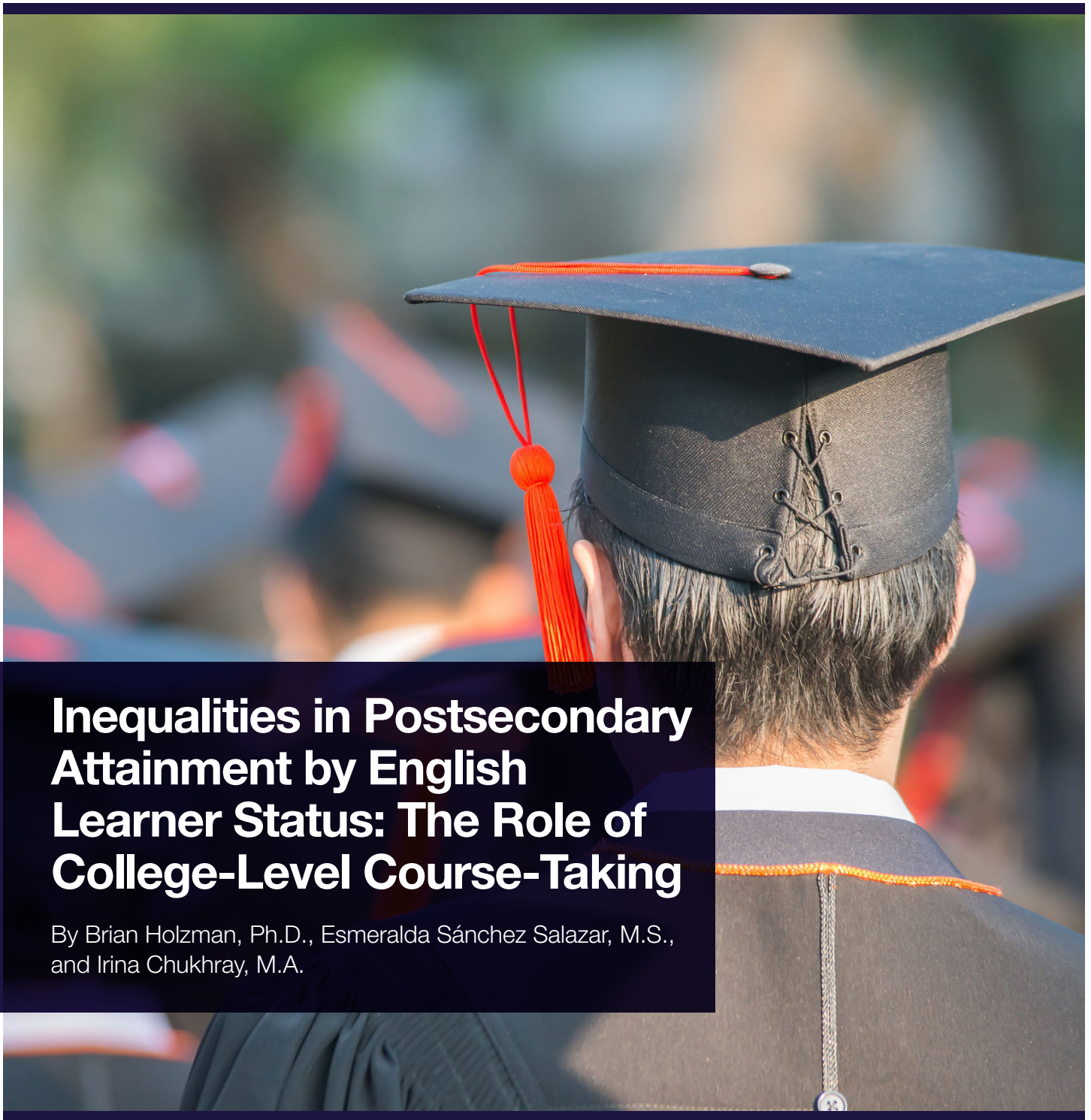# RICE | HERC
Houston Education Research Consortium

Inequalities in Postsecondary Attainment by English Learner Status: The Role of College-Level Course-Taking

By Brian Holzman, Ph.D., Esmeralda Sánchez Salazar, M.S., and Irina Chukhray, M.A.

Abstract

This report examines the role of English learner (ELs) status in four-year college enrollment and bachelor's degree completion among Houston Independent School District (HISD) high school graduates. We divide students into four groups: students who are never classified as EL (hereafter referred to as "Never EL" students), ELs who are reclassified during elementary school (grades kindergarten-five), ELs who are reclassified during middle school (grades six-eight), and ELs who are reclassified during high school (grades nine-11) or still EL as of the fall of 12th grade (hereafter referred to as "reclassified during high school" students). We find that gaps in four-year college outcomes by EL status are large, but are entirely explained by differences in sociodemographic, academic, and school characteristics. Of the academic characteristics we consider, differences in college-level course taking during the junior and senior years of high school explain 7 percent of the gap in four-year college enrollment between Never EL students and students reclassified in elementary school, 18 percent of gap between Never EL students and students reclassified in middle school, and 22 percent of the gap between Never EL students and students reclassified in high school. We also find that differences in college-level course-taking explain 14 percent of the gap in four-year college completion between Never EL students and students reclassified in middle school and 40 percent of the gap between Never EL students and students reclassified in high school.

Note on the authors: Irina Chukhray, M.A. is currently a doctoral student at the University of California - Davis.

Executive Summary

Results

- While gaps in four-year college enrollment and bachelor's degree completion by EL status appear large, they are fully explained by sociodemographic, academic, and school characteristics.

- After controlling for sociodemographic and school characteristics, EL students reclassified in middle and high school take fewer college-level courses during the junior and senior years of high school than Never EL students.

- In contrast, Never EL students and EL students reclassified in elementary school appear to take similar numbers of college-level courses.

- When considering academic characteristics like reading test scores, math test scores, average course grades, and the number of college-level courses taken, differences in college-level course-taking explain 7 to 22 percent of the gap in four-year college enrollment between Never EL students and students reclassified in elementary, middle, and high school.

- In terms of four-year college completion, differences in college-level course-taking explain 14 percent of the gap between Never EL students and students reclassified in middle school and 40 percent of the gap between Never EL students and students reclassified in high school.

Note
Throughout this report, whenever we refer to English Learner (EL) status, we intend to focus on four groups of students:
1. Never EL
2. Former EL reclassified during elementary school (grades kindergarten-five)
3. Former EL reclassified during middle school (grades six-eight)
4. Former EL reclassified during high school (grades nine-12)[1]

---

[1] Students reclassified in grades nine-11 are in the same category as students still classified as EL in grade 12 because the latter category is too small to analyze separately. More substantively, students still classified as EL are in the "reclassified during high school" group because, like students reclassified in grades 9-11, they enter high school as ELs.

Background

According to the Migration Policy Institute, 26 percent of U.S. children under 18 years old are first- or second-generation immigrant.[2] This growing diversity brings linguistic and cultural challenges as well as opportunities for public schools to identify novel strategies to serve ELs. According to the National Center for Education Statistics, it is estimated that 10 percent of U.S. public school students are currently classified as EL.[3]

While the English language instruction that current ELs receive may be valuable, it may inadvertently limit students' access to advanced courses[4]: students may not have time to take advanced courses in addition to their English language courses or they may not be ready to take advanced courses.[5] Given the importance of college preparatory classes for applying to and enrolling in college after graduating from high school, these English language instruction classes, while important, may simultaneously contribute to long-term educational inequalities.

In a previous Houston Education Research Consortium (HERC) report[6], we found HISD students who earned college-level credits during their senior year of high school were more likely to attend and graduate from college. In fact, only 14 percent of HISD students who earned no college-level credits during the senior year of high school obtained any kind of postsecondary credential within six years of high school graduation. If access to college-level high school coursework varies by EL status[7], then that may partly explain disparities in educational attainment.[8]

This study will inform HISD on how EL status affects students after high school graduation. Our deep dive into the factors explaining gaps by EL status can aid the district's ongoing efforts to increase educational opportunities for students who have ever been classified as EL.

2   Migration Policy Institute. (2017). *Children in U.S. Immigrant Families: Number and Share of Total Child Population, by Age Group and State*. Retrieved from http://www.migrationpolicy.org/programs/data-hub/charts/children-immigrant-families.

3   National Center for Education Statistics, U.S. Department of Education. (2018). Table 204.20: Number and percentage of public school students participating in English language learner (EL) programs, by state: Selected years, fall 2000 through fall 2016. In National Center for Education Statistics, U.S. Department of Education, (Ed.), *Digest of Education Statistics* (2018 ed.). Retrieved from https://nces.ed.gov/programs/digest/d18/tables/dt18_204.20.asp.

4   Umansky, I. M. (2016). Leveled and Exclusionary Tracking: English Learners' Access to Academic Content in Middle School. *American Educational Research Journal, 53*(6), 1792–1833.

5   For example, to take geometry in ninth grade, algebra I may need to be taken in eighth grade and pre-algebra may need to be taken in seventh grade.

6   Holzman, B. (2018). *Transitioning to College and Work (Part 1: Where are high school seniors from 2006-2008 now?)* (HERC Research Report, Volume 6, Issue 1). Houston, TX: Houston Education Research Consortium, Kinder Institute for Urban Research, Rice University.

7   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.

8   Kanno, Y., & Cromley, J. G. (2015). English Language Learners' Pathways to Four-Year Colleges. *Teachers College Record, 117*(12), 1–44.

Research Questions

1. What is the role of EL status, specifically the grade level during which a student is reclassified, in predicting four-year college enrollment and completion five years after high school graduation?

2. Is there a relationship between the grade level of reclassification and college-level course-taking during high school?

3. Is any share of the gaps in four-year college enrollment and completion by EL status (i.e., grade level of reclassification) attributable to gaps in college-level course-taking during high school by EL status? If so, how much? Does the explained variance differ across EL group?

Preview of Results

First, we demonstrate that there are large gaps in four-year college outcomes by EL status, with students reclassified in elementary, middle, and high school attending and graduating from college at lower rates than their Never EL peers. However, these gaps are no longer meaningful once we account for sociodemographic, academic, and school characteristics.

Next, we show that the number of college-level courses taken during the junior and senior years of high school are key to understanding why gaps in four-year college outcomes by EL status exist. Even after accounting for background characteristics, students reclassified in middle and high school take fewer college courses than Never EL students and students reclassified in elementary school. These course-taking gaps appear to feed into gaps in four-year college outcomes—differences in the number of college-level courses that Never EL and former EL students take explain large portions of the gaps in four-year college outcomes.

Data And Methods

Source

We use HISD data from the HERC longitudinal database, 1995-2017. The data contain information on students' sociodemographic and academic characteristics and have been matched to postsecondary records from the National Student Clearinghouse, which allow us to measure college attendance and graduation.

Sample

Our sample consists of five cohorts of HISD students who were in the 12th grade during the fall semesters between 2007 and 2011 and who graduated high school during the spring semesters between 2008 and 2012. We restrict the sample to students who have complete data for the covariates included in the statistical models. In the end, our sample includes 31,374 students, approximately three-fourths of high school graduates.[9]

Dependent Variables

For Research Questions #1 and #3, we examine two four-year college outcomes: 1) whether a student enrolled in a four-year college within five years of high school graduation and 2) whether a student completed a bachelor's degree within five years of high school graduation. For this second outcome, we limit the sample to high school graduates who enrolled in a four-year college or university the fall immediately following graduation. Therefore, the sample size is reduced to 13,348 students. For Research Question #2, the outcome measures the number of college-level courses (see next subsection for a definition) that a student takes during the junior and senior years of high school.

While this study focuses on four-year college outcomes, results are also available upon request for models looking at any college enrollment or any degree completion. We do not discuss these outcomes in this report because the mechanism being considered—college-level course-taking—is geared toward four-year college enrollment and completion. While college-level course-taking during high school may have a role in college access and success more broadly, there are important differences in the pathways to and through two-year, technical, and vocational programs that should not be conflated with the processes involved in successfully starting and finishing four-year degrees.

Key Independent Variables

The main variable of interest is EL status, a four-category measure of the grade level during which a student is reclassified:

- Never EL

- Former EL reclassified during elementary school (grades kindergarten-five)

- Former EL reclassified during middle school (grades six-eight)

- Former EL reclassified during high school (grades nine-12)[10]

This variable allows us to calculate raw gaps in four-year college outcomes by EL status and how those gaps change after accounting for sociodemographic, academic, and school characteristics (Research Question #1). We also use this variable to understand gaps in college-level course-taking by EL status (Research Question #2) and how gaps in sociodemographic, academic, and school characteristics

---

9   Of the 41,525 students who graduated from HISD high schools in the spring semesters between 2008 and 2012, we excluded 86 observations missing data on four-year college outcomes. These students could not be located in the National Student Clearinghouse files. We excluded an additional 3,480 who were classified as special education. Given the focus on academic characteristics (reading test scores, math test scores, average course grades, number of college courses taken), we excluded 3,862 students missing test score data or grades and courses data. More students were missing test score data (3,650) than grades and courses data (1,854). The wide majority of students missing test score data or grades and courses data were not in the data files at all; the minority were in the data files and had missing values. Finally, we excluded an additional 2,723 students missing data from the enrollment form. The majority of these students had no information on their household structure (1,441) or the number of years they had lived in the United States (965). Nearly all these students were in the enrollment form data file but were missing values, so their parents either provided no information or inaccurate information on the measures or their addresses could not be geocoded. Additional details on sample selection and missing data are available from the authors upon request.

---

10   Students reclassified in grades nine-11 are in the same category as students still classified as EL in grade 12 because the latter category is too small to analyze separately. More substantively, students still classified as EL in grade 12 are part of the "reclassified during high school" group because, like students reclassified during grades nine-11, they entered high school as ELs.

feed into gaps in four-year college outcomes (Research Question #3). Figure 1 shows how the sample is divided into the four groups. Most students are never classified as EL, while the fewest number of the students are reclassified in high school.

Many students attend other districts before HISD. Because we do not observe their EL status before they come to HISD, we account for the years they are enrolled in HISD schools. Moreover, HISD enrolls a large share of immigrant students. Because we do not know the kind of education students receive before entering HISD, we account for the years they attend HISD schools and the number of years they have lived in the United States. Please see Appendix A for details on these issues.

Figure 1. English Learner Status Among HISD High School Graduates, 2008-2012

- Never EL
- Reclassified in Elementary School
- Reclassified in Middle School
- Reclassified in High School
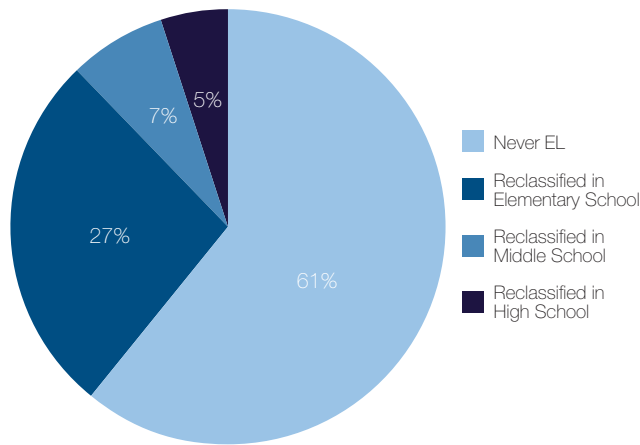
The other independent variable of interest is the number of college-level courses that a student takes during the junior and senior years of high school. We define college-level courses as Advanced Placement (AP), International Baccalaureate (IB), or academic dual credit courses.[11]

Control Variables

Our statistical analyses account for a variety of student- and school-level factors that may influence whether students attend and graduate from a four-year college or whether they take college-level courses during high

11   Academic dual credit courses refer to dual credit courses that are also not considered Career and Technical Education (CTE) courses.

school. The full list of variables included in the models is available in Appendix C. In the analyses, we group these variables into three categories: sociodemographic controls (e.g., gender, race/ethnicity, economic disadvantage), academic controls (e.g., reading test scores, math test scores, average course grades, number of college courses taken), and school controls (e.g., percentage of economically disadvantaged students attending a student's school). Summary statistics on the sample are available in Appendix B.

Analytic Strategy

To address Research Question #1, we estimate multilevel logistic regression models to predict four-year college enrollment and bachelor's degree completion. The key variable of interest is EL status. The reference group is the Never EL category, so the gaps calculate the difference in four-year college enrollment between Never EL students and students reclassified in elementary school and so on. The first model controls for EL status as well as controls for the cohort (e.g., high school graduates in spring 2008) and the number of years that we observe a student in our dataset; details on this latter variable may be found in Appendix A. Results from this first model should be interpreted as the raw gap in four-year college outcomes by EL status. In the second model, we add controls for sociodemographic, academic, and school characteristics to trace how gaps change after accounting for these additional measures. The results from this model should be interpreted as the adjusted gap in four-year college outcomes by EL status. If there are lingering and statistically significant differences between groups, then there may be other factors that we do not account for that may explain why students in certain EL groups attain lower levels of education.

For Research Question #2, we use a zero-inflated negative binomial count model to examine whether, compared to Never EL students, students reclassified in elementary, middle, or high school complete fewer college-level

courses during the junior and senior years of high school. The model controls for sociodemographic and school characteristics.[12]

Finally, for Research Question #3, we use the Karlson-Holm-Breen (KHB) decomposition method to understand how gaps in sociodemographic, academic, and school characteristics feed into gaps in four-year college outcomes.[13, 14,15] (Additional details on the KHB method are available in Appendix D.) For example, after taking a 100-point math test, we find that, on average, Group A students score 20 points higher than Group B students. We can divide that gap into its constituent parts (e.g., 12 points may be attributed to group differences in sociodemographic characteristics, 8 points may be attributed to group differences in academic characteristics). We can further disaggregate the academic characteristics (e.g., six points may be attributed to group differences in math scores from the prior year, two points may be attributed to group differences in course-taking).

12   We do not control for academic factors like test scores or grades because they are potentially endogenous with respect to the outcome; that is, it is unclear whether, for example, differences in grades drive differences in courses taken or vice versa.

13   Breen, R., Karlson, K. B., & Holm, A. (2013). Total, Direct, and Indirect Effects in Logit and Probit Models. *Sociological Methods & Research, 42*(2), 1–28.

14   Karlson, K. B., & Holm, A. (2011). Decomposing primary and secondary effects: A new decomposition method. *Research in Social Stratification and Mobility, 29*(2), 221–237.

15   Karlson, K. B., Holm, A., & Breen, R. (2012). Comparing Regression Coefficients Between Same-sample Nested Models Using Logit and Probit: A New Method. *Sociological Methodology, 42*(1), 286–313.

Results

Research Question #1

What is the role of EL status, specifically the grade level during which a student is reclassified, in predicting four-year college enrollment and completion five years after high school graduation?

In this section, we explore gaps in four-year college outcomes by EL status and how student and school characteristics explain those gaps. Table 1 presents results from multilevel logistic regression models predicting four-year college enrollment.[16] Because the results from logistic regressions can be difficult to interpret, marginal effects (ME) were calculated. By multiplying the numbers in Table 1 by 100, the results may be interpreted as percentage point differences in four-year college enrollment. Model 1 measures the raw difference in four-year college enrollment between students classified as Never EL (reference group) and students reclassified during elementary school, students reclassified during middle school, and students reclassified during high school. First, there are large gaps in four-year college enrollment. Compared to students classified as Never EL, students reclassified during elementary school are 17 percentage points less likely to enroll in a four-year college, while students reclassified during middle school are 32 percentage points less likely to do so. Students reclassified during high school show the lowest four-year college enrollment rates; they are 42 percentage points less likely to enroll than students never classified as EL.

Model 2 presents the gap in four-year college enrollment once sociodemographic (e.g., gender, race/ethnicity, economic disadvantage), academic (i.e., reading test scores, math test scores, the average course grades that a student earns, and the number of college-level courses that a student takes), and school (e.g., percent of economically disadvantaged students, number of AP and IB courses offered) are taken into account. With these variables, we find that differences between students classified as Never EL and students reclassified in elementary, middle, and high school are no longer meaningful. This means that gaps in four-year college enrollment are largely driven by differences in background characteristics.

Table 1: Marginal Effects from Multilevel Logistic Regression Models Predicting Four-Year College Enrollment within Five Years of High School Graduation	Model 1		Model 2	
Variable	ME	Sig.	ME	Sig.
EL Status (ref. = Never EL)				
Reclassified in Elementary School	-0.17	***	0.02	
Reclassified in Middle School	-0.32	***	-0.01	
Reclassified in High School	-0.42	***	-0.02	
Cohort Fixed-Effects	X		X	
Years in HISD Fixed-Effects	X		X	
Sociodemographic Controls			X	
Academic Controls			X	
School Controls			X	

Source: HERC longitudinal database, 1995–2017.

Note: Sample is limited to high school seniors who graduated in the spring semesters between 2008 and 2012 and who had non-missing data (N = 31,374).

* p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests)

Table 2 shows results from a statistical model predicting whether students earned a bachelor's degree.[17] For this model, the sample was limited to students who attended a four-year college the fall after high school. By multiplying the numbers in Table 2 by 100, the results may be interpreted as percentage point differences in four-year college completion five years after high school graduation. The first model presents the raw gaps in bachelor's degree completion between students classified as Never EL and students reclassified in elementary, middle, and high school. All three gaps are statistically significant. Compared to students classified as Never EL, students reclassified during elementary school are 5 percentage points less likely to earn a bachelor's degree, while students reclassified during middle and high school are 12 percentage points less likely to do so.

In Model 2, we account for sociodemographic, academic, and school characteristics and, again, show that differences between by EL status are no longer meaningful. This means that gaps in four-year college completion are largely driven by differences in background characteristics.

---

16   The complete regression results are available from the authors upon request.

17   The complete regression results are available from the authors upon request.

Table 2: Marginal Effects from Multilevel Logistic Regression Models Predicting Bachelor's Completion within Five Years of High School Graduation Conditional on Immediate Enrollment in a Four-Year College

	Model 1		Model 2	
Variable	ME	Sig.	ME	Sig.
EL Status (ref. = Never EL)				
Reclassified in Elementary School	-0.05	***	0.00	
Reclassified in Middle School	-0.12	***	-0.03	
Reclassified in High School	-0.12	***	-0.01	
Cohort Fixed-Effects	X		X	
Years in HISD Fixed-Effects	X		X	
Sociodemographic Controls			X	
Academic Controls			X	
School Controls			X	

Source: HERC longitudinal database, 1995–2017.

Note: Sample is limited to high school seniors who graduated in the spring semesters between 2008 and 2012, who enrolled in a four-year college the fall immediately after high school graduation, and who had non-missing data (N = 13,443).

* p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests)

Research Question #2

Is there a relationship between the grade level of reclassification and college-level course-taking during secondary school?
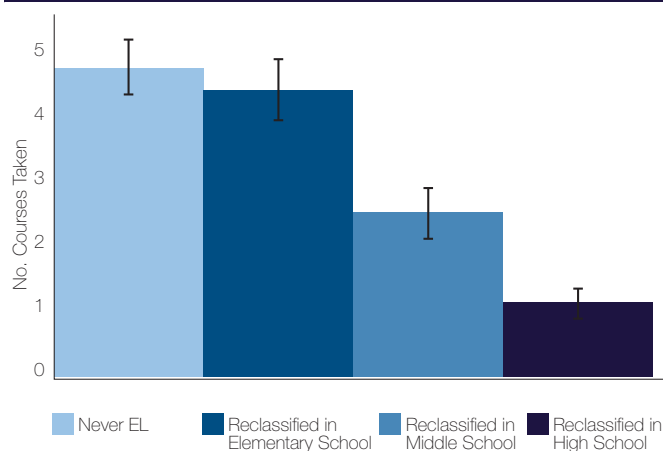
For the remainder of this report, we focus on one particular characteristic that may explain gaps in four-year college outcomes: college-level course-taking during high school. As mentioned earlier, it is possible that former EL students have less access to AP, IB, or academic dual credit courses. They may either lack the English language skills to be successful or not enough time to enroll because they are busy taking English language acquisition courses. Regardless of the reason, we first test whether there are meaningful differences in college-level course-taking among the four EL groups. To do this, we estimate a zero-inflated negative binomial count model which predicts the number of college-level courses a student takes and controls for EL status, sociodemographic factors, and school factors.[18]

Figure 2 graphs the results from this analysis. The y-axis shows the average number of college courses a student is predicted to take during the junior and senior years of high school. Students classified as Never EL (light blue) are predicted to take 4.8 college courses, while students reclassified during elementary school (dark blue) take 4.4 courses. The difference between these two groups is not statistically significant since their error bars, represented by 95 percent confidence intervals, overlap. Even though students reclassified during elementary school face unique challenges, they appear able to overcome those challenges and take a similar number of college-level courses during the junior and senior years of high school.

Students reclassified during middle school (bright blue) are predicted to take 2.6 college courses, while students reclassified during high school (navy) take 1.1 courses. The graph shows that both groups take fewer college courses than Never EL students and students reclassified during elementary school, and that these differences are meaningful since the error bars do not overlap. In addition, because the error bar on the middle school bar does not overlap with the error bar on the high school bar, we can also conclude students reclassified during high school take fewer college courses than students reclassified during middle school.

Figure 2. Junior and Senior Year College Course-Taking by EL Status[19]

Legend: Never EL | Reclassified in Elementary School | Reclassified in Middle School | Reclassified in High School

y-axis: No. Courses Taken

18　We do not control for academic factors like test scores or grades because they are potentially endogenous with respect to the outcome; that is, it is unclear whether, for example, differences in grades drive differences in courses taken, or vice versa.

19　The graph is produced using the `margins` command in Stata after estimating a zero-inflated negative binomial count model. The full regression results are available in Appendix E.

We want to note that a key reason why students reclassified during middle and high school take fewer college-level courses than Never EL students and students reclassified during elementary school may be tied to English language ability: students may not have the reading, writing, speaking, and listening skills to be successful in those courses. It is beyond the scope of this study to determine what level of English language ability must be achieved in order to take and be successful in a college-level course. Although English language ability may be a consideration in the course recommendation process (even if it limits access for ELs), there may be alternative strategies to broaden ELs' access to college-level coursework. We will return to this idea in the "Potential Future Directions as Outlined by HISD" section.

Research Question #3

Is any share of the gap in four-year college enrollment or completion by EL status (i.e., grade level of reclassification) attributable to gaps in college-level course-taking during high school by EL status? If so, how much? Does the explained variance differ across EL group—former EL reclassified during elementary school, former EL reclassified during middle school, and former EL reclassified during high school compared to Never EL, respectively?

In this section, we take a deeper dive into college-level course-taking by addressing how EL gaps in course-taking feed into EL gaps in four-year college outcomes. This information may be useful to district practitioners interested in increasing college enrollment and completion. If, for example, gaps in course-taking explain a significant share of the gap in four-year college outcomes, then the district may wish to identify strategies that can level the course-taking playing field among EL groups.

Using the KHB method of non-linear variance decomposition, we disentangle the EL gaps presented in Tables 1 and 2. Results from the decomposition of gaps in four-year college enrollment are shown in Figure 3. In the left panel, each full bar represents the total gap between Never EL students and students reclassified in elementary, middle, or high school. For example, Model 1 in Table 1 shows that there is a raw gap of 42 percentage

points between Never EL students and students reclassified in high school; we see that reflected on the graph. The four sections of each bar show how gaps in background characteristics feed into gaps in four-year college enrollment. When we discuss gaps due to differences in:

Sociodemographic Characteristics
We refer to overall differences between Never EL students and former EL students in terms of characteristics like gender, race/ethnicity, and economic disadvantage. A full list of sociodemographic characteristics is available in Appendix C.

Academic Characteristics
We refer to overall differences between Never EL students and former EL students in terms of the following academic characteristics: reading test scores, math test scores, average course grades, and the number of college courses taken.

School Characteristics
We refer to overall differences between Never EL students and former EL students in terms of the following school characteristics: the percentage of EL students, the percentage of economically disadvantaged students, the number of AP and IB courses offered, and the percentage of high school graduates who attended college the fall following graduation.

Unexplained Characteristics
We refer to overall differences between Never EL students and former EL students that we cannot explain with the characteristics considered in the analyses.

Please note that sections above the zero line mean that the characteristic contributes to gaps between Never EL students and students reclassified during elementary, middle, and high school. In contrast, sections below the zero line reduce gaps; practically speaking, the characteristic advantages former EL students relative to students in the Never EL reference group.

The largest section is green; this means that gaps in enrollment appear to flow from differences in student sociodemographic characteristics like gender, race/ethnicity, and economic disadvantage.[20] Since school administrators, district leaders, and policymakers have little power to change these characteristics, we will focus on the differences in academic characteristics.

The right side of the panel shows how differences in academic characteristics as a whole (i.e., the combination of differences in reading test scores, math test scores, average course grades, and the number of college courses taken) contribute to the total gap in four-year college enrollment. Differences in academic characteristics amount to 13 percent of the total gap between Never EL students and students reclassified in elementary school, 30 percent of the total gap between Never EL students and students reclassified in middle school, and 38 percent of the total gap between Never EL students and students reclassified in high school. However, the portion that is tied to differences in college-level course-taking constitutes the majority of the portion due to differences in academic characteristics: 59 percent of the gap between Never EL students and students reclassified in elementary school, 60 percent of the gap between Never EL students and students reclassified in middle school, and 57 percent of the gap between Never EL students and students reclassified in high school. If, for example, students reclassified during high school complete the same number of college-level courses as Never EL students (if the course-taking gap is null), the four-year college enrollment gap would decrease from 42 percentage points to 33 percentage points. Differences in college-level

course-taking during high school play a strong role in explaining differences in enrollment.

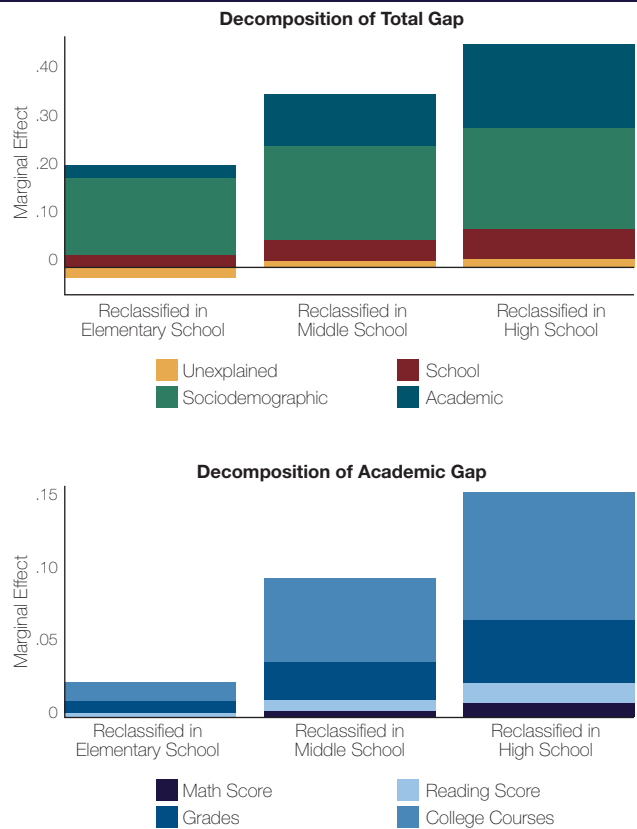Figure 3. Decomposition of EL Gaps in Four-Year College Enrollment

Figure 4 presents results from the decomposition of gaps in bachelor's degree completion among students who enrolled in a four-year college the fall after high school. For students reclassified during elementary and middle school, we see that differences in academic characteristics (in blue) — which include differences attributed to reading and math scores, grades, and college courses — fall below the zero line. As a whole, gaps in academic characteristics are not contributing to their lower levels of bachelor's degree completion; rather, they are helping their completion rates. In contrast, for students reclassified during high school, differences in academic characteristics still contribute to differences in bachelor's degree completion and amount to 16 percent of the total gap between Never EL students and students reclassified in high school.

The story is more complicated when we separate differences in academic characteristics into constituent parts: there are academic characteristics that fall above and below the zero line. For example, for all three groups, grades fall below the zero line. This means that differences in grades are not contributing to gaps in bachelor's degree completion between Never EL students and students reclassified in elementary, middle, and high school. In fact, former EL students' grades are higher than Never EL students' grades and push former ELs through college. However, for students reclassified in middle and high school, the college-level course-taking portion is above the zero line, which means that differences in course-taking contribute to their lower levels of bachelor's degree completion. If students reclassified in middle and high school take the same number of college-level courses as Never EL students, we may see their bachelor's degree completion rates increase by two and five percentage points, respectively.[21] Taken together, differences in college-level course-taking explain a sizeable share of the gap in bachelor's degree completion between Never EL students and students reclassified during middle and high school.

Figure 4. Decomposition of EL Gaps in Bachelor's Degree Completion Conditional on Immediate Enrollment in a Four-Year College

Decomposition of Total Gap

Legend: Unexplained, School, Sociodemographic, Academic

Decomposition of Academic Gap

Legend: Math Score, Reading Score, Grades, College Courses

---

21   In results not shown, we find that differences in school characteristics, particularly the percentage of economically disadvantaged students and the percentage of high school graduates who enroll in college the fall following graduation, play a significant role in bachelor's degree completion. While these results are interesting, they do not pertain to the goal of the study, which is to examine the role of college-level course-taking during high school in producing gaps in four-year college enrollment and completion by EL status. However, additional details on these results and others are available from the authors upon request.

Discussion

The findings in this report show that while gaps in four-year college enrollment and bachelor's degree completion by EL status appear large, they are fully explained by sociodemographic, academic, and school characteristics. After adding these variables to the statistical models, there are no meaningful differences in enrollment or completion between Never EL students and students reclassified in elementary, middle, and high school.

The analyses also show that college-level course-taking varies by EL status. After controlling for sociodemographic and school characteristics, students reclassified in middle and high school take fewer college-level courses than Never EL students. In contrast, Never EL students and students reclassified in elementary school appear to take similar numbers of college-level courses. The gap between Never EL students and students reclassified during high school is intuitive since the analyses focus on college-level course-taking during the junior and senior years. Why students reclassified during middle school take fewer of these courses needs additional study. However, one possibility is that taking college-level courses during the junior and senior years of high school is the culmination of a course-taking pathway that begins much earlier in students' educational careers. While still many years away from their junior or senior year of high school, EL students in middle school may have their coursework trajectories diverted in a way that eventually contributes to a lower likelihood of enrolling in a four-year college or completing a four-year college degree within five years of high school graduation.

Finally, the decomposition analyses show that, when considering academic characteristics, gaps in college-level course-taking explain a significant share of the gap in four-year college outcomes between Never EL students and students reclassified in elementary, middle, and high school. In fact, if students reclassified during high school take the same number of college-level courses as Never EL students, the gap in four-year college enrollment would decrease from 42 to 33 percentage points, while the gap in bachelor's degree completion gap would reduce from 12 to 7 percentage points.

While our analyses suggest gaps in college-level course-taking by EL status feed into EL gaps in four-year college outcomes, we cannot predict how increasing EL access to college-level coursework during high school will affect outcomes. As mentioned earlier in this section, course-taking during the junior and senior years is possibly part of a pathway that begins during middle school. Students who are not able to take pre-AP, pre-IB, gifted, and honors courses in middle school may be set on a different educational trajectory than those who do take these courses. Because of the complicated interplay between course sequencing, academic skills, and information, simply placing a student in an AP, IB, or academic dual credit course may not automatically increase four-year college enrollment and completion. Long-term solutions that identify ways to ensure students are adequately prepared for and take more advanced courses during middle and high school are likely necessary to increase postsecondary attainment.

Potential Future Directions as Outlined by HISD

If schools and districts wish to make four-year college enrollment and bachelor's degree completion more accessible by EL status, they may want to consider increasing access to college-level course-taking for all students, regardless of EL status. Through conversations with staff in HISD's Multilingual Education Department, we have developed the following list of potential future directions:

- **Design ways in which EL students can receive advanced course content that can adequately prepare them for college:** For example, schools may consider delivering college-level coursework in students' native language (if there are qualified teachers available for this instruction) or in English, but provide EL students with a tutor for assistance if they struggle with the curriculum.

- **Ensure EL students have adequate access to advanced course content as early as middle school:** It is possible that students who take pre-AP, pre-IB, gifted, and honors courses in middle school are set on a different course-taking and college-going trajectory from students who do not take those courses. Therefore, it may be important to have EL students complete advanced coursework in middle school.

- **Encourage EL students who meet a certain proficiency level in English reading, writing, speaking, and listening to take advanced courses:** This can include courses which may not rely on high levels of English language comprehension, like studio art or Spanish literature.

- **Integrate teacher recommendations into the course recommendation process:** Currently, students are placed into college-level courses based on academic measures like grades and completing prerequisites. These academic measures may not reflect a student's potential, interest, or motivation for completing college-level coursework or their educational goals. Through their day-to-day interactions with students, teachers may have a deep understanding of whether EL students are capable of succeeding in these rigorous courses.

- **Have members of each school's Language Proficiency Assessment Committee (LPAC) communicate with counselors and registrars about EL students' academic performance:** The LPAC regularly reviews EL students' academic performance in English language acquisition as well as other subjects. The LPAC may be able to inform school counselors and registrars on EL students' areas of strength and identify potential courses for students.

- **Offer trainings for counselors and registrars on course-taking for EL students:** In order to assist staff in making informed decisions about appropriate placement as they register students for courses, it may be helpful to embed information about the types of courses that EL students are allowed to enroll in as part of existing trainings.

- **Strengthen connections between staff in the Multilingual Education Department and the College Readiness Department:** HISD has implemented a number of programs which aim to ease the transition from high school to college. Staff from the Multilingual Education Department can collaborate with staff from these programs to provide guidance on how to prepare EL students for college or to develop strategies targeting EL students for specialized advising.

- **Help EL students reclassify early by identifying evidence-based instructional practices which can aid English language acquisition:** Many instructors, particularly those working in high schools, may have little exposure to EL students and possess inadequate training to teach them. Regardless of their experiences and specialties, teachers may benefit from additional training to integrate literacy practices into coursework like sheltered English instruction, which is "an instructional approach that engages [ELs] above the beginner level in developing grade-level content-area knowledge, academic skills, and increased English proficiency.[22]

---

22 The Education Alliance at Brown University. (n.d.). Sheltered English Instruction. Retrieved March 11, 2019, from https://www.brown.edu/academics/education-alliance/teaching-diverse-learners/strategies-0/sheltered-english-instruction-0.

Appendix A: Sample Selection

Many EL students are immigrants and may have less exposure to United States public schools and American society. A student who arrives in the United States during their 10th grade year has no opportunity to be reclassified in elementary or middle school, and their EL status may make it more difficult to access college-level high school coursework from that point forward. We address this data limitation in three ways:

1. First, we include years in HISD fixed-effects, which account for the fact that students attend HISD schools for different numbers of years. Many of the students who attend HISD schools fewer than 13 years may be immigrants, but many are not (e.g., a student who attends Aldine Independent School District in elementary school and then transfers to HISD schools prior to graduation). These fixed-effects are useful for two reasons. One, we do not observe EL status prior to HISD attendance, so they partially address the problem with unobservables. And two, they allow us to isolate the relationship between EL status and postsecondary outcomes without confounding partial attendance in HISD. While the years in HISD fixed-effects are useful beyond immigrant students, they do, in part, account for the fact that some immigrant students may have less exposure to HISD schools because they come to the United States after kindergarten.

2. Second, using enrollment form data, we can calculate how many years each student has lived in the United States and include that variable as a control in the statistical models. We may be worried that the years in HISD fixed-effects do not adequately account for immigration. For example, two immigrant students who entered HISD schools in kindergarten may have come to the United States at different times (e.g., one immigrated at age 1, while the other immigrated weeks before school started). By adding this continuous control variable, we can further account for differences in immigration and U.S. arrival.

3. Third, we conduct a series of robustness checks, one of which is to limit the models to students who attended HISD schools continuously from kindergarten through 12th grade. We wish to note that limiting the sample to students who attended HISD schools for 13 years is a narrower and more selective sample; it simply serves as a robustness check to verify our findings.[23]

---

23   These results are available from the authors upon request. The findings are substantively similar but imprecisely estimated.

Appendix B: Sample Characteristics

Variable	All Students	Never EL Students	Students Reclassified in Elementary School	Students Reclassified in Middle School	Students Reclassified in High School
Dependent Variables					
Four-Year College Enrollment	0.54	0.64	0.45	0.28	0.21
Bachelor's Degree Completion Conditional on Immediate Enrollment in a Four-Year College	0.45	0.47	0.42	0.33	0.36
Sociodemographic Control Variables					
Age on September 1 of Gr. 12	17.64	17.61	17.64	17.72	17.93
Female	0.54	0.55	0.53	0.53	0.48
Race/Ethnicity					
White	0.13	0.21	0.01	0.01	0.02
Black	0.29	0.47	0.01	0.01	0.04
Hispanic	0.52	0.27	0.92	0.95	0.86
Asian	0.05	0.05	0.06	0.04	0.08
Immigrant	0.13	0.02	0.17	0.42	0.75
Years in U.S.	16.49	17.39	16.56	14.06	8.79
Non-English Home Language	0.41	0.06	0.95	0.98	1.00
Economically Disadvantaged	0.64	0.51	0.83	0.88	0.86
Socioeconomic Status in Census Block Group	-0.28	-0.06	-0.61	-0.66	-0.64
One-Parent Household	0.36	0.43	0.23	0.26	0.31
Academic Control Variables					
TAKS Reading Score (Gr. 11)	0.07	0.20	0.08	-0.39	-0.98
TAKS Math Score (Gr. 11)	0.15	0.17	0.23	-0.04	-0.32
Junior/Senior Course Grades (in 10s)	8.41	8.44	8.39	8.31	8.27
Junior/Senior College Courses Taken	4.61	5.13	4.53	2.65	1.55
School Control Variables					
School-Level % English Learner (in 10s)	0.86	0.67	1.06	1.31	1.52
School-Level % Econ. Disadv. (in 10s)	6.73	6.22	7.44	7.67	7.70
School-Level No. AP/IB Courses Offered	28.66	28.86	28.18	28.26	29.38
School-Level % Immediate Enrollment (in 10s)	5.77	6.14	5.37	4.91	4.76
N	31,374	19,087	8,533	2,097	1,657

Source: HERC longitudinal database, 1995–2017.

Note: Sample is limited to high school seniors who graduated in the spring semesters between 2008 and 2012 and who had non-missing data.

Appendix C: Variables Used in the Analyses

Variable	All Students
Dependent Variables	
Four-Year College Enrollment	Binary: Whether a student enrolled in a four-year college or university within five years of high school graduation.
Bachelor's Degree Completion Conditional on Immediate Enrollment in a Four-Year College	Binary: Among students who enrolled in a four-year college or university, whether a student completed a bachelor's degree within five years of high school graduation.
Sociodemographic Control Variables	
Age on September 1 of Gr. 12	Continuous.
Female	Binary.
Race/Ethnicity	Categorical: White (ref.), Black, Hispanic, and Asian.
Immigrant	Binary: Foreign-born.
Years in U.S.	Continuous.
Non-English Home Language	Binary: Primary language spoken at home is a foreign language.
Economically Disadvantaged	Binary: Eligible for free and reduced-price lunch program or other federal poverty programs or living below the federal poverty line.
Socioeconomic Status in Census Block Group	Continuous: Index created using American Community Survey data and principal components analysis. Based on six Census Block Group-level measures: the proportion without a high diploma, the proportion with a bachelor's degree, the proportion working in professional/managerial occupations, the median household income, the proportion living in poverty, and the proportion of female-headed households (alpha = 0.90).
One-Parent Household	Binary.
Academic Control Variables	
TAKS Reading Score (Gr. 11)	Continuous: Reading test score from the 11th grade Texas Assessment of Knowledge and Skills (TAKS). Reported in standard deviation units.
TAKS Math Score (Gr. 11)	Continuous: Math test score from the 11th grade Texas Assessment of Knowledge and Skills (TAKS). Reported in standard deviation units.
Junior/Senior Course Grades (in 10s)	Continuous: Average percentage grade among all courses taken in the 11th and 12th grades (reported in 10s).
Junior/Senior College Courses Taken	Continuous: Number of Advanced Placement (AP), International Baccalaureate (IB), and academic dual credit courses taken in the 11th and 12th grades.
School Control Variables	
School-Level Percent English Learner (in 10s)	Continuous: Percentage of English Learner students (reported in 10s). Calculated by aggregating student data to the school level.
School-Level Percent Economically Disadvantaged (in 10s)	Continuous: Percentage of economically disadvantaged students (reported in 10s). Calculated by aggregating student data to the school level.
School-Level Number of AP and IB Courses Offered	Continuous: Number of AP and IB courses offered during the 11th and 12th grades. Calculated by aggregating student data to the school level.
School-Level Percent Immediate Enrollment (in 10s)	Continuous: Percentage of high school graduates who attended college the fall following graduation (reported in 10s). Calculated by aggregating student data to the school level.
Other Control Variables	
Years in HISD	Categorical: Number of years a student enrolled in a HISD school.
Cohort	Categorical: High school graduation class of spring 2008 (ref.), 2009, 2010, 2011, and 2012.

Source: HERC longitudinal database, 1995–2016.

Appendix D: Statistical Models

To answer Research Question #1, we estimate multilevel logistic regression models of the following form:

Level 1

$$logit(Y_{ij}) = \beta_{0j} + \beta_{1j}RECLASS\_ELEM_{1ij} + \beta_{2j}RECLASS\_MID_{2ij}$$
$$+ \beta_{3j}RECLASS\_HIGH_{3ij} + \sum_{q=4}^{Q} \beta_{qj}X_{qij}$$

Level 2

$$\beta_{0j} = \gamma_{00} + \sum_{s=1}^{S} \gamma_{0s}W_{sj} + u_{0j}$$
$$\beta_{1j} = \gamma_{10}$$
$$\beta_{2j} = \gamma_{20}$$
$$\vdots$$
$$\beta_{qj} = \gamma_{q0}$$

where Y_{ij} is the dichotomous indicator of enrolling in a four-year college or completing a bachelor's degree. The coefficient β_{0j} is a school-level random intercept and $\sum_{q=4}^{Q}\beta_{qj}X_{qij}$ is a vector of student-level covariates. The coefficients on β_{1j}, β_{2j}, and β_{3j} represent the difference in four-year college outcomes between Never EL students and former EL students reclassified during elementary, middle, and high school, respectively. Each school-level intercept is modeled as a function of school-level covariates $\sum_{s=1}^{S}\gamma_{0s}W_{sj}$ and a normally distributed error term u_{0j}. Model 1 controls for EL status only, while groups of sociodemographic (Model 2), academic (Model 3), and school (Model 4) are added subsequently.

To address Research Question #2, we estimate a zero-inflated negative binomial count model and cluster standard errors by school.

To address Research Question #3, we use the Karlson-Holm-Breen (KHB) decomposition method. Like Research Question #1, we use a multilevel logistic model. The KHB approach is traditionally used to examine mediation — to distinguish between direct and indirect effects. The focus of our study is on the indirect effects of EL status on four-year college outcomes, particularly academic characteristics and, more specifically college-level course-taking. Our study may be considered a gaps-on-gaps analysis since we seek to decompose the total gap in enrollment or completion by EL status into its constituent parts — gaps attributed to gaps in sociodemographic characteristics, gaps attributed to gaps in academic characteristics, and gaps attributed to gaps in school characteristics. In effect, we estimate the following for students reclassified in elementary, middle, and high school:

$$(\bar{Y}_{NeverEL} - \bar{Y}_{EL})$$
$$= (\hat{\alpha}_{NeverEL} - \hat{\alpha}_{EL}) + (\overline{SOCDEM}_{NeverEL} - \overline{SOCDEM}_{EL})\hat{\delta}$$
$$+ (\overline{ACADEMIC}_{NeverEL} - \overline{ACADEMIC}_{EL})\hat{\gamma} + (\overline{SCHOOL}_{NeverEL}$$
$$- \overline{SCHOOL}_{EL})\hat{\lambda}$$

In this model, the outcome $\bar{Y}_{NeverEL} - \bar{Y}_{EL}$ is the gap in enrollment or completion between Never EL students and students reclassified in elementary, middle, and high school. The outcome is modeled as a function of gaps in background characteristics (e.g., the Never EL-EL difference in economic disadvantage, the Never EL-EL difference in college-level course-taking). The terms $\hat{\delta}$, $\hat{\gamma}$, and $\hat{\lambda}$ represent vectors of estimated coefficients and show how gaps in background characteristics explain gaps in four-year college outcomes.

Research Questions #1 and #3 are estimated in Stata using the KHB commands, `khbtab` and `khb`, respectively.[24] In both linear and logistic regression models, confounder variables can change a coefficient's magnitude. However, in a logistic model, a coefficient's magnitude can also change due to rescaling; basically, adding variables to a logistic model increases the outcome's latent variance. Therefore, when comparing coefficients across nested logistic regression models, any changes observed may be due to the addition of confounders or the rescaling problem.[25] The KHB approach addresses the problem of rescaling, making it more possible to compare coefficients across models and to decompose group differences based on these rescaled results.

24   Kohler, U., Karlson, K. B., & Holm, A. (2011). Comparing coefficients of nested nonlinear probability models. *The Stata Journal, 11*(3), 420–438.
25   Karlson, K. B., & Holm, A. (2011). Decomposing primary and secondary effects: A new decomposition method. Research in Social Stratification and Mobility, 29(2), 221–237.

Appendix E: Zero-Inflated Negative Binomial Count Model Predicting Junior and Senior Year College Course-Taking

Negative Binomial Regression			
Variable	β	S.E.	Sig.
EL Status (ref. = Never EL)			
Reclassified in Elementary School	-0.04	(0.02)	
Reclassified in Middle School	-0.30	(0.03)	***
Reclassified in High School	-0.76	(0.05)	***
Age on September 1 of Gr. 12	-0.03	(0.01)	
Female	0.04	(0.02)	**
Race/Ethnicity (ref. = White)			
Black	-0.30	(0.06)	***
Hispanic	-0.25	(0.04)	***
Asian	0.21	(0.14)	
Immigrant	-0.10	(0.06)	
Years in U.S.	-0.03	(0.01)	***
Non-English Home Language	0.04	(0.03)	
Economically Disadvantaged	-0.05	(0.03)	
Socioeconomic Status in Census Block Group	0.06	(0.01)	***
One-Parent Household	-0.06	(0.01)	***
School-Level % English Learner (in 10s)	0.05	(0.05)	
School-Level % Econ. Disadv. (in 10s)	0.05	(0.03)	*
School-Level No. AP/IB Courses Offered	0.00	(0.00)	
School-Level % Immediate Enrollment (in 10s)	0.13	(0.03)	***
Intercept	0.76	(0.34)	*
Ln(Dispersion Parameter)	-1.22	(0.08)	***
—2 Log-Likelihood	-72,628		

Logistic Regression for Zero Inflation			
Variable	β	S.E.	Sig.
EL Status (ref. = Never EL)			
Reclassified in Elementary School	0.18	(0.11)	
Reclassified in Middle School	0.82	(0.16)	***
Reclassified in High School	1.45	(0.15)	***
Age on September 1 of Gr. 12	0.23	(0.03)	***
Female	-0.35	(0.05)	***
Race/Ethnicity (ref. = White)			
Black	1.15	(0.14)	***
Hispanic	1.01	(0.13)	***
Asian	-0.40	(0.32)	
Immigrant	0.25	(0.15)	
Years in U.S.	0.07	(0.02)	***
Non-English Home Language	-0.21	(0.08)	**
Economically Disadvantaged	0.12	(0.07)	
Socioeconomic Status in Census Block Group	-0.09	(0.05)	
One-Parent Household	0.26	(0.05)	***
School-Level % English Learner (in 10s)	-0.49	(0.13)	***
School-Level % Econ. Disadv. (in 10s)	-0.12	(0.08)	
School-Level No. AP/IB Courses Offered	0.04	(0.01)	***
School-Level % Immediate Enrollment (in 10s)	-0.51	(0.11)	***
Intercept	0.64	(1.13)	

Source: HERC longitudinal database, 1995–2017.

Note: Sample is limited to high school seniors who graduated in the spring semesters between 2008 and 2012 and who had non-missing data (N = 31,374). The models include years in HISD and cohort fixed-effects and standard errors are clustered at the school level.

* p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests)

Mission

Focusing on the most pressing issues facing the region, the Houston Education Research Consortium (HERC) is a research-practice partnership between Rice University and 11 area school districts. HERC, housed within the Kinder Institute for Urban Research, develops research directly alongside district leaders with findings shared with decision makers – culminating in long-term, equity-minded solutions, opportunities and growth for Houston and beyond.

RICE | HERC
Houston Education Research Consortium