Transitioning to College and Work
Part 2: A Study of Potential Enrollment Indicators

by Brian Holzman, Ph.D., and Horace Duffy, M.A. July 2020

Appendix A. Developing the HERC Indicator of College Enrollment

Although the Chicago indicator was based on prior research (Allensworth & Easton, 2005) and the state indicator was based on education policy (Texas Education Agency, 2007), there are relatively few early warning indicators of college enrollment. Analytic decisions were made by consulting academic publications and research reports, as well as exploring data; the researchers engaged in an iterative, conceptual, and data-driven process.

Using the Chicago indicator as a reference point, it seemed important to incorporate measures of grades and credits into a potential college enrollment indicator. Higher education research shows grades are more important to college success than standardized test scores (Geiser & Santelices, 2007). In the researchers' view, grades are useful not only because they serve as a measure of academic performance, but also because they, in part, reflect student motivation and effort. These non-cognitive skills, for instance, are highlighted in Conley's framework of college readiness (2010). While passing classes and earning credits are surely important to college enrollment, they may be a low bar, especially since research shows the rigor of one's high school curriculum is correlated with enrollment outcomes (Long, Conger, & Iatarola, 2012). Focusing on advanced credits earned, rather than overall credits earned, also aligns with the state's new accountability system. Its College, Career, and Military Readiness component can be satisfied by earning high scores on Advanced Placement or International Baccalaureate exams or earning specified numbers of dual credits (Texas Education Agency, 2018).

Using the HERC data, a series of logistic regression models that predicted college enrollment and controlled for a variety of student background characteristics, cohort fixed-effects, and school fixed-effects was estimated. From these models, it was determined attendance, grades, and advanced credits earned were most strongly associated with enrollment. The research team reviewed the distributions of these three variables to try to identify tipping points — points at which the distribution changed in a way that affected college enrollment. The tipping point for attendance appeared to be around 90 percent.[1] In terms of grades, the tipping point was around 80 percent, which corresponded to a B-average.[2]

Distributional analyses provided no clear tipping point for advanced credits. A number of potential tipping points were tested by regressing college enrollment on dummy variables that indicated whether a student earned a certain number of advanced credits between grades seven and 12: any, three, six,

[1] A common definition of chronic absenteeism is missing 10 percent or more of the days a student is enrolled in school (García & Weiss, 2018).

[2] Researchers at the University of Chicago Consortium on School Research found, "[s]tudents who earn Cs or Ds in high school are unlikely to graduate from college, while those with a B average (a 3.0 GPA) have about a 50/50 chance of earning a four-year college degree" (Allensworth, Gwynne, Moore, & de la Torre, 2014, p. 55).

nine, 12, 15, and 18. The pseudo R-squareds from these models suggested earning any advanced credits explained the least amount of variation (0.11). In contrast, the pseudo R-squareds from the other tipping points ranged from 0.14-0.15. The analyses proceeded with the three-credit tipping point. Earning three advanced credits between grades seven and 12 corresponded to an average of one-half credit per year, which was equivalent to passing a semester-long advanced course.

In the end, the binary HERC indicator was based on satisfying three binary measures in each grade:

1. Having an attendance rate greater than or equal to 90 percent (i.e., non-chronic absenteeism)
2. Earning an average grade percentage greater than or equal to 80 percent (i.e., B-average or higher)
3. Earning at least 0.5 advanced credits

The development of the potential indicator was an iterative process that underwent many versions. Throughout the different versions, the findings were generally consistent. The potential indicator was developed to move the conversation on early warning systems toward college enrollment and attempt to provide practitioners with a tool that might help achieve goals. The HERC indicator, however, might not be appropriate in all contexts and settings. For example, earning at least 0.5 advanced credits might not be a useful measure in a school or district where no advanced courses were offered or where all students took advanced courses. The HERC indicator might also not be appropriate if the goal was to increase four-year or selective college enrollment rather than any college enrollment. In that case, changing the grades and/or credits components (e.g., earning an A-average or higher, earning two or more advanced credits) might be warranted.

Additional details on the HERC indicator's development and its history are available from the authors upon request.

Appendix B. Control Variables Used in Analyses

Variable	Description
Grade Level	Categorical: Seventh (ref.), Eighth, Ninth, 10th, 11th, and 12th.
Number of Total Credits Earned in Sixth Grade	Continuous: Courses failed (percentage grade below 69.5) were not counted.
Number of Core Courses Failed in Sixth Grade	Continuous (reported in semesters).
Attendance Rate in Sixth Grade	Continuous: Percentage of school days attended (reported in 10s).
Average Grade Percentage in Sixth Grade	Continuous: Average percentage grade among all courses taken (reported in 10s).
Number of Advanced Credits Earned in Sixth Grade	Continuous: Advanced courses include Pre-Advanced Placement (Pre-AP), Pre-International Baccalaureate (Pre-IB), AP, IB, and academic dual credit courses (dual credit courses that were not Career & Technical Education).
English/Language Arts Test Score in Sixth Grade	Continuous: English/Language Arts test score from the Texas Assessment of Knowledge and Skills (TAKS; reported in standard deviation units).
Mathematics Test Score in Sixth Grade	Continuous: Mathematics test score from the Texas Assessment of Knowledge and Skills (TAKS; reported in standard deviation units).
Age	Continuous.
Female	Binary.
Race/Ethnicity	Categorical: White (ref.), Black, Hispanic, and Asian.
English Learner	Binary.
Special Education	Binary.
Economically Disadvantaged	Binary: Was eligible for the free and reduced-price lunch program or other federal poverty programs, or living below the federal poverty line.
Number of In-School Suspensions	Continuous.
Number of Out-of-School Suspensions	Continuous.

Source: HERC multi-year data.

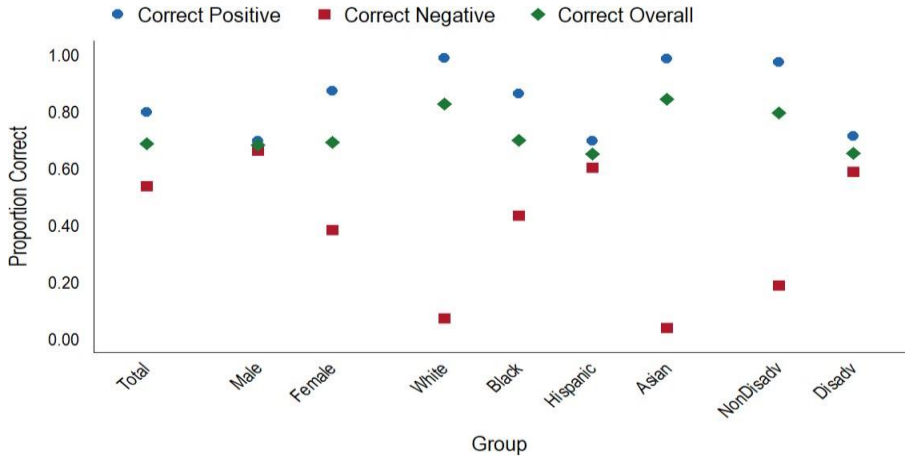
Appendix C. Research Question 1 Summary Statistics

Variable	Total	Enrolled	Not Enrolled
Age	12.30	12.20	12.43
Female	0.52	0.57	0.45
Black	0.24	0.26	0.22
Hispanic	0.62	0.54	0.73
Asian	0.04	0.06	0.01
English Learner	0.12	0.07	0.18
Special Education	0.05	0.03	0.07
Economically Disadvantaged	0.76	0.68	0.87
N (students)	12,011	6,918	5,093

Source: HERC multi-year data.
Note: Sample was limited to non-Native American students who were not missing data.

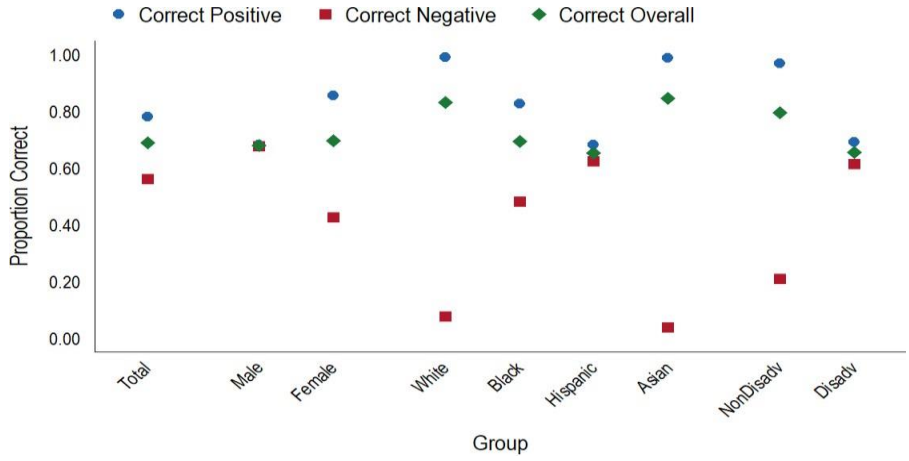
Appendix D. Correct Predictions of College Enrollment by Gender, Race, and SES

Figure D1. Chicago Indicator in Seventh Grade



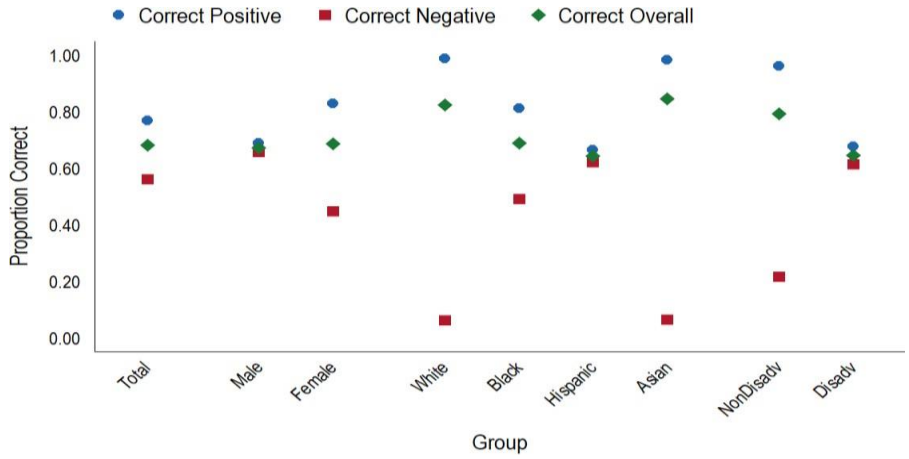
Source: HERC multi-year data.
Note: Sample was limited to 12,011 non-Native American students who were not missing data. Results came from logistic regression models that predicted college enrollment; controlled for the Chicago, HERC, or state indicator in grade seven; controlled for student background characteristics (age, gender, race/ethnicity, English learner, special education, economic disadvantage); and controlled for cohort and school fixed-effects (dummy variables for each cohort, dummy variables for each school attended in seventh grade).

Figure D2. HERC Indicator in Seventh Grade



Source: HERC multi-year data.
Note: Sample was limited to 12,011 non-Native American students who were not missing data. Results came from logistic regression models that predicted college enrollment; controlled for the Chicago, HERC, or state indicator in grade seven; controlled for student background characteristics (age, gender, race/ethnicity, English learner, special education, economic disadvantage); and controlled for cohort and school fixed-effects (dummy variables for each cohort, dummy variables for each school attended in seventh grade).

Figure D3. State Indicator in Seventh Grade



Source: HERC multi-year data.

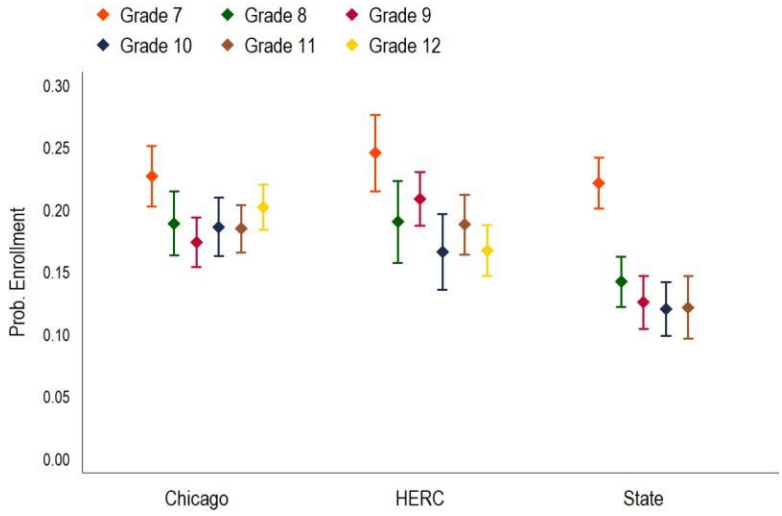
Note: Sample was limited to 12,011 non-Native American students who were not missing data. Results came from logistic regression models that predicted college enrollment; controlled for the Chicago, HERC, or state indicator in grade seven; controlled for student background characteristics (age, gender, race/ethnicity, English learner, special education, economic disadvantage); and controlled for cohort and school fixed-effects (dummy variables for each cohort, dummy variables for each school attended in seventh grade).

Appendix E. Additional Potential Indicator Analysis — What is the Total Effect of Meeting a Potential Indicator on College Enrollment?

There is something different about ninth grade — students were more likely to fall off-track from college enrollment that year. If school or district practitioners notice a student is falling off-track, should they wait until ninth grade to help that student get back on-track to college enrollment? Student academic performance in ninth grade is likely tied to performance in earlier grades. Although ninth-grade students may be more at risk of falling off-track, a student's on- or off-track status in ninth grade is probably related to their status in the seventh and eighth grades. If a teacher or counselor notices a student is falling off-track in seventh grade, interventions may be necessary to help the student get back on-track before entering high school.

Figure E1 illustrates how each grade-specific indicator increased the likelihood of college enrollment, accounting for the fact that earlier grade indicators predicted later grade indicators. For example, in looking at the Chicago indicator, the seventh-grade measure (orange line) increased enrollment by the largest amount, while the ninth-grade measure increased enrollment by the smallest amount (magenta line). In looking at the HERC and state indicators, the seventh-grade measures increased enrollment by the largest amounts, as well. While the role of the other grade-specific indicators varied by grade and type of potential indicator, a pattern emerged — the seventh-grade indicator had the strongest relationship with enrollment.

Figure E1. Total Effect of Potential Grade Level Indicators on College Enrollment



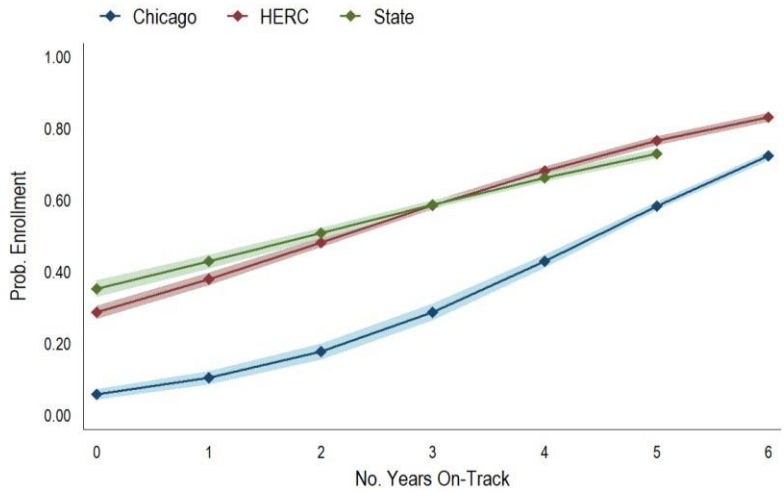
Source: HERC multi-year data.
Note: Sample was limited to 26,061 non-Native American students. Results came from structural equation models that predicted college enrollment; controlled for the Chicago, HERC, or state indicator in each grade; controlled for student background characteristics (age, gender, race/ethnicity, English learner, special education, economic disadvantage); and controlled for cohort fixed-effects (dummy variables for each cohort).

The intention is not to suggest there is something special about seventh grade — the strong role of seventh grade was likely due to the fact that it was the first grade examined and had cumulative effects on later grades. Practically, identifying students for early intervention might pay off more than identifying students for intervention in later grades. That said, the event history analyses suggested the ninth-grade transition was a challenging time, at least as defined by the Chicago and HERC indicators. Although identifying students for intervention as early as possible might have the greatest impact, additional assistance in the ninth grade might also be necessary since that appeared to be when students were at the greatest risk of falling off-track.

Appendix F. Additional Potential Indicator Analysis — Did Meeting a Potential Indicator Multiple Grades Increase College Enrollment?

While meeting the Chicago, HERC, and state indicators in specific grades was positively associated with college enrollment, meeting the potential indicators in multiple grades also increased enrollment. Figure F1 shows students who met the Chicago, HERC, and state indicators for additional years were more likely to enroll than students who met the potential indicators for fewer years. For example, 29 percent of students who never met the HERC indicator enrolled in college, while 83 percent of students who met the potential indicator in each grade — from seventh through 12th — did so. Interestingly, trend lines for the HERC and state indicators lay higher than the trend line for the Chicago indicator. This might reflect that the potential indicator was designed to predict high school graduation, not college enrollment; students might meet the potential indicator, but have little interest or preparation to go on to college.

Figure F1. College Enrollment by the Number of Years On-Track



Source: HERC multi-year data.
Note: Sample was limited to 11,425 non-Native American students who were not missing data. Results came from logistic regression models that predicted college enrollment; controlled for the number of years a student met the Chicago, HERC, or state indicator; controlled for student background characteristics (age, gender, race/ethnicity, English learner, special education, economic disadvantage); and controlled for cohort and school fixed-effects (dummy variables for each cohort, dummy variables for each school attended in seventh grade).

Appendix G. Research Question 2 Statistical Models

Falling off-track was modeled using discrete-time event history analyses, a class of models used to understand whether and when events occur. Specifically, the following logistic regression was estimated:

$$\text{logit}(h_{isg}) = \Gamma_g + \boldsymbol{X}_{is}\Phi + \boldsymbol{Z}_{isg}\Psi + \Delta_c + \Lambda_s$$

where h_{ist} was the hazard function of falling off-track for student *i* in school *s* in grade *g* (conditional on being on-track). The logit of the hazard function was modeled as a function of grade fixed-effects Γ_g, time-invariant student characteristics $\boldsymbol{X}_{is}$, time-variant student characteristics $\boldsymbol{Z}_{isg}$, cohort fixed-effects Δ_c, and seventh-grade school fixed-effects Λ_s. The models used robust standard errors.

Appendix H. Research Question 2 Summary Statistics

Variable	Mean	SD
Did Not Meet the Chicago Indicator at Least Once	0.51	(0.50)
Did Not Meet the HERC Indicator at Least Once	0.77	(0.42)
Did Not Meet the State Indicator at Least Once	0.65	(0.48)
Number of Total Credits Earned in Sixth Grade	14.00	(2.98)
Number of Core Courses Failed in Sixth Grade (semesters)	0.60	(1.43)
Attendance Rate in Sixth Grade (in 10s)	9.66	(0.41)
Average Grade Percentage in Sixth Grade (in 10s)	8.41	(0.66)
Number of Advanced Credits Earned in Sixth Grade	1.73	(1.57)
English/Language Arts Test Score in Sixth Grade (standardized)	-0.21	(0.99)
Mathematics Test Score in Sixth Grade (standardized)	-0.23	(0.94)
Age	12.44	(0.68)
Female	0.52	(0.50)
Black	0.26	(0.44)
Hispanic	0.62	(0.49)
Asian	0.03	(0.17)
English Learner	0.13	(0.34)
Special Education	0.04	(0.19)
Economically Disadvantaged	0.79	(0.41)
Number of In-School Suspensions	0.52	(1.31)
Number of Out-of-School Suspensions	0.43	(1.10)
N (students)	17,879	

Source: HERC multi-year data.

Notes: Sample was limited to non-Native American students with non-missing data.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests)

Appendix I. Research Question 2 Regression Results

Table I1. Log-Odds from a Discrete-Time Hazard Model of Not Meeting the Chicago On-Track Indicator

Variable	Coef.	Sig.
Grade (ref. = Seventh)		
Eighth	-0.46	***
Ninth	0.24	***
10th	0.08	+
11th	0.01	
12th	-0.85	***
Number of Total Credits Earned in Sixth Grade	0.01	
Number of Core Courses Failed in Sixth Grade (semesters)	0.05	**
Attendance Rate in Sixth Grade (in 10s)	-0.30	***
Average Grade Percentage in Sixth Grade (in 10s)	-1.30	***
Number of Advanced Credits Earned in Sixth Grade	0.00	
English/Language Arts Test Score in Sixth Grade (standardized)	-0.06	**
Mathematics Test Score in Sixth Grade (standardized)	-0.21	***
Age	0.12	***
Female	-0.13	***
Race/Ethnicity (ref. = White)		
Black	-0.15	*
Hispanic	0.32	***
Asian	-0.14	
English Learner	-0.17	***
Special Education	-0.37	***
Economically Disadvantaged	0.14	***
Number of In-School Suspensions	0.24	***
Number of Out-of-School Suspensions	0.31	***
Log-likelihood	-20,542	
Pseudo R-squared	0.22	

Source: HERC multi-year data.

Note: Sample was limited to 17,879 non-Native American students (63,352 student-grade observations) with non-missing data. All models included cohort and seventh-grade school fixed-effects and used robust standard errors.

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests)

Table I2. Log-Odds from a Discrete-Time Hazard Model of Not Meeting the HERC On-Track Indicator

Variable	Coef.	Sig.
Grade (ref. = Seventh)		
Eighth	-0.47	***
Ninth	1.34	***
10th	0.42	***
11th	0.28	***
12th	0.48	***
Number of Total Credits Earned in Sixth Grade	0.04	***
Number of Core Courses Failed in Sixth Grade (semesters)	0.02	
Attendance Rate in Sixth Grade (in 10s)	-0.79	***
Average Grade Percentage in Sixth Grade (in 10s)	-1.39	***
Number of Advanced Credits Earned in Sixth Grade	-0.22	***
English/Language Arts Test Score in Sixth Grade (standardized)	-0.21	***
Mathematics Test Score in Sixth Grade (standardized)	-0.38	***
Age	0.22	***
Female	0.03	
Race/Ethnicity (ref. = White)		
Black	-0.15	*
Hispanic	0.05	
Asian	-0.16	
English Learner	0.79	***
Special Education	0.16	*
Economically Disadvantaged	0.06	
Number of In-School Suspensions	0.27	***
Number of Out-of-School Suspensions	0.42	***
Log-likelihood	-18,522	
Pseudo R-squared	0.34	

Source: HERC multi-year data.

Note: Sample was limited to 17,879 non-Native American students (46,228 student-grade observations) with non-missing data. All models included cohort and seventh-grade school fixed-effects and used robust standard errors.

Table I3. Log-Odds from a Discrete-Time Hazard Model of Not Meeting the State On-Track Indicator

Variable	Coef.	Sig.
Grade (ref. = Seventh)		
Eighth	-0.60	***
Ninth	-1.81	***
10th	-1.91	***
11th	-1.47	***
Number of Total Credits Earned in Sixth Grade	0.12	*
Number of Core Courses Failed in Sixth Grade (semesters)	-0.08	***
Attendance Rate in Sixth Grade (in 10s)	-0.16	***
Average Grade Percentage in Sixth Grade (in 10s)	-0.59	***
Number of Advanced Credits Earned in Sixth Grade	0.04	***
English/Language Arts Test Score in Sixth Grade (standardized)	0.27	***
Mathematics Test Score in Sixth Grade (standardized)	-1.31	***
Age	-0.79	***
Female	0.24	***
Race/Ethnicity (ref. = White)		
Black	0.37	***
Hispanic	0.12	
Asian	-0.17	
English Learner	0.50	***
Special Education	0.44	***
Economically Disadvantaged	0.05	
Number of In-School Suspensions	0.06	**
Number of Out-of-School Suspensions	0.16	***
Log-likelihood	-12,758	
Pseudo R-squared	0.48	

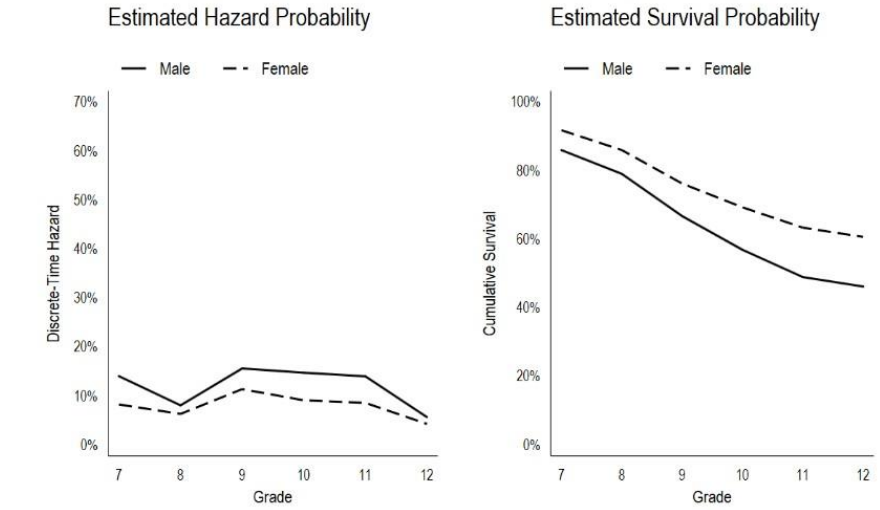
Source: HERC multi-year data.

Note: Sample was limited to 17,879 non-Native American students (41,936 student-grade observations) with non-missing data. All models included cohort and seventh-grade school fixed-effects and used robust standard errors.

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests)

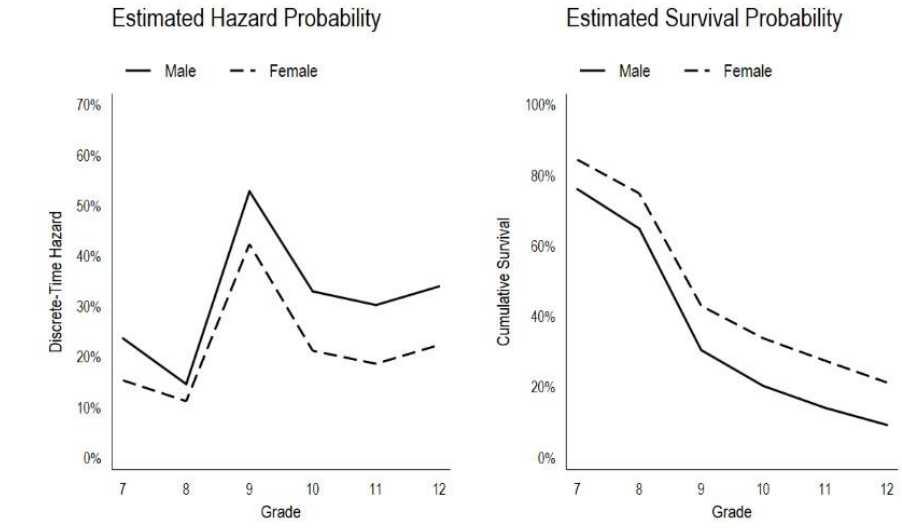
Appendix J. R Hazard and Survival Functions of Falling Off-Track by Gender, Race, and SES

Figure J1. Falling Off-Track Based on the Chicago Indicator, by Gender



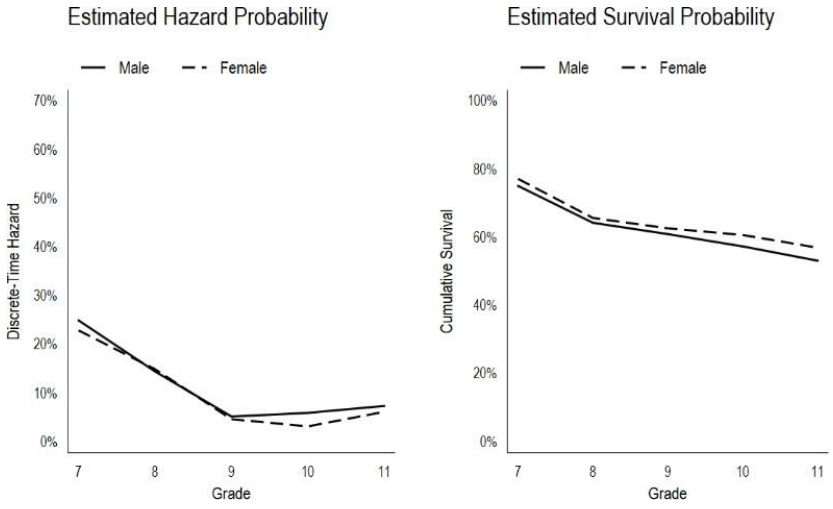
Source: HERC multi-year data.
Note: Sample was limited to 17,879 non-Native American students with non-missing data. Results came from discrete-time hazard models with control variables, cohort fixed-effects, and seventh-grade school fixed-effects.

Figure J2. Falling Off-Track Based on the HERC Indicator, by Gender



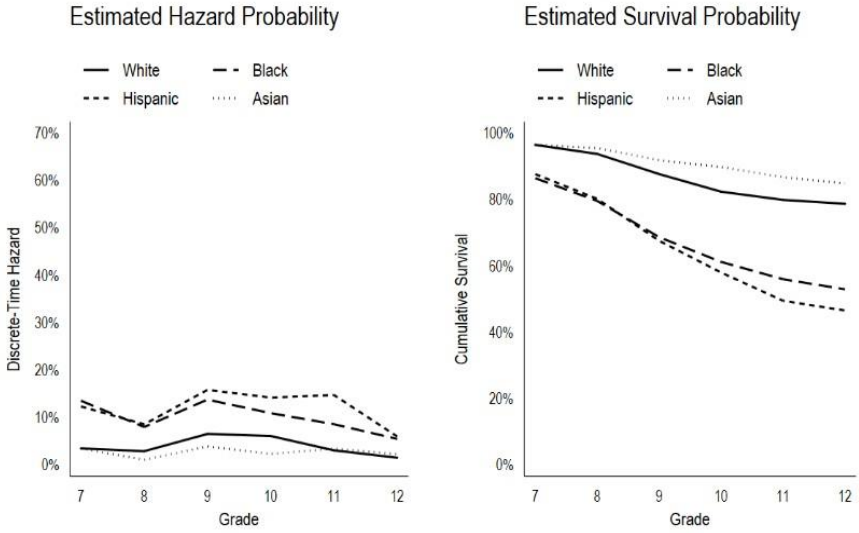
Source: HERC multi-year data.
Note: Sample was limited to 17,879 non-Native American students with non-missing data. Results came from discrete-time hazard models with control variables, cohort fixed-effects, and seventh-grade school fixed-effects.

Figure J3. Falling Off-Track Based on the State Indicator, by Gender



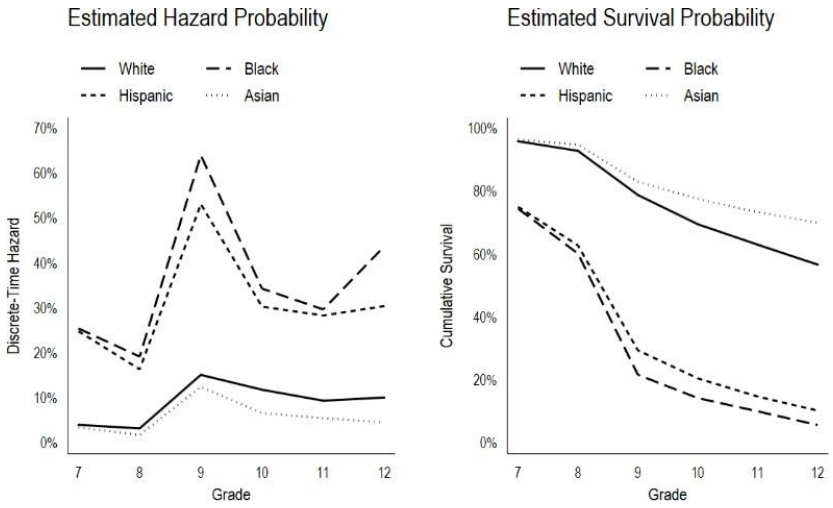
Source: HERC multi-year data.
Note: Sample was limited to 17,879 non-Native American students with non-missing data. Results came from discrete-time hazard models with control variables, cohort fixed-effects, and seventh-grade school fixed-effects.

Figure J4. Falling Off-Track Based on the Chicago Indicator, by Race



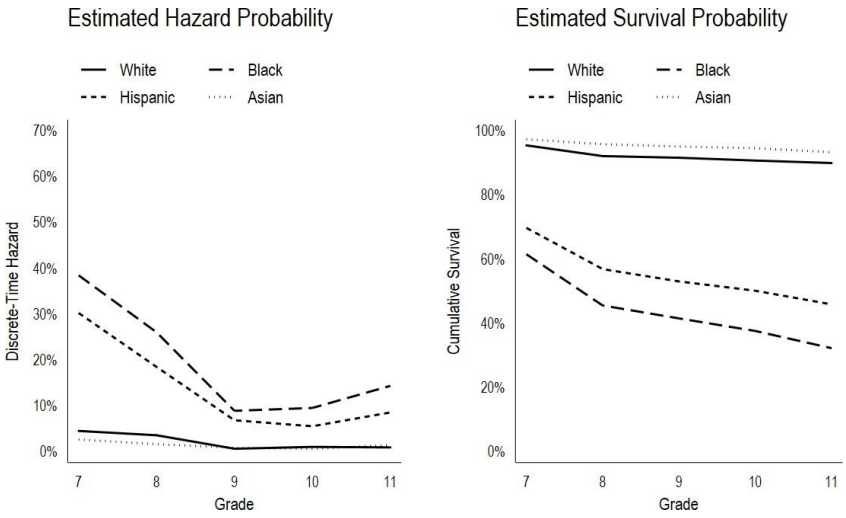
Source: HERC multi-year data.
Note: Sample was limited to 17,879 non-Native American students with non-missing data. Results came from discrete-time hazard models with control variables, cohort fixed-effects, and seventh-grade school fixed-effects.

Figure J5. Falling Off-Track Based on the HERC Indicator, by Race



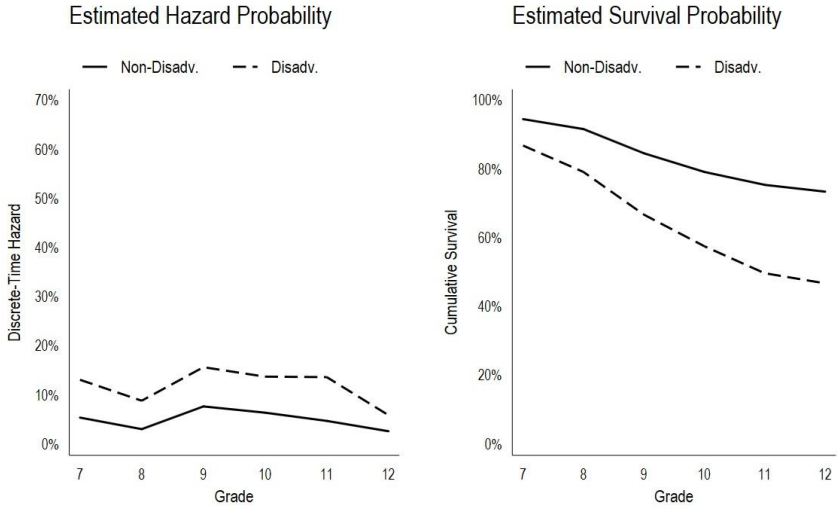
Source: HERC multi-year data.
Note: Sample was limited to 17,879 non-Native American students with non-missing data. Results came from discrete-time hazard models with control variables, cohort fixed-effects, and seventh-grade school fixed-effects.

Figure J6. Falling Off-Track Based on the State Indicator, by Race



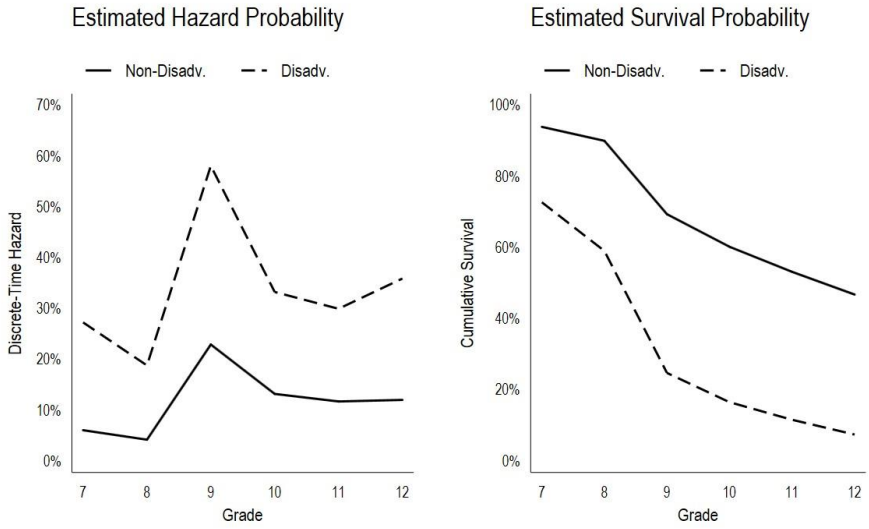
Source: HERC multi-year data.
Note: Sample was limited to 17,879 non-Native American students with non-missing data. Results came from discrete-time hazard models with control variables, cohort fixed-effects, and seventh-grade school fixed-effects.

Figure J7. Falling Off-Track Based on the Chicago Indicator, by SES



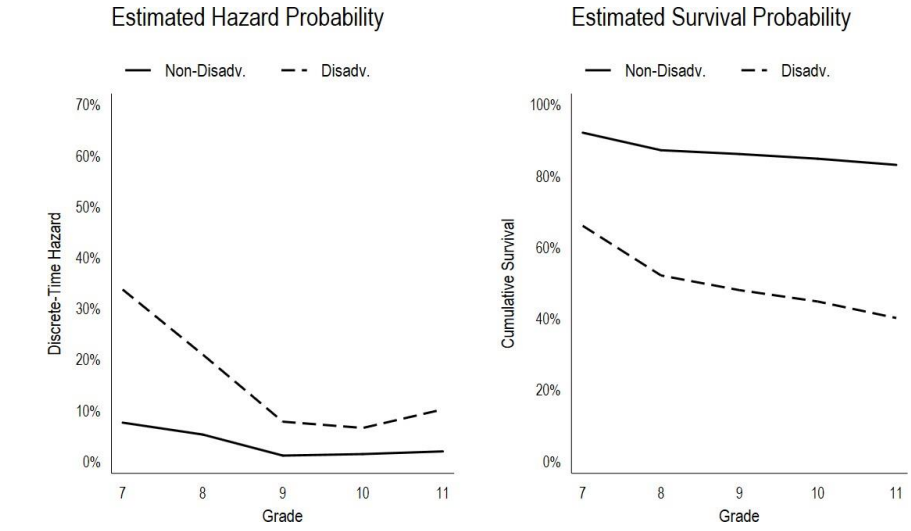
Source: HERC multi-year data.
Note: Sample was limited to 17,879 non-Native American students with non-missing data. Results came from discrete-time hazard models with control variables, cohort fixed-effects, and seventh-grade school fixed-effects.

Figure J8. Falling Off-Track Based on the HERC Indicator, by SES



Source: HERC multi-year data.
Note: Sample was limited to 17,879 non-Native American students with non-missing data. Results came from discrete-time hazard models with control variables, cohort fixed-effects, and seventh-grade school fixed-effects.

Figure J9. Falling Off-Track Based on the State Indicator, by SES



Source: HERC multi-year data.
Note: Sample was limited to 17,879 non-Native American students with non-missing data. Results came from discrete-time hazard models with control variables, cohort fixed-effects, and seventh-grade school fixed-effects.

About HERC. Focusing on the most pressing challenges facing the region, the Houston Education Research Consortium (HERC) is a research-practice partnership between Rice University and 11 Houston-area school districts. HERC research is developed 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

Houston Education Research Consortium
a program of the Kinder Institute for Urban Research
MS-258 Rice University | Houston, Texas 77005
713-348-2532 | herc@rice.edu
Find us online: kinder.rice.edu/herc