	0.040	0.069	
All ninth grade indicators	0.341	0.265	0.306	0.221	0.327	0.163	0.201	0.154	0.110	0.214	0.205	0.238	0.149	0.188	0.284
All ninth grade indicators + background characteristics	0.351	0.297	0.323	0.245	0.338	0.180	0.283	0.204	0.129	0.228	0.229	0.315	0.205	0.214	0.302

Note. Indicators were entered non-parametrically as a series of dummy variables allowing for non-linear relationships. The models at the bottom of the table include all ninth-grade indicators displayed above. The final model also includes variables for gender, race/ethnicity, free/reduced-price lunch eligibility, special education status, social status, concentration of poverty, crime rate, age, and mobility in the middle grades. ACCESS proficiency and growth are sometimes abbreviated in the table as proficiency or growth because of space constraints.

Table 5

Models Predicting Attainment With and Without School Fixed Effects and Selected Ninth Grade Achievement Indicators

Ninth Grade Indicators	High School Graduation				College Enrollment				College Degree Completion			
	No Fixed Effects		Fixed Effects		No Fixed Effects		Fixed Effects		No Fixed Effects		Fixed Effects	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Never EL	2.449***	0.0163	3.331***	0.113	0.112***	0.0084	-0.104	0.0770	-1.376***	0.0174	-1.459***	0.188
Attendance	0.595***	0.0148	0.593***	0.0165	0.552***	0.0156	0.435***	0.0161	0.539***	0.0335	0.234***	0.0335
GPA	1.187***	0.0182	1.249***	0.0194	0.635***	0.0111	0.740***	0.0119	0.931***	0.0202	1.173***	0.0223
Test score	0.256***	0.0160	0.122***	0.0193	0.543***	0.0101	0.331***	0.0122	0.736***	0.0164	0.295***	0.0205
Former EL	2.335***	0.0309	3.351***	0.118	0.142***	0.0180	-0.265***	0.0798	-1.328***	0.0376	-1.356***	0.193
Attendance	0.807***	0.0386	0.802***	0.0397	0.763***	0.0355	0.696***	0.0359	0.517***	0.0700	0.275***	0.0702
GPA	1.297***	0.0334	1.313***	0.0342	0.546***	0.0190	0.601***	0.0194	0.881***	0.0381	1.076***	0.0396
Test score	0.263***	0.0354	0.162***	0.0369	0.535***	0.0201	0.359***	0.0210	0.655***	0.0370	0.325***	0.0389
Late-arriving EL	1.986***	0.109	3.041***	0.159	-0.251***	0.0826	-0.411***	0.115	-1.184***	0.181	-1.174***	0.264
Attendance	0.587***	0.105	0.603***	0.106	0.725***	0.113	0.696***	0.115	0.403	0.284	0.260	0.281
GPA	1.084***	0.0848	1.097***	0.0856	0.596***	0.0600	0.646***	0.0604	0.994***	0.141	1.090***	0.144
Test score	0.344***	0.100	0.282***	0.102	0.620***	0.0772	0.484***	0.0785	1.025***	0.146	0.876***	0.150
LTEL without IEP	2.245***	0.132	3.315	0.176	-0.368***	0.0881	-0.419***	0.119	-1.406***	0.173	-1.313***	0.256
Attendance	0.868***	0.0941	0.881***	0.0950	0.989***	0.111	0.933***	0.112	0.349	0.233	0.157	0.232
GPA	1.407***	0.0854	1.427***	0.0866	0.636***	0.0582	0.705***	0.0587	0.875***	0.160	1.012***	0.161
Test score	0.154	0.119	0.0780	0.121	0.526***	0.0911	0.380***	0.0933	0.727***	0.200	0.494***	0.195
LT EL with IEP	2.336***	0.174	3.414***	0.210	-0.364***	0.133	-0.438***	0.154	-1.275***	0.429	-1.271***	0.468
Attendance	0.771***	0.0824	0.777***	0.0840	0.878***	0.122	0.823***	0.121	0.498*	0.275	0.394	0.273
GPA	0.821***	0.0891	0.845***	0.0909	0.409***	0.0642	0.460***	0.0646	0.903***	0.203	1.097***	0.373
Test score	0.515***	0.132	0.486***	0.134	0.593***	0.107	0.532***	0.107	0.1229***	0.374	1.097***	0.373

Note. * p<0.10, ** p<0.05, *** p<0.01. Predictors come from students' ninth grade year, standardized and entered as continuous variables.

Main effects for each EL group represent a student with average values on attendance, GPA, and test score in each group.

Table 6

High School Graduation Rates for Long-Term ELs Without IEPs and Late-Arriving ELs

Ninth Grade Attendance Rate	Ninth Grade GPA			
	A-B (3.01-4.00)	B-C (2.01-3.00)	C-D (1.01-2.00)	D-F (0.00-1.00)
98% - 100%	98%	97%	87%	* 49%
95% - <98%	95%	91%	83%	40%
90% - <95%	93%	86%	69%	48%
85% - <90%	92%	83%	64%	29%
<85%	* 87%	73%	46%	14%

Note. Cells with an asterisk (*) are based on fewer than 50 students. The top category is set at 98% attendance because of the large number of students with greater than 95% attendance.

Table 7

High School Graduation Rates for Long-Term ELs With IEPs

Ninth Grade ACCESS Level	Ninth Grade GPA			
	A-B (3.01-4.00)	B-C (2.01-3.00)	C-D (1.01-2.00)	D-F (0.00-1.00)
5.1-6.0	* 100%	95%	86%	* 36%
4.1-5.0	97%	92%	78%	35%
3.1-4.0	89%	88%	70%	21%
2.1-3.0	60%	80%	66%	26%
1.0-2.0	10%	42%	67%	* 6%

Note. Cells with an asterisk (*) are based on fewer than 50 students.

Table 8

College Degree Completion Rates for Select Groups of ELs

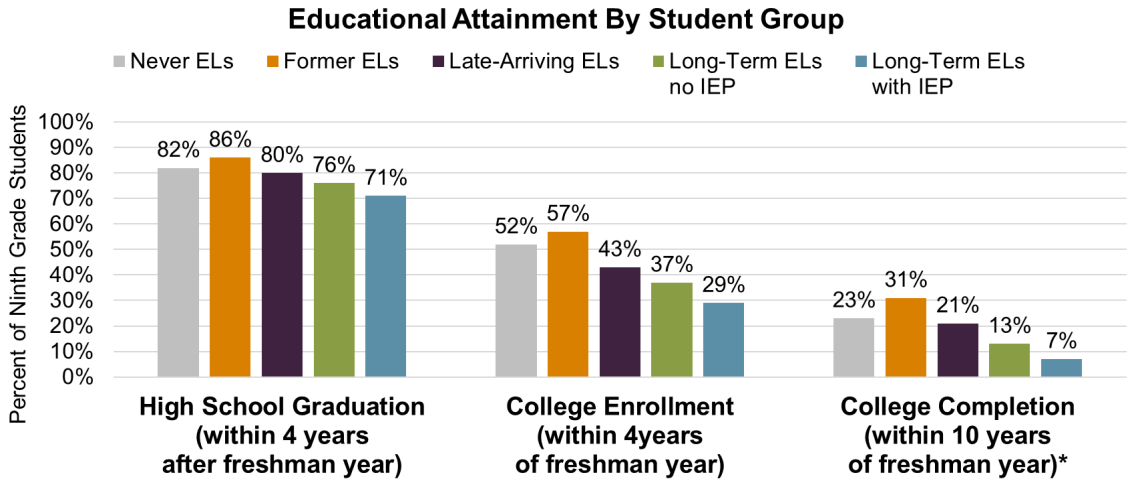
Ninth Grade Test Score	Ninth Grade GPA – Late-Arriving ELs				Ninth Grade GPA – LTELs With IEPs			
	A-B (3.01-4.00)	B-C (2.01-3.00)	C-D (1.01-2.00)	D-F (0.00-1.00)	A-B (3.01-4.00)	B-C (2.01-3.00)	C-D (1.01-2.00)	D-F (0.00-1.00)
Very high	* 100%	N/A	N/A	N/A	N/A	N/A	N/A	N/A
High	* 65%	* 0%	* 0%	* 0%	* 100%	N/A	* 0%	N/A
Mid	54%	27%	* 21%	* 0%	* 50%	* 32%	* 22%	* 0%
Low	33%	12%	5%	* 3%	17%	7%	4%	* 0%
Very low	* 13%	5%	0%	* 0%	* 14%	4%	1%	* 3%

Note. Cells with an asterisk (*) are based on fewer than 50 students.

Figures

Figure 1

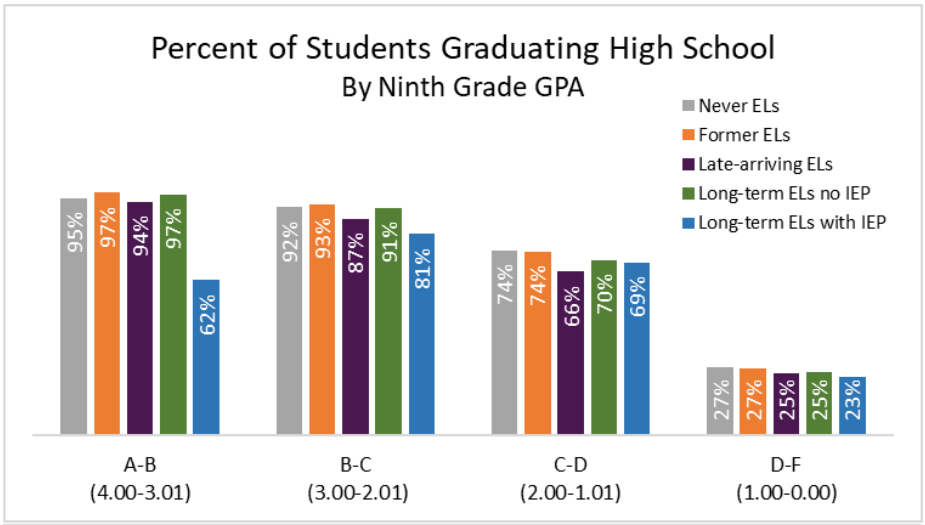
Educational Attainment Among Different Groups of Students

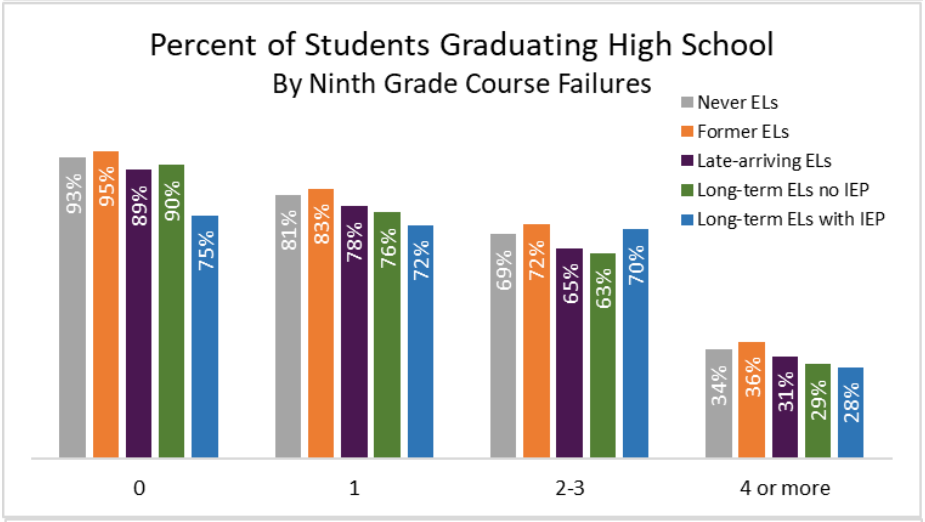


Note. The figure is based on cohorts beginning high school from the fall of 2008 to the fall of 2017. Only the earliest four cohorts are included for college completion, as the recent cohorts have not been out of high school long enough to complete college in 10 years.

Figure 2

Four-Year High School Graduation Rates by Ninth Grade GPAs and Attendance





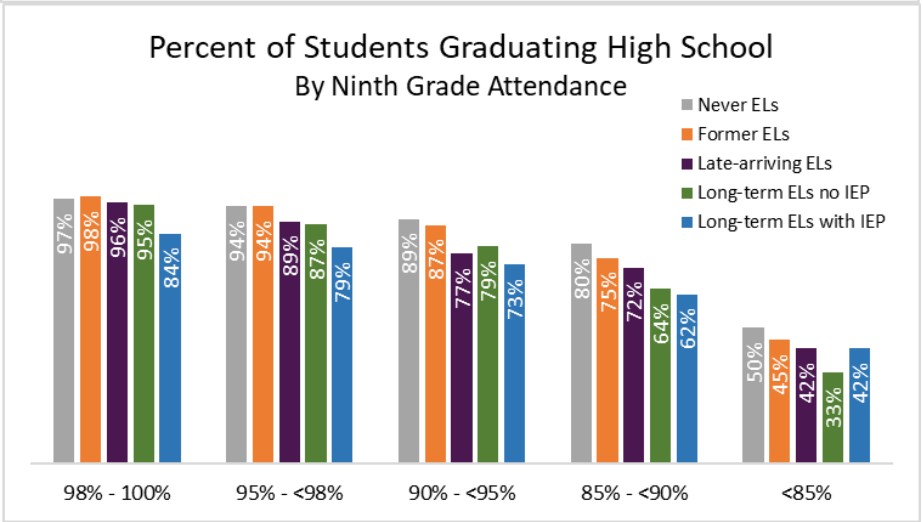
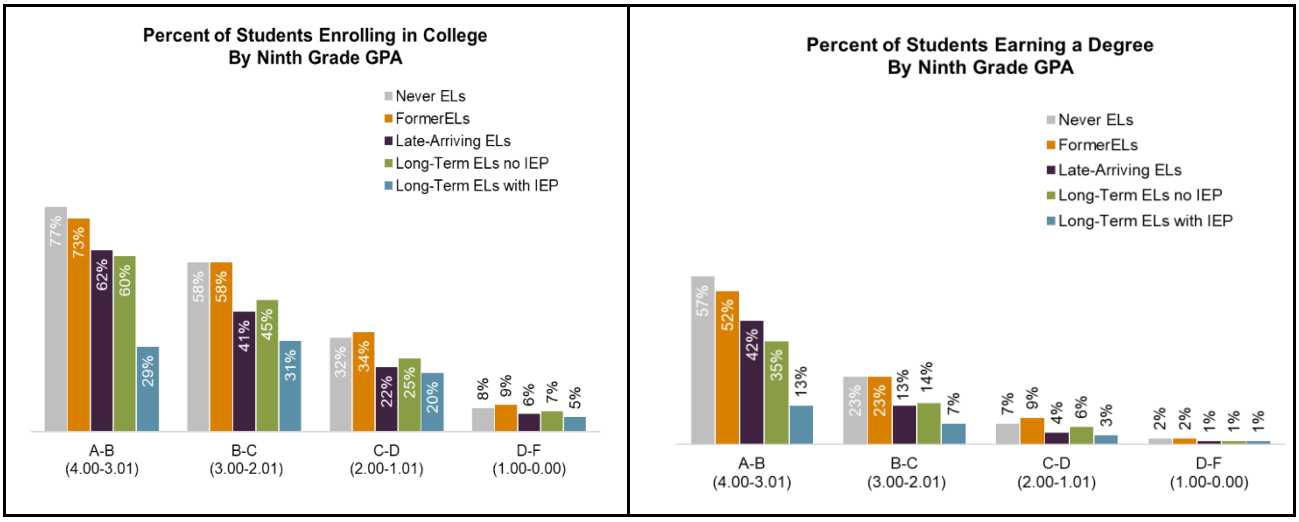
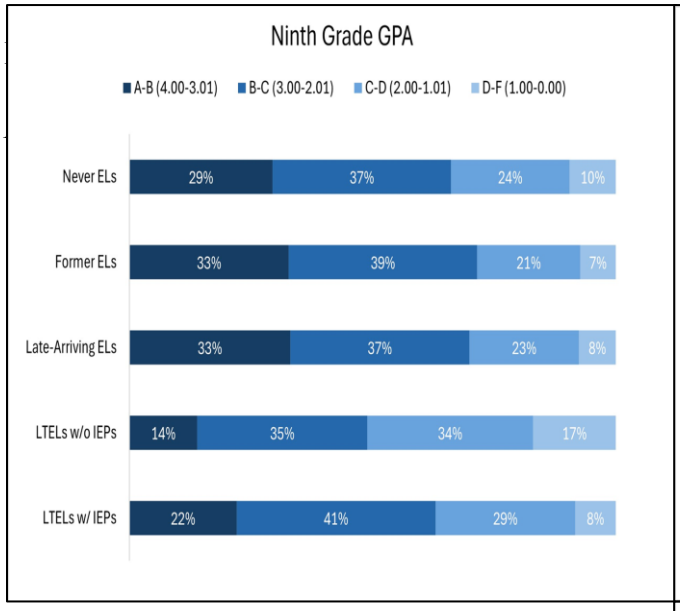


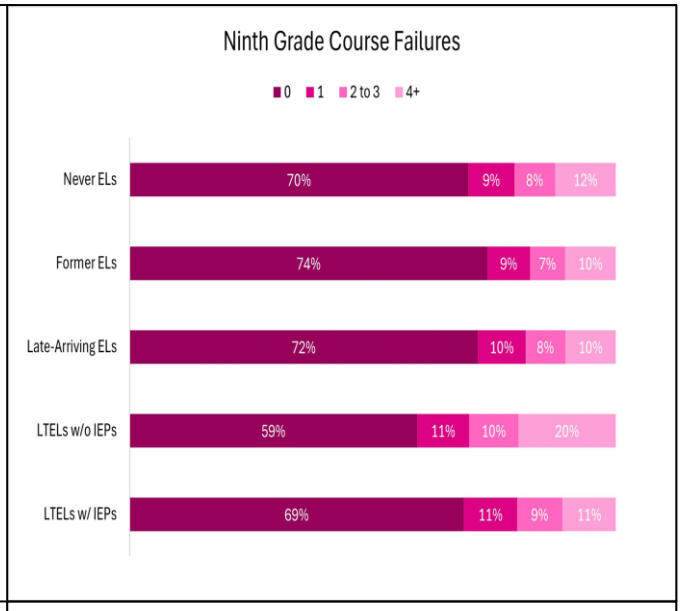
Figure 3

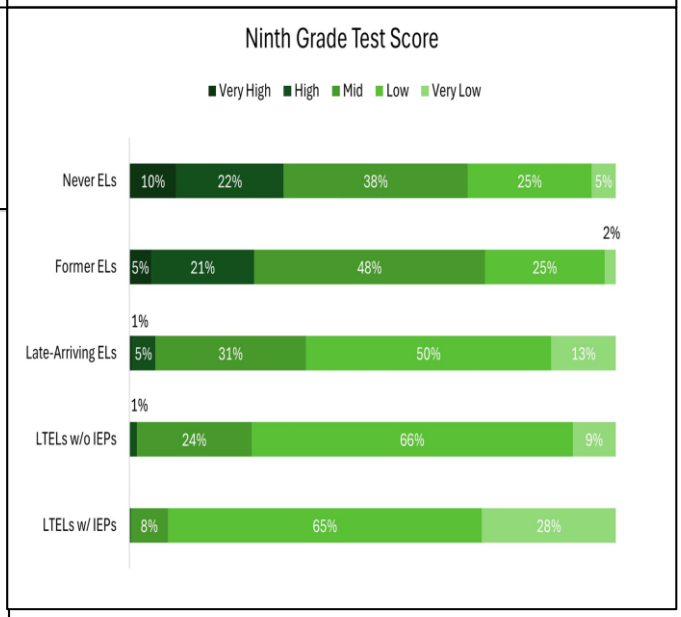
College Enrollment and College Degree Completion by Ninth Grade GPA

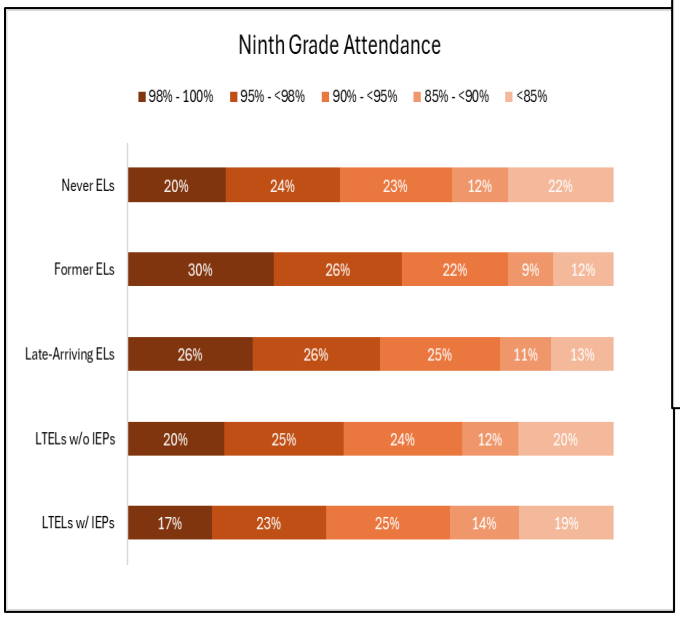


Note. Percentages for college enrollment are based on all students who enrolled in ninth grade, followed for four years after beginning ninth grade. Percentages for degree attainment are based on the cohorts that started ninth grade in the fall of 2008 through the fall of 2011, followed for ten years after beginning ninth grade.


Ninth Grade GPA
A-B (4.00-3.01)
B-C (3.00-2.01)
C-D (2.00-1.01)
D-F (1.00-0.00)
Never ELs
29%
37%
24%
10%
Former ELs
33%
39%
21%
7%
Late-Arriving ELs
33%
37%
23%
8%
LTELs w/o IEPs
14%
35%
34%
17%
LTELs w/ IEPs
22%
41%
29%
8%


Ninth Grade Course Failures
0
1
2 to 3
4+
Never ELs
70%
9%
8%
12%
Former ELs
74%
9%
7%
10%
Late-Arriving ELs
72%
10%
8%
10%
LTELs w/o IEPs
59%
11%
10%
20%
LTELs w/ IEPs
69%
11%
9%
11%


Ninth Grade Test Score
Very High
High
Mid
Low
Very Low
Never ELs
10%
22%
38%
25%
5%
Former ELs
5%
21%
48%
25%
2%
Late-Arriving ELs
5%
31%
50%
13%
1%
LTELs w/o IEPs
24%
66%
9%
1%
LTELs w/ IEPs
8%
65%
28%


Ninth Grade Attendance
98% - 100%
95% - <98%
90% - <95%
85% - <90%
<85%
Never ELs
20%
24%
23%
12%
22%
Former ELs
30%
26%
22%
9%
12%
Late-Arriving ELs
26%
26%
25%
11%
13%
LTELs w/o IEPs
20%
25%
24%
12%
20%
LTELs w/ IEPs
17%
23%
25%
14%
19%

Online Appendices

Appendix A: Additional Tables and Figures

Table A1

Pseudo R2 Statistics From Single-Indicator Models including Never-ELs With IEPs

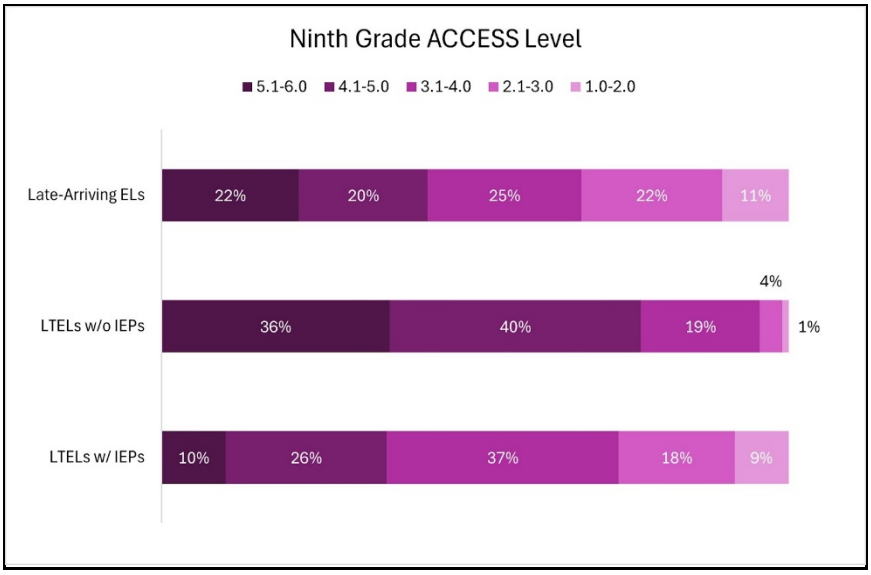
Indicator	High School Graduation				College Enrollment				College Degree Completion			
	LTELs Without IEPs	LTELs With IEPs	Never ELs Without IEPs	All Never ELs	LTELs Without IEPs	LTELs With IEPs	Never ELs Without IEPs	All Never ELs	LTELs Without IEPs	LTELs With IEPs	Never ELs Without IEPs	All Never ELs
Background characteristics	0.038	0.016	0.025	0.094	0.051	0.027	0.057	0.116	0.072	0.025	0.124	0.169
Standardized test score												
8th grade - English	0.021	0.017	0.021	0.074	0.021	0.019	0.063	0.105	0.033	0.053	0.107	0.149
8th grade - math	0.020	0.010	0.019	0.085	0.025	0.012	0.054	0.109	0.050	0.068	0.121	0.162
9th grade	0.033	0.021	0.030	0.100	0.031	0.023	0.079	0.128	0.071	0.084	0.165	0.194
10th grade	0.033	0.015	0.032	0.094	0.031	0.017	0.080	0.122	0.043	0.056	0.172	0.196
11th grade	0.027	0.021	0.031	0.079	0.034	0.032	0.090	0.123	0.038	0.043	0.180	0.204
GPA												
8th grade	0.114	0.007	0.028	0.129	0.079	0.017	0.042	0.119	0.065	0.013	0.107	0.167
9th grade	0.260	0.086	0.113	0.236	0.107	0.033	0.064	0.147	0.133	0.059	0.120	0.207
10th grade	0.308	0.107	0.129	0.270	0.138	0.043	0.070	0.161	0.162	0.068	0.106	0.208
11th grade	0.291	0.101	0.129	0.246	0.131	0.046	0.067	0.153	0.153	0.031	0.083	0.174

Indicator	High School Graduation				College Enrollment				College Degree Completion			
	LTELs Without IEPs	LTELs With IEPs	Never ELs Without IEPs	All Never ELs	LTELs Without IEPs	LTELs With IEPs	Never ELs Without IEPs	All Never ELs	LTELs Without IEPs	LTELs With IEPs	Never ELs Without IEPs	All Never ELs
12th grade	0.247	0.111	0.109	0.215	0.108	0.043	0.054	0.115	0.124	0.045	0.070	0.130
Attendance rate												
8th grade	0.083	0.055	0.055	0.105	0.040	0.029	0.034	0.055	0.026	0.031	0.043	0.058
9th grade	0.205	0.077	0.092	0.214	0.100	0.068	0.090	0.143	0.067	0.067	0.138	0.173
10th grade	0.238	0.084	0.085	0.223	0.116	0.068	0.086	0.150	0.102	0.063	0.121	0.175
11th grade	0.207	0.067	0.062	0.192	0.112	0.068	0.081	0.138	0.116	0.051	0.107	0.161
12th grade	0.189	0.044	0.050	0.137	0.099	0.054	0.055	0.092	0.056	0.057	0.061	0.093
Courses failed												
9th grade	0.245	0.072	0.105	0.230	0.104	0.040	0.066	0.121	0.083	0.025	0.087	0.129
10th grade	0.283	0.074	0.097	0.249	0.114	0.043	0.062	0.130	0.108	0.056	0.091	0.153
11th grade	0.227	0.037	0.059	0.200	0.082	0.017	0.050	0.108	0.102	0.060	0.074	0.129
12th grade	0.175	0.045	0.046	0.130	0.051	0.028	0.031	0.051	0.075	0.024	0.041	0.076
Days suspended												
9th grade	0.031	0.016	0.033	0.074	0.014	0.009	0.025	0.044	0.019	0.003	0.047	0.056
10th grade	0.026	0.009	0.015	0.045	0.019	0.008	0.018	0.034	0.015	0.000	0.031	0.048
11th grade	0.011	0.003	0.002	0.013	0.006	0.005	0.010	0.017	0.005	0.000	0.020	0.036
12th grade	0.000	0.000	0.000	0.001	0.002	0.000	0.003	0.005	0.000	0.000	0.016	0.019
ACCESS proficiency level												

Indicator	High School Graduation				College Enrollment				College Degree Completion			
	LTELs Without IEPs	LTELs With IEPs	Never ELs Without IEPs	All Never ELs	LTELs Without IEPs	LTELs With IEPs	Never ELs Without IEPs	All Never ELs	LTELs Without IEPs	LTELs With IEPs	Never ELs Without IEPs	All Never ELs
8th grade	0.016	0.055			0.024	0.044			0.010	0.058		
9th grade	0.032	0.078			0.035	0.060			0.040	0.074		
10th grade	0.033	0.052			0.030	0.043			0.026	0.057		
11th grade	0.007	0.045			0.008	0.024			0.085	0.021		
12th grade	0.008	0.026			0.033	0.030			0.056	0.031		
ACCESS score growth												
8th to 9th grade	0.003	0.005			0.001	0.003			0.009	0.006		
9th to 10th grade	0.016	0.007			0.014	0.011			0.020	0.033		
10th to 11th grade	0.003	0.010			0.006	0.013			0.019	0.033		
11th to 12th grade	0.004	0.001			0.002	0.014			0.014	0.030		
ACCESS proficiency												
9th grade	0.016	0.013			0.014	0.014			0.022	0.031		
10th grade	0.005	0.008			0.003	0.006			0.007	0.021		
11th grade	0.000	0.003			0.000	0.001			0.025	0.004		
12th grade	0.000	0.000			0.002	0.001			0.023	0.002		

Figure A1

ACCESS Proficiency Levels by EL Subgroup



Note: The categories for the ninth-grade test score variable were created based on the mean and standard deviation of all students in the sample who took either the EXPLORE or the PSAT in ninth grade. "Very High" scores are at or above 1.5 standard deviations above the mean score, "High" scores are between 1.5 and 0.5 standard deviations above the mean, "Mid" scores are at or between 0.5 standard deviations above the mean and 0.5 standard deviations below the mean, "Low" scores are between 0.5 and 1.5 standard deviations below the mean, and "Very Low" scores are at or below 1.5 standard deviations below the mean.

Table A2

Pseudo R2 Statistics From Single-Indicator Models for 2010 and 2016 Cohorts

Ninth Grade Indicator(s)	2010 Cohort										2016 Cohort				
	High School Graduation					College Degree Attainment					High School Graduation				
	Former ELs	Late-Arriving Els	LTELs Without IEPs	LTELs With IEPs	Never ELs	Former ELs	Late-Arriving Els	LTELs Without IEPs	LTELs With IEPs	Never ELs	Former ELs	Late-Arriving Els	LTELs Without IEPs	LTELs With IEPs	Never ELs
GPA	0.287	0.272	0.208	0.097	0.261	0.170	0.119	0.110	0.002	0.195	0.286	0.136	0.287	0.086	0.201
Courses failed	0.256	0.223	0.225	0.116	0.230	0.115	0.053	0.038	0.003	0.122	0.262	0.120	0.248	0.045	0.205
Attendance rate	0.199	0.189	0.251	0.057	0.193	0.133	0.088	0.087	0.009	0.181	0.233	0.104	0.204	0.129	0.192
Standardized test score	0.103	0.030	0.039	0.088	0.127	0.149	0.035	0.077	0.014	0.186					
Days suspended	0.038	0.021	0.066	0.015	0.055	0.019	0.008	0.000	0.000	0.037	0.025	0.013	0.004	0.014	0.034
ACCESS proficiency level		0.054	0.062	0.105			0.070	0.106	0.028			0.078	0.058	0.060	
ACCESS growth		0.037	0.035	0.021			0.019	0.028	0.027			0.008	0.019	0.026	
GPA + failures	0.297	0.284	0.243	0.133	0.271	0.174	0.110	0.144	0.004	0.198	0.309	0.151	0.302	0.094	0.225
GPA + attendance	0.317	0.297	0.306	0.119	0.299	0.199	0.130	0.129	0.012	0.247	0.342	0.184	0.334	0.187	0.260
GPA + test score	0.308	0.282	0.220	0.196	0.296	0.231	0.131	0.146	0.016	0.272					
GPA + suspensions	0.289	0.274	0.225	0.100	0.266	0.172	0.119	0.094	0.003	0.201	0.290	0.141	0.288	0.091	0.206
GPA + ACCESS level		0.284	0.211	0.223			0.143	0.149	0.030			0.205	0.273	0.142	
GPA + ACCESS growth		0.291	0.192	0.154			0.135	0.063	0.027			0.173	0.282	0.138	
Failures + attendance	0.304	0.270	0.326	0.131	0.282	0.172	0.089	0.096	0.011	0.211	0.334	0.179	0.309	0.133	0.263
Failures + test score	0.287	0.242	0.235	0.214	0.284	0.209	0.074	0.102	0.017	0.243					
Failures + suspensions	0.260	0.227	0.252	0.117	0.239	0.119	0.054	0.029	0.004	0.134	0.271	0.127	0.251	0.051	0.210
Failures + ACCESS level		0.241	0.235	0.254			0.098	0.120	0.034			0.195	0.248	0.139	

Ninth Grade Indicator(s)	2010 Cohort										2016 Cohort				
	High School Graduation					College Degree Attainment					High School Graduation				
	Former ELs	Late-Arriving Els	LTELs Without IEPs	LTELs With IEPs	Never ELs	Former ELs	Late-Arriving Els	LTELs Without IEPs	LTELs With IEPs	Never ELs	Former ELs	Late-Arriving Els	LTELs Without IEPs	LTELs With IEPs	Never ELs
Failures + ACCESS growth		0.251	0.249	0.191			0.069	0.058	0.043			0.140	0.244	0.124	
Attendance + test score	0.229	0.210	0.261	0.133	0.226	0.209	0.115	0.138	0.015	0.254					
Attendance + suspensions	0.201	0.190	0.261	0.059	0.198	0.134	0.088	0.074	0.013	0.182	0.237	0.107	0.204	0.130	0.194
Attendance + ACCESS level		0.219	0.269	0.171			0.136	0.158	0.026			0.159	0.224	0.210	
Attendance + ACCESS growth		0.208	0.268	0.102			0.102	0.087	0.035			0.092	0.207	0.198	
Test score + suspensions	0.126	0.051	0.092	0.098	0.148	0.159	0.043	0.072	0.013	0.195					
Test score + ACCESS level		0.066	0.079	0.127			0.088	0.143	0.028						
Test score + ACCESS growth		0.068	0.068	0.062			0.063	0.035	0.030						
Suspensions + ACCESS level		0.065	0.106	0.089			0.069	0.130	0.051			0.077	0.055	0.070	
Suspensions + ACCESS growth		0.051	0.081	0.053			0.019	0.029	0.011			0.019	0.019	0.033	
ACCESS level + ACCESS growth		0.074	0.060	0.039			0.091	0.099	0.030			0.059	0.053	0.048	
All ninth grade indicators	0.338	0.341	0.352	0.288	0.326	0.251	0.196	0.167	0.138	0.296	0.358	0.258	0.341	0.259	0.272
All ninth grade indicators + background characteristics	0.358	0.391	0.391	0.386	0.341	0.278	0.325	0.288	0.338	0.316	0.377	0.319	0.395	0.306	0.303

Note. There is no ninth grade test score data for the 2016 cohort because neither the EXPLORE nor the PSAT were administered in 2016.

Table A3

Pseudo R2 Statistics From Single-Indicator Models Including Students With Missing Data on Indicators

Indicator	High School Graduation					College Enrollment					College Degree Completion				
	Former ELs	Late-Arriving Els	LTELs Without IEPs	LTELs With IEPs	All Never ELs	Former ELs	Late-Arriving ELs	LTELs Without IEPs	LTELs With IEPs	All Never ELs	Former ELs	Late-Arriving ELs	LTELs Without IEPs	LTELs With IEPs	All Never ELs
Background characteristics	0.065	0.064	0.038	0.016	0.094	0.065	0.094	0.051	0.027	0.116	0.090	0.149	0.072	0.025	0.169
Standardized test score															
8th grade - English	0.062	0.017	0.019	0.146	0.065	0.061	0.025	0.020	0.049	0.092	0.087	0.047	0.037	0.070	0.136
8th grade - math	0.077	0.037	0.019	0.138	0.074	0.072	0.042	0.024	0.042	0.095	0.107	0.083	0.053	0.076	0.147
9th grade - composite	0.082	0.040	0.031	0.092	0.101	0.074	0.046	0.023	0.028	0.110	0.118	0.102	0.063	0.096	0.200
10th grade - composite	0.094	0.060	0.044	0.096	0.107	0.063	0.042	0.019	0.016	0.091	0.152	0.085	0.071	0.079	0.231
11th grade - composite	0.263	0.204	0.211	0.234	0.283	0.118	0.083	0.072	0.060	0.159	0.190	0.171	0.103	0.110	0.276
GPA															
8th grade	0.124	0.060	0.100	0.007	0.108	0.080	0.052	0.069	0.016	0.102	0.093	0.066	0.052	0.011	0.139
9th grade	0.273	0.194	0.260	0.087	0.237	0.123	0.105	0.107	0.033	0.147	0.171	0.177	0.133	0.059	0.207
10th grade	0.339	0.298	0.319	0.128	0.297	0.144	0.137	0.143	0.050	0.174	0.194	0.230	0.162	0.068	0.217
11th grade	0.391	0.386	0.369	0.181	0.347	0.159	0.163	0.167	0.076	0.202	0.178	0.180	0.176	0.031	0.208
12th grade	0.440	0.477	0.411	0.229	0.398	0.167	0.165	0.177	0.091	0.207	0.163	0.181	0.175	0.045	0.192
Attendance rate															
8th grade	0.103	0.048	0.081	0.054	0.100	0.044	0.020	0.038	0.029	0.056	0.047	0.040	0.029	0.031	0.065
9th grade	0.235	0.158	0.205	0.077	0.215	0.112	0.087	0.100	0.068	0.143	0.121	0.079	0.067	0.067	0.173
10th grade	0.274	0.219	0.242	0.092	0.234	0.129	0.102	0.118	0.068	0.155	0.137	0.111	0.102	0.063	0.178
11th grade	0.317	0.298	0.265	0.118	0.261	0.144	0.127	0.131	0.068	0.162	0.147	0.113	0.116	0.051	0.176
12th grade	0.322	0.322	0.307	0.143	0.281	0.125	0.105	0.134	0.078	0.141	0.112	0.114	0.087	0.057	0.128
Courses failed															
9th grade	0.252	0.165	0.245	0.072	0.231	0.104	0.072	0.104	0.040	0.122	0.109	0.096	0.083	0.025	0.129
10th grade	0.310	0.260	0.297	0.100	0.281	0.122	0.092	0.121	0.050	0.146	0.134	0.139	0.108	0.056	0.163

Indicator	High School Graduation					College Enrollment					College Degree Completion				
	Former ELs	Late-Arriving Els	LTELs Without IEPs	LTELs With IEPs	All Never ELs	Former ELs	Late-Arriving ELs	LTELs Without IEPs	LTELs With IEPs	All Never ELs	Former ELs	Late-Arriving ELs	LTELs Without IEPs	LTELs With IEPs	All Never ELs
11th grade	0.359	0.353	0.322	0.135	0.321	0.132	0.106	0.123	0.052	0.164	0.143	0.146	0.127	0.060	0.166
12th grade	0.403	0.415	0.374	0.184	0.360	0.127	0.104	0.129	0.079	0.156	0.124	0.122	0.131	0.024	0.144
Days suspended															
9th grade	0.046	0.022	0.027	0.014	0.066	0.020	0.015	0.014	0.009	0.040	0.019	0.017	0.019	0.003	0.056
10th grade	0.034	0.015	0.023	0.008	0.041	0.015	0.007	0.018	0.009	0.030	0.021	0.004	0.015	0.000	0.052
11th grade	0.032	0.042	0.040	0.016	0.041	0.011	0.013	0.019	0.010	0.022	0.010	0.004	0.006	0.000	0.024
12th grade	0.056	0.102	0.084	0.041	0.079	0.016	0.026	0.032	0.017	0.023	0.005	0.003	0.028	0.003	0.015
ACCESS proficiency level															
8th grade		0.025	0.016	0.055			0.033	0.024	0.044			0.053	0.010	0.058	
9th grade		0.042	0.027	0.062			0.050	0.029	0.048			0.066	0.040	0.070	
10th grade		0.027	0.016	0.040			0.045	0.019	0.028			0.068	0.018	0.043	
11th grade		0.018	0.001	0.038			0.042	0.006	0.011			0.064	0.025	0.013	
12th grade		0.020	0.003	0.033			0.023	0.007	0.011			0.050	0.018	0.017	
ACCESS score growth															
8th to 9th grade		0.006	0.003	0.103			0.001	0.001	0.034			0.019	0.014	0.007	
9th to 10th grade		0.012	0.010	0.087			0.006	0.006	0.011			0.027	0.011	0.004	
10th to 11th grade		0.024	0.022	0.134			0.004	0.002	0.013			0.032	0.012	0.002	
11th to 12th grade		0.052	0.027	0.121			0.001	0.002	0.009			0.030	0.007	0.002	
ACCESS proficiency															
9th grade		0.022	0.016	0.013			0.016	0.014	0.014			0.035	0.022	0.031	
10th grade		0.023	0.005	0.008			0.015	0.003	0.006			0.039	0.007	0.021	
11th grade		0.003	0.000	0.003			0.009	0.000	0.001			0.037	0.025	0.004	
12th grade		0.009	0.000	0.000			0.004	0.002	0.001			0.053	0.023	0.002	

Appendix B: Lasso results

We employed machine learning techniques for all cohorts to ensure the results presented in the main paper were reliable and consistent across different methodologies. Specifically, we used the Least Absolute Shrinkage and Selection Operator (Lasso) method. This approach minimizes the residual sum of squares (similar to OLS) while incorporating a penalty term to enhance variable selection and prevent overfitting (Belloni et al., 2014; Chernozhukov et al., 2018).

Lasso regression offers several advantages over traditional methods, particularly in high-dimensional settings where there are many covariates. Unlike standard OLS, which includes all predictors, Lasso automatically selects relevant variables by shrinking some coefficients to zero, effectively performing feature selection. This improves model interpretability and prevents overfitting, especially when dealing with correlated variables. Moreover, Lasso enhances computational efficiency by reducing the number of variables in the model without requiring manual selection. These statistical and computational benefits make Lasso a useful robustness check, complementing traditional regression techniques.

This can be mathematically expressed as follows, where the objective function balances goodness-of-fit with a penalty term that encourages sparsity in the model:

$$\text{minimize} \left(\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right)$$

where the penalty term is equal to the sum of the absolute values of the non-intercept coefficients multiplied by the regularization parameter λ. When λ is small, the Lasso solution approaches the ordinary least squares (OLS) solution, as the penalty term has negligible effect on coefficient shrinkage. However, as λ increases, more coefficients are shrunk towards zero, with some

potentially becoming exactly zero, effectively removing those covariates from the model and avoiding overfitting, especially when dealing with a large number of predictors. This method produces sparse models with fewer covariates by selecting only the most important predictors and eliminating the less relevant ones. Therefore, we use the models to identify the most important and relevant factors that predict high school graduation, immediate college enrollment, and college degree attainment, and compare the results with the more traditional techniques presented in the main text.

The optimal value of λ was determined through cross-validation to minimize prediction error. We used a five-fold cross-validation, splitting the data into 5 equal parts with the model estimated with 80% of the data and then cross-validating on the remaining 20% of the data. This process is repeated for each fold, and the average deviance across all folds becomes the cross-validation function. The deviance measures the average discrepancy between the predicted probabilities and the actual outcomes in the held-out validation set.

The figures in the remainder of this Appendix provide the coefficient paths for one of the groups (former ELs), which illustrate how the estimated parameters change as the regularization parameter λ varies. These paths show how different covariates enter or exit the model as λ changes, with smaller values of λ including more variables (similar to OLS) and larger values shrinking coefficients toward zero, leading to fewer variables in the model. The value under the 'one standard error rule' represents the largest λ whose cross-validated error is within one standard error of the minimum, offering a solution that balances prediction accuracy and model complexity, making it robust to noise in the data. For example, in Figure B.1, which predicts high school graduation for Former ELs, the dashed vertical line corresponds to the standard-error-rule λ_{SE}=0.038, where only three variables are retained: GPA, attendance rate and course

failures in ninth grade. This suggests that those three variables are the strongest predictors of high school graduation for this subgroup, while other variables, such as demographic and socioeconomic factors, are excluded as λ increases. Similar figures are available for each of the subgroups from the authors.

There are two differences with respect to the results we presented in the main part of the paper. First, instead of entering the potential achievement indicators as a series of categorical variables, they were entered as continuous variables–otherwise LASSO would treat the different categories as different variables. Second, a smaller set of demographic variables were included in the LASSO regressions because each was entered separately, and many were redundant with others in their prediction of attainment outcomes. The variables included in the LASSO models were ninth grade standardized test scores, ninth grade GPA, ninth grade attendance rate, number of days suspended in ninth grade, number of courses failed in ninth grade, whether a male student, free or reduced-price lunch eligible, Latinx student, Black student, white student, Asian student, and whether identified for special education services. For active EL students, we also included ACCESS English language proficiency score in grade 9, ACCESS scores growth from grades 8 to 9, and whether the student reached English proficiency in grade 9.

In general, the LASSO models identified the same indicators as most important for predicting the educational attainment outcomes as we found with the more traditional methods, with a few differences described below. Table B1 shows the relevant ninth grade indicators identified for each outcome and each group.

For predicting high school graduation for most groups of students, course failures, GPA and attendance are the variables selected through the LASSO models, which is similar to the results presented with the exploratory analyses. A difference here is that course failures showed

up as an important predictor even after including GPA. We wonder if this occurred because the variables are entered as continuous, while in the main analysis GPA was entered as a series of dummy variables allowing for nonlinear relationships that cannot be reflected in the LASSO model. Figures B.1 through B.3 show the coefficient paths by the different values of λ for each of the outcomes for former ELs and the values for λ_{CV} and λ_{SE}.[2] While course failures, GPA and attendance were the three variables selected for the λ_{SE}, attendance did not remain for late-arriving ELs and standardized test scores were included for long-term ELs with an IEP. Through this method we continued to find that English proficiency indicators were not as relevant for active ELs as grades and attendance.

When it comes to key indicators for college enrollment, GPA is the most predictive variable for all students. This is the only indicator that is selected by the model for long-term English learners. For other groups of students, standardized test scores are also selected, and attendance is also relevant for former and never ELs. Some demographic variables, such as gender, race, and free or reduced-price lunch status, improve the prediction of college enrollment.

GPA and standardized test scores are the model-selected indicators for predicting college degree attainment for all groups of students except for long-term English learners. Using LASSO, we do not find any variables that meet the criteria for inclusion for long-term ELs. If any were to be included, GPA is the indicator that is chosen to be first. Some of the demographic variables were also relevant for this outcome for other students.

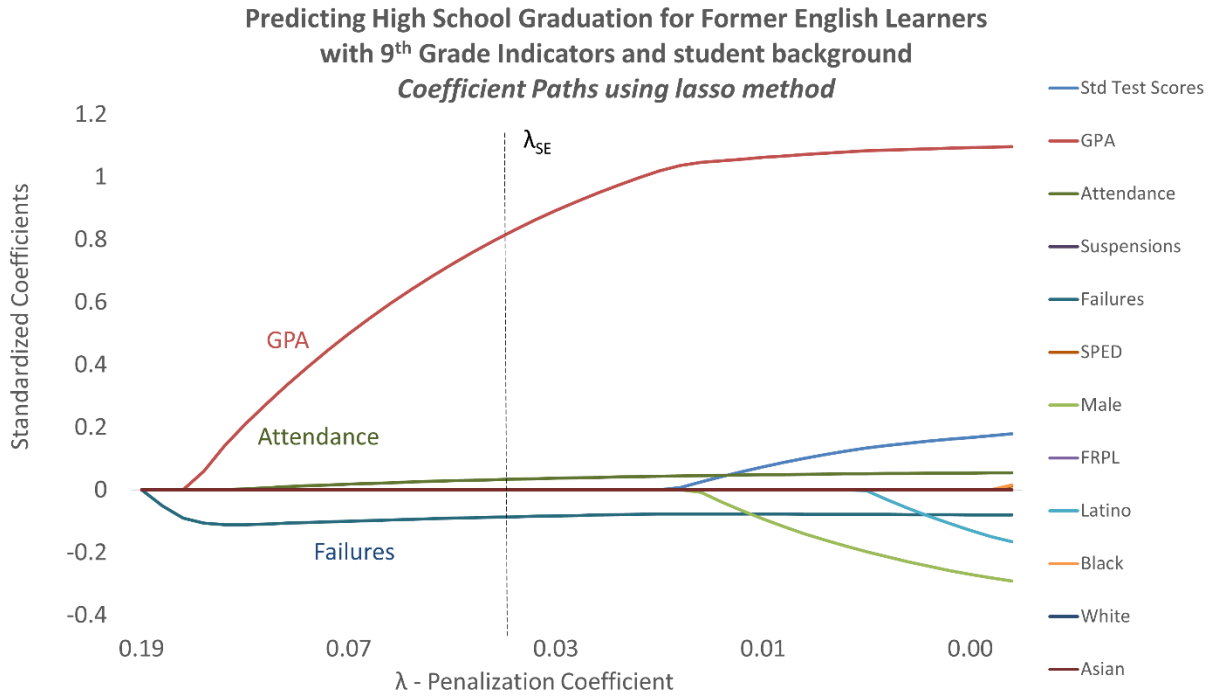
[2] Figures for other groups of students are available upon request from the authors.

Table B1: Key indicators identified using LASSO and the "one standard error rule"

	High School Graduation	Immediate College Enrollment	College Degree Attainment
Former ELs	Course failures, GPA & attendance	GPA & standardized test scores & attendance & Latino & male	GPA & standardized test scores & Latino
Late-arriving ELs	Course failures & GPA	GPA & standardized test scores & Latino, Asian & white	GPA & standardized test scores & Latino
LTELs no IEP	Course failures, GPA & attendance	GPA	
LTELs with IEP	Course failures, GPA & attendance	GPA	
Never ELs	Course failures, GPA, attendance & standardized test scores	GPA& standardized test scores, attendance, & course failures & FRPL& Asian& Black & male	GPA& standardized test scores & attendance & FRPL& Asian & Black & male

Figure B1

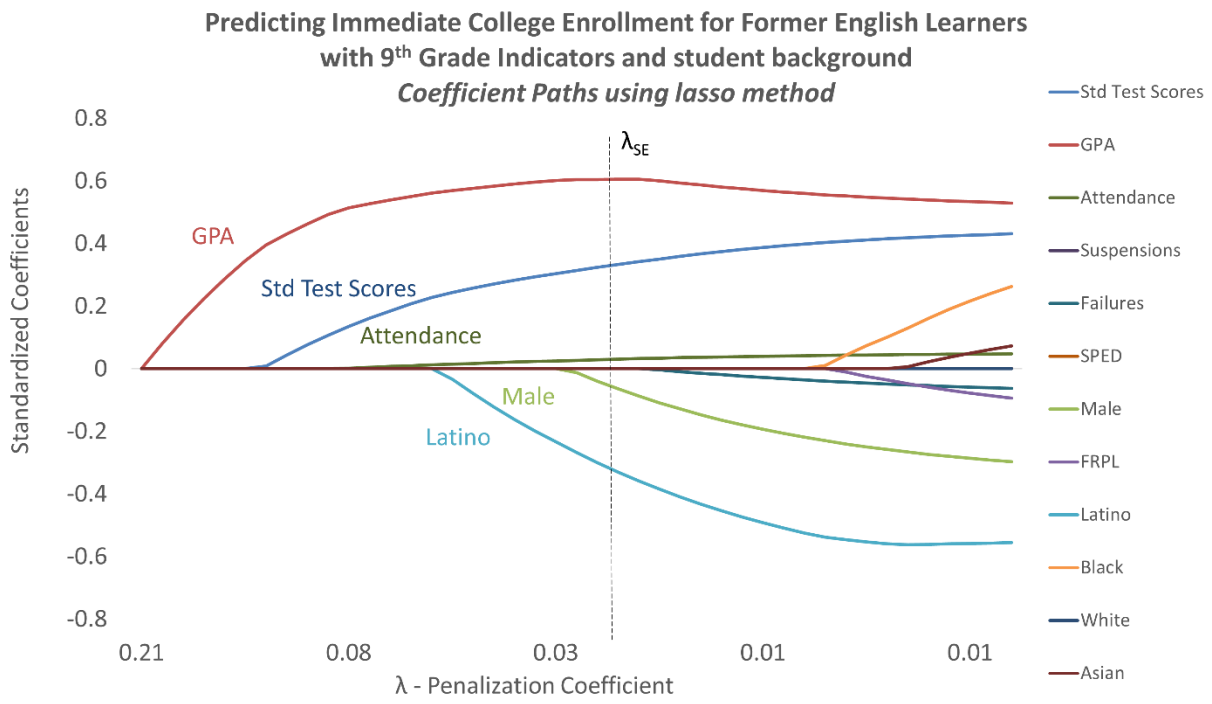
Predicting High School Graduation for Former ELs



Note. $\lambda_{SE} = 0.038$ is the standard-error-rule λ; the number of selected coefficients is 3.

Figure B2

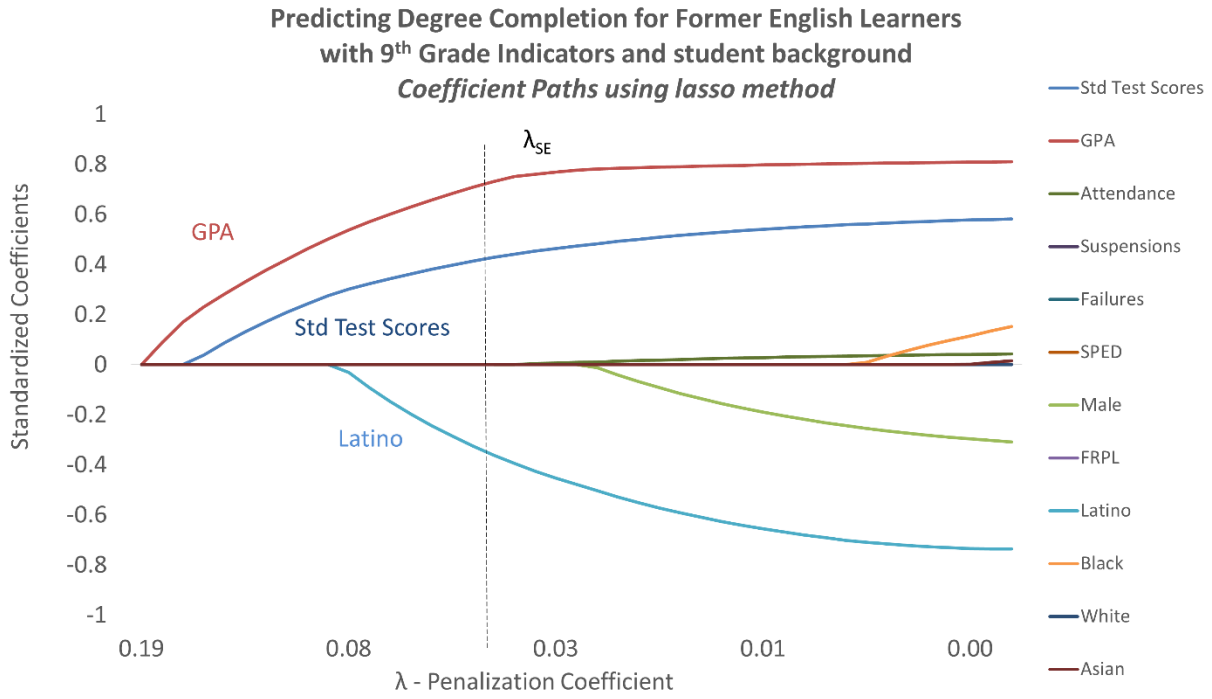
Predicting Immediate College Enrollment for Former ELs



Note. $\lambda_{SE} = 0.025$ is the standard-error-rule λ; the number of selected coefficients is 5.

Figure B3

Predicting Degree Completion for Former ELs



Note. $\lambda_{SE} = 0.043$ is the standard-error-rule λ; the number of selected coefficients is 3.